



HAL
open science

Contribution à l'identification récursive des systèmes par l'approche des sous-espaces

Guillaume Mercère

► **To cite this version:**

Guillaume Mercère. Contribution à l'identification récursive des systèmes par l'approche des sous-espaces. Automatique / Robotique. Université des Sciences et Technologie de Lille - Lille I, 2004. Français. NNT: . tel-00007958

HAL Id: tel-00007958

<https://theses.hal.science/tel-00007958>

Submitted on 7 Jan 2005

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

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

UNIVERSITÉ DES SCIENCES ET TECHNOLOGIES DE LILLE

Numéro attribué par la bibliothèque : 3546

THÈSE

présentée à

l'Université des Sciences et Technologies de LILLE

en vue de l'obtention du grade de

DOCTEUR D'UNIVERSITÉ

Spécialité

Automatique et Informatique Industrielle

par

Guillaume MERCÈRE

**CONTRIBUTION À L'IDENTIFICATION RÉCURSIVE DES
SYSTÈMES PAR L'APPROCHE DES SOUS-ESPACES**

Soutenue publiquement le 07 décembre 2004 à 10h30

JURY

| | | |
|----------------------|---------------|--|
| Rapporteurs : | A. RICHARD | Professeur des Universités Université Henri Poincaré Nancy I |
| | F. ROTELLA | Professeur des Universités École Nationale d'Ingénieurs de Tarbes |
| Examineurs : | M. LOVERA | Professeur associé École Polytechnique de Milan (Italie) |
| | J. P. RICHARD | Professeur des Universités École Centrale de Lille |
| Directeur de thèse : | C. VASSEUR | Professeur des Universités Université des Sciences et Technologies de Lille |
| Co-encadrant : | S. LECŒUCHE | Enseignant chercheur École Nationale Supérieure des Mines de Douai |

A mes parents et à Miss...

Remerciements

Je tiens tout d'abord à remercier les membres du jury qui m'ont fait l'honneur d'examiner ce travail de thèse.

Je remercie M. Alain RICHARD, professeur de l'Université Henri Poincaré de Nancy, pour avoir accepté d'être rapporteur de ma thèse et avoir étudié avec minutie ce travail. Ses commentaires constructifs et ses explications sur la théorie de l'estimation m'ont permis de mieux appréhender certaines notions qui étaient obscures jusqu'alors.

Je tiens à exprimer ma gratitude à M. Frédéric ROTELLA, professeur de l'École Nationale d'Ingénieurs de Tarbes, d'avoir bien voulu évaluer mon manuscrit de thèse. Je le remercie pour ses remarques avisées sur l'ensemble de ce document.

J'adresse mes vifs remerciements à M. Marco LOVERA, professeur associé de l'Université polytechnique de Milan, d'avoir eu la gentillesse d'être examinateur de ma thèse. J'ai eu la chance qu'il s'intéresse à mon travail lorsque je lui ai demandé par email son manuscrit de thèse en 2002. J'ai ensuite pu le rencontrer lors de la conférence SYSID de 2003 à Rotterdam. De cette rencontre est née une collaboration importante qui a finalement débouché sur un article dans la revue International Journal of Adaptive Control and Signal Processing. Encore merci d'avoir accepté de travailler avec un modeste doctorant en identification. J'espère que cette coopération durera de nombreuses années. Je n'ai en effet pas pu encore visiter entièrement Milan !-).

Je suis également très reconnaissant envers M. Jean-Pierre RICHARD, professeur de l'École Centrale de Lille, pour l'intérêt qu'il a porté à mon travail et pour l'avoir examiné avec attention.

Je tiens à remercier M. Christian VASSEUR, professeur de l'Université des Sciences et Technologies de Lille et ancien directeur du laboratoire I^3D , pour avoir dirigé ce travail de thèse. Il a accepté de m'encadrer et m'a de suite fait confiance alors que je n'étais pas originaire de son université. Malgré la centaine de kilomètres qui nous sépare, travailler sous sa direction a été une expérience enrichissante.

Je veux tout particulièrement exprimer ma gratitude à Stéphane LECŒUCHE pour avoir co-encadré cette thèse et m'avoir permis de travailler à l'EIPC. Bien que depuis deux ans permanent de l'École d'Ingénieurs de Mines de Douai, il a toujours été présent lorsque j'avais besoin de lui. Il m'a guidé tout au long de ces trois années et m'a fait gravir avec succès toutes les étapes qui jalonnent les trois années de thèse. Je le remercie également de ne pas avoir ménagé ses efforts pour étudier les méthodes des sous-espaces. Les nombreuses questions qu'il m'a posées ont maintes fois donné lieu à d'importantes améliorations des algorithmes développés. Encore merci pour ta confiance et ta sympathie. J'aurais également une petite pensée pour Victor et Anne Gaëlle qui n'ont pas pu voir leur papa aussi souvent qu'ils le souhaitaient à cause de moi.

Je tiens à présent à remercier les permanents de l'EIPC qui m'ont permis de réaliser cette thèse dans les meilleures conditions. Je remercie tout d'abord M. Emmanuel LECLERCQ de m'avoir accepté en tant qu'enseignant chercheur à l'EIPC, tout d'abord en CDD puis plus récemment en CDI. Merci à Patrick pour ces trois années de co-voiturage et de discussions quotidiennes. Merci aussi à Fred, Guillaume, Nicolas D., Bernard, Olivier, Benoît et Audrey pour leur aide en informatique, en mécanique, en électronique et surtout pour leur bonne humeur et leurs blagues ou jeux de mots si célèbres à l'EIPC. Je tiens également à remercier Nicolas V. pour ses connaissances « latexiennes ». Je remercie enfin tous les membres de l'EIPC que je n'ai

pas encore cités pour leurs encouragements et les services qu'ils m'ont rendus avec gentillesse tout au long de ces trois années.

Je n'oublierai pas Sylvain LALOT, mon cher « vieux singe », avec qui nous avons passé, Stéphane et moi, de nombreuses heures à nous « chamailler » sur des points particuliers d'articles et d'expériences réalisés depuis trois ans.

Au terme de ces trois années, je pense bien évidemment à ma famille et plus particulièrement à mes parents sans qui rien de tout cela n'aurait pu arriver. Merci pour vos sacrifices et votre soutien.

Mes plus belles pensées seront destinées à Sandrine qui a accepté d'être mon épouse il y a quelques mois. Tu as permis, par ta joie de vivre et ton réconfort, de rendre ces trois années de thèse et de vie commune les plus délicieuses possibles.

Table des matières

| | | |
|----------|--|-----------|
| 1 | Introduction | 1 |
| 1.1 | Motivations | 1 |
| 1.2 | Contributions et organisation du document | 3 |
| 2 | Méthodes classiques d'identification récursive | 7 |
| 2.1 | Motivations | 7 |
| 2.2 | Problématique et notations | 8 |
| 2.3 | Méthodes classiques d'identification récursive | 11 |
| 2.3.1 | Moindres carrés récursifs | 11 |
| 2.3.2 | Variable instrumentale récursive | 18 |
| 2.4 | Limitations des méthodes classiques et conclusion | 21 |
| 3 | Méthodes d'identification des sous-espaces : une vue d'ensemble des algorithmes MOESP | 23 |
| 3.1 | Problématique | 23 |
| 3.2 | Bref historique | 25 |
| 3.3 | « MIMO Output Error State sPace model identification » | 29 |
| 3.3.1 | Algorithme Ordinary MOESP | 30 |
| 3.3.2 | Algorithmes PI et PO MOESP | 37 |
| 3.3.3 | Algorithme MOESP pour les modèles avec « erreurs en les variables » | 42 |
| 3.3.4 | Estimation des matrices d'état | 43 |
| 3.3.5 | Propriétés asymptotiques des algorithmes MOESP et choix des paramètres structuraux | 47 |
| 3.4 | Variable instrumentale et théorèmes unifiés | 53 |
| 3.5 | Conclusion | 59 |
| 4 | Méthodes récursives d'identification des sous-espaces fondées sur le critère de Yang | 61 |
| 4.1 | Motivations | 61 |
| 4.2 | Problématique et méthodes du traitement d'antennes | 63 |
| 4.2.1 | Modèle d'antenne et modélisation des signaux | 64 |
| 4.2.2 | MULTiple SIGNAL Classification (MUSIC) | 67 |
| 4.2.3 | Critère de Yang et algorithmes dérivés | 68 |
| 4.2.4 | Méthode du propagateur | 76 |
| 4.2.5 | Conclusion | 78 |
| 4.3 | Adaptation des algorithmes fondés sur le critère de Yang au problème d'identification récursive des sous-espaces | 78 |

| | | |
|----------|---|------------|
| 4.3.1 | Algorithmes développés par T. Gustafsson | 79 |
| 4.3.2 | Algorithmes développés par M. Lovera | 82 |
| 4.3.3 | Adaptation des algorithmes COPAST et COPASTd | 85 |
| 4.3.4 | Algorithmes développés par H. Oku | 88 |
| 4.3.5 | Algorithme développé par M. Lovera (suite) | 90 |
| 4.3.6 | Synthèse | 91 |
| 4.4 | Conclusion | 92 |
| 5 | Adaptation de la méthode du propagateur à l'identification récursive des sous-espaces | 93 |
| 5.1 | Problématique | 93 |
| 5.2 | Analogie entre l'identification des sous-espaces et le traitement d'antennes . . . | 95 |
| 5.3 | Mise à jour du vecteur d'observation | 97 |
| 5.3.1 | Dans un contexte déterministe | 100 |
| 5.3.2 | Dans un contexte stochastique | 104 |
| 5.3.3 | Complexité numérique et fiabilité des algorithmes de mise à jour du vecteur d'observation | 107 |
| 5.3.4 | Conclusion | 108 |
| 5.4 | Estimation récursive d'une base de l'espace d'observabilité | 109 |
| 5.4.1 | Adaptation de la méthode du propagateur au problème d'identification . | 109 |
| 5.4.2 | Estimation à partir d'un vecteur d'observation épuré des effets des perturbations | 113 |
| 5.4.3 | Introduction d'une variable instrumentale | 116 |
| 5.4.4 | Synthèse | 124 |
| 5.4.5 | Propriétés asymptotiques de la méthode du propagateur | 125 |
| 5.4.6 | Complexité numérique des algorithmes d'estimation de la matrice d'observabilité | 132 |
| 5.5 | Complexité numérique des algorithmes d'identification récursive : une étude de cas | 134 |
| 5.6 | Estimation des matrices d'état | 136 |
| 5.7 | Conclusion | 138 |
| 6 | Applications | 139 |
| 6.1 | Exemples de simulation | 140 |
| 6.1.1 | Étude d'un système invariant. Comparaison des algorithmes | 140 |
| 6.1.2 | Étude d'un système variant | 153 |
| 6.1.3 | Conclusion | 162 |
| 6.2 | Étude de systèmes réels | 162 |
| 6.2.1 | Étude d'un système de transmission flexible | 162 |
| 6.2.2 | Étude d'un système électronique | 169 |
| 6.3 | Conclusion | 180 |
| 7 | Conclusion et perspectives | 181 |

| | |
|---|------------|
| A Preuves | 187 |
| A.1 Preuve de la formule (3.133) | 187 |
| A.2 Preuve de la formule (3.154) | 188 |
| A.3 Preuve de la formule (4.96) | 189 |
| A.4 Preuve de la proposition 5.1 | 191 |
| A.5 Preuve de l'algorithme (5.81) | 198 |
| A.6 Preuve de l'algorithme (5.122) | 200 |
| A.7 Preuve de l'algorithme (5.138) | 202 |
| B Étude du système de transmission flexible | 207 |
| B.1 Étude du moteur à courant continu | 207 |
| B.2 Étude mécanique du système de transmission flexible | 209 |
| C Abréviations, acronymes et notations | 211 |
| C.1 Abréviations et acronymes | 211 |
| C.2 Symboles et notations | 212 |
| D Bibliographie personnelle | 215 |
| Bibliographie | 217 |

Chapitre 1

Introduction

Sommaire

| | |
|--|----------|
| 1.1 Motivations | 1 |
| 1.2 Contributions et organisation du document | 3 |

1.1 Motivations

La construction et l'utilisation de modèles constituent, de nos jours, des étapes incontournables de nombreuses disciplines scientifiques et technologiques (physique, chimie, biologie, économie...). La modélisation permet en effet de formaliser le comportement du processus étudié à l'aide d'une représentation, baptisée « modèle », à partir de laquelle il est possible de comprendre, commander ou améliorer le fonctionnement du procédé analysé. Deux approches, souvent complémentaires, peuvent être envisagées pour modéliser un système :

- La première, qui fait appel aux connaissances des spécialistes du domaine considéré, consiste à regrouper, généralement sous forme de systèmes différentiels, algébriques ou graphiques, les lois et relations de la physique qui décrivent le comportement du processus.
- Dans de nombreuses situations pratiques, la complexité du procédé à modéliser est telle qu'il est difficile, voire impossible, de connaître ou d'associer les lois physiques gouvernant la dynamique du système. Il est alors nécessaire de faire appel à un modèle « boîte noire¹ », construit à partir de mesures des données d'entrée-sortie du procédé. L'objectif majeur de ce modèle est de reproduire au mieux le comportement du processus identifié. On parle plus particulièrement, dans ce cas, « d'identification du système ». Le modèle de représentation ainsi obtenu est généralement constitué d'un ensemble de relations mathématiques liant les variables du système analysé. Les paramètres de ce type de modèle sont, la plupart du temps, difficilement interprétables.

La modélisation est souvent considérée comme la phase initiale de l'automatique moderne. Elle permet en effet de déterminer le modèle du procédé qui sera utilisé pour prédire son comportement futur, diagnostiquer ses évolutions de fonctionnement ou synthétiser son régulateur

¹Par opposition à « boîte blanche » lorsque les lois de la physique sont connues et exclusivement utilisées pour la construction du modèle.

et sa loi de commande. Puisque les processus industriels sont généralement complexes, il est nécessaire de posséder des connaissances approfondies dans des domaines aussi variés que la mécanique, l'électronique, l'énergétique, la biologie ou la chimie pour obtenir un modèle boîte blanche efficace. La modélisation expérimentale ou identification des systèmes [SS89, BDR⁺92, Lju99] est donc devenue une alternative logique à la modélisation classique fondée sur l'exploitation des lois physiques.

Dans de nombreuses situations, la complexité croissante des procédés a pour conséquence de voir les propriétés dynamiques du système évoluer, ne serait-ce que lentement, au cours du temps. Le comportement d'une automobile dépend ainsi de sa vitesse, de la quantité d'essence transportée, de variables environnementales telles que le vent ou la qualité de la chaussée. Les caractéristiques physiques et dynamiques d'un système tel qu'un chauffe-eau sont liées à des facteurs tels que la température extérieure, l'entartrage de la résistance chauffante et du ballon, le nombre de personnes utilisant, au même instant, de l'eau chaude... Il est alors très intéressant, voire nécessaire, de considérer un modèle de procédé lentement variable afin de tenir compte d'évolutions non prédictibles. Les algorithmes d'identification récursive² sont particulièrement adaptés à ce type de problématique. De nombreuses méthodes adaptatives ont été proposées depuis les années 70 [LS83, You84, Lju99] pour fournir des modèles dont les paramètres varient relativement aux changements de processus. La majorité d'entre elles reposent sur des représentations de type entrée-sortie. Ces dernières se sont initialement développées en raison de la disponibilité d'algorithmes hors ligne fondés sur ce type d'écriture [SS89]. Bien que particulièrement efficace pour des systèmes monovariables, l'identification récursive d'un modèle entrée-sortie se révèle être plus difficile lorsque le procédé comporte plusieurs entrées et plusieurs sorties. Elle demande en effet de connaître *a priori* l'ordre des polynômes composant les matrices polynomiales du modèle. De plus, lorsque le nombre d'entrées et de sorties devient conséquent, la gestion de matrices de fonctions de transfert en est que plus délicate. En identification hors ligne, ces difficultés ont motivé le développement de techniques estimant directement un modèle d'état du système sans phase transitoire d'identification de représentations d'entrée-sortie : les méthodes des sous-espaces [Vib95, VD96b, BSR98, BRS98]. Le premier avantage de l'approche d'état repose sur le fait qu'elle gère, de manière équivalente, les systèmes mono et multivariables³. Elle ne nécessite aucune forme canonique et est basée sur un seul indice structurel : l'ordre du système. De par son écriture, elle permet d'analyser la stabilité, l'observabilité et la commandabilité du processus à identifier. Elle est enfin particulièrement bien adaptée à la théorie de la commande [AM90, ZD98]. Ces qualités ont rapidement conduit les chercheurs à utiliser la représentation d'état pour des procédés requérant une modélisation en ligne. Puisque les méthodes des sous-espaces ont rapidement été considérées, en identification hors ligne, comme une alternative logique aux techniques d'erreur de prédiction⁴, il a été tenté, dès la fin des années 80 [MDVV89], de les appliquer sur des procédés variants. Malheureusement, ces premiers essais ont rapidement été vains puisque, par construction, les méthodes hors ligne des sous-espaces emploient des outils d'algèbre linéaire tels que la décomposition en valeurs propres ou singulières qui ne sont pas appropriés aux applications temps réel du fait de leur complexité numérique. La fin des années 90 et le début des années 2000 ont vu

²On parle également « d'identification temps réel » ou « d'algorithmes adaptatifs ».

³Seules les dimensions des matrices d'état du modèle sont modifiées.

⁴L. Ljung a écrit en 1996 que « The subspace methods for state space models are, in my mind, the most interesting development in the past decade » [Lju96].

l'apparition de nouvelles approches évitant l'utilisation de ces lourds outils de décomposition tout en étant capables de traiter des systèmes perturbés à la fois par des bruits de mesure et de procédé [Lov98, Gus99, Oku00, Lov03]. Ces algorithmes récursifs ont en commun d'être fondés sur l'adaptation, au problème d'identification, d'un critère particulier de traitement d'antennes [KV96] : le critère de Yang [Yan95b]. Ces méthodes, qui se sont avérées efficaces dans le domaine de la surveillance de systèmes [ONVV01, Oku03], présentent néanmoins trois inconvénients majeurs :

- ils reposent tous sur le même critère ;
- ils nécessitent des approximations pour rendre le critère de Yang quadratique afin de faciliter sa minimisation [Gus97a, Lov98, LGV00] ;
- ils utilisent des algorithmes itératifs de type gradient et de fortes contraintes sur les estimées pour minimiser le critère d'ordre 4 originel de Yang [OK99, OK02, Lov03].

L'objectif de cette thèse consiste à développer des algorithmes d'identification récursive originaux fondés sur de nouveaux critères ayant comme propriétés d'être *quadratiques, sans approximation et sans contrainte*. Ces fonctions coût sont obtenues en s'inspirant de la méthode du propagateur [MD91] appliquée initialement en traitement d'antennes. Son principe repose sur l'utilisation d'un opérateur linéaire particulier et la décomposition de l'espace d'observation en deux sous-espaces complémentaires : le sous-espace signal et le sous-espace bruit. Son avantage majeur réside dans le fait qu'aucune recherche d'éléments propres de la matrice interspectrale des signaux reçus par l'antenne n'est réalisée. Son adaptation au problème d'identification s'appuie sur une hypothèse classique : le système étudié est observable. Nous verrons plus précisément que cette condition est nécessaire et suffisante à une estimation récursive et consistante d'une base de l'espace d'observabilité.

1.2 Contributions et organisation du document

Cette thèse a pour objectif de :

- proposer un état de l'art des méthodes d'identification récursive des sous-espaces ;
- développer de nouvelles techniques d'identification récursive des sous-espaces remédiant aux principaux défauts des algorithmes énoncés jusqu'à aujourd'hui.

De nombreuses méthodes d'identification récursive des sous-espaces ont en effet été exposées depuis le milieu des années 90. A notre connaissance, aucun article dressant un état de l'art de ces dernières n'a été proposé, à ce jour, dans la littérature. Celles-ci méritent pourtant d'être présentées dans un contexte unifié puisqu'elles résultent toutes d'adaptations d'algorithmes particuliers de traitement d'antennes. Bien que ces techniques aient démontré leur efficacité sur de nombreux exemples de simulation, il nous a semblé intéressant de développer des méthodes alternatives à l'ajustement du critère de Yang [Yan95b] au problème d'identification. Ce dernier souffre en effet d'un désagrément notable : son ordre 4. L'utilisation d'un critère quadratique sans approximation et sans contrainte est donc proposée au sein de ce manuscrit. Les performances et l'applicabilité des algorithmes qui en émanent sont mises en évidence sur des exemples de simulation et via des données de maquettes de laboratoire. Le plan d'étude de ce document est le suivant :

Chapitre 2 : État de l'art de l'identification récursive

Ce chapitre présente les méthodes classiques d'identification récursive fondées sur le principe des moindres carrés. Il permet au lecteur non coutumier de cette thématique d'appréhender

les principales notions liées à cette branche de la modélisation. Une attention particulière est apportée à la technique de la variable instrumentale. Cette méthode est en effet fondamentale pour la compréhension des algorithmes d'identification récursive des sous-espaces.

Chapitre 3 : Méthodes d'identification des sous-espaces : une vue d'ensemble des algorithmes MOESP

L'objectif de cette thèse consiste à présenter et développer des algorithmes d'identification récursive des sous-espaces. Il semble donc essentiel de proposer, dans un premier temps, une synthèse des méthodes hors ligne des sous-espaces. Les techniques regroupées sous l'acronyme MOESP peuvent être considérées comme le point de départ de tout algorithme récursif des sous-espaces. Une présentation des principales méthodes composant cette famille de techniques est donc exposée au sein de ce chapitre. Les fondements mathématiques et géométriques, les caractéristiques théoriques et les principales propriétés asymptotiques de cette classe de méthodes d'identification sont plus particulièrement commentés.

Chapitre 4 : Méthodes récursives d'identification des sous-espaces fondées sur le critère de Yang

Un état de l'art des techniques récursives des sous-espaces fondées sur le critère de Yang [Yan95b] est proposé. Une description de ces algorithmes dans leur contexte d'origine (le traitement d'antennes) et leur ajustement au problème d'identification sont plus précisément présentés. L'adaptation des méthodes COPAST et COPASTd en identification récursive des sous-espaces est également développée.

Chapitre 5 : Adaptation de la méthode du propagateur à l'identification récursive des sous-espaces

Ce chapitre contient les principales contributions de cette thèse. Ce dernier débute par une étude de l'analogie entre l'identification des sous-espaces et le traitement d'antennes. Cette analyse permet de justifier l'approche proposée jusqu'ici en identification récursive des sous-espaces i.e. adapter des techniques particulières de traitement d'antennes au problème d'estimation de la matrice d'observabilité du système. Deux étapes sont ainsi mises en évidence :

1. la mise à jour d'un vecteur nommé vecteur d'observation à chaque nouvelle acquisition ;
2. l'estimation d'une base de la matrice d'observabilité, à partir de ce vecteur, en adaptant des méthodes particulières de poursuite des sous-espaces.

Deux nouveaux algorithmes d'estimation du vecteur d'observation sont introduits et comparés aux techniques déjà existantes. Une analyse de la complexité numérique de ces méthodes est exposée. Le problème de l'estimation récursive d'une base de l'espace d'observabilité est ensuite considéré. Sept algorithmes sont ainsi développés. Ceux-ci reposent sur l'adaptation au problème d'identification de la méthode du propagateur [MD91]. Un examen des premières propriétés asymptotiques de ces techniques et une étude pratique et théorique du coût calculatoire de ces dernières sont également présentés. Une description des méthodes d'extraction des matrices d'état du procédé conclut ce chapitre.

Chapitre 6 : Applications

La mise en application des algorithmes d'identification récursive des sous-espaces fondés sur le propagateur est menée au sein de ce chapitre. Deux situations sont considérées :

- Une analyse des performances et des capacités de poursuite de ces méthodes est tout d'abord étudiée sur des exemples de simulation. Une comparaison avec celles basées sur l'approximation PAST est également proposée [Yan95b].

- L'utilisation de données réelles mesurées sur deux maquettes de laboratoire révèle, dans un second temps, les possibilités des techniques développées.

Chapitre 7 : Conclusion et perspectives

Ce dernier chapitre expose une synthèse des contributions de cette thèse et présente les travaux de recherche futurs en lien direct avec la problématique traitée dans ce manuscrit.

Chapitre 2

Méthodes classiques d'identification récurrente

Sommaire

| | |
|--|-----------|
| 2.1 Motivations | 7 |
| 2.2 Problématique et notations | 8 |
| 2.3 Méthodes classiques d'identification récurrente | 11 |
| 2.3.1 Moindres carrés récurrents | 11 |
| 2.3.2 Variable instrumentale récurrente | 18 |
| 2.4 Limitations des méthodes classiques et conclusion | 21 |

2.1 Motivations

Lors du 13^e congrès mondial IFAC¹ de San Francisco, L. Ljung a montré, à travers son « historique de l'identification des systèmes » [Lju96], que cette spécialité de l'automatique a connu un développement considérable durant les deux dernières décennies du XX^e siècle tant sur le plan pratique [HK86, ZVDL94, GSMR96, BPP96, LLM97, AVVD98, FDV⁺98, CR01] que théorique [LS83, SS89, VD96b, Lju99, JKDW01, LBB⁺01]. Cette progression s'explique par le fait que l'identification est une discipline commune à de nombreux domaines des sciences de l'ingénieur (statistique, économie, biologie, traitement du signal...). Cette diversité d'applications implique inévitablement une grande variété de problèmes posés ainsi qu'une multiplicité de solutions proposées.

Parmi l'ensemble des techniques d'identification développées depuis le début des années 60, les méthodes récurrentes ont rapidement vu le jour [ÅE71]. Leur principal objectif consiste à estimer, de manière consistante² et en ligne, le modèle du processus étudié. Les deux raisons fondamentales de ce développement sont que :

- ces techniques permettent de suivre les évolutions (lentes) de la dynamique du système étudié en adaptant continuellement le modèle de procédé. Il est alors possible de prendre

¹International Federation of Automatic Control.

²Nous considérons qu'un modèle est consistant si le vecteur de paramètres caractérisant ce dernier converge vers les « vrais » paramètres avec une probabilité de 1.

des décisions en temps réel en fonction des dernières informations acquises sur le processus ;

- ces méthodes sont, la plupart du temps, moins gourmandes en ressource mémoire et en temps de calcul que leurs homologues hors ligne tout en étant capables de fournir un modèle fiable du procédé.

De nombreux auteurs ont appliqué avec succès des algorithmes d'identification récursive sur des applications plus ou moins complexes (thermique [Che02, BSM03], génie des procédés [TBDM91, WKL03], aéronautique [KKKN01, UB03], robotique [ÖG02]...) dans des domaines aussi variés que la commande de procédés [ÅHHH93, ÅW95, LLM97, PC98, LT01], le filtrage [GS84, Shy89, SK94, GBT99, Gus00] ou la surveillance de systèmes [Abi88, Ger98, Tia01, DBL⁺01, Öst02, Oku03]. Des ouvrages de référence [LS83, BDR⁺92, Lju99, LBB⁺01] présentent de manière détaillée les tenants et aboutissants de cette discipline. Notre objectif au sein de ce chapitre n'est donc pas de décrire (si cela était possible) l'ensemble des méthodes récursives d'identification mais d'exposer les techniques qui seront utiles à la compréhension des algorithmes récursifs des sous-espaces développés dans les prochains chapitres de ce manuscrit.

2.2 Problématique et notations

Lorsque nous interagissons avec un système, il est utile, voire nécessaire, de connaître les relations entre les variables observées. Ces dernières peuvent être divisées en trois catégories [Lju99] :

- les entrées de commande u qui permettent d'influer sur le processus ;
- les perturbations n , qui agissent également sur le système, mais qui sont en général non contrôlables par l'utilisateur ;
- les sorties y *i.e.* des variables mesurables qui caractérisent l'action du procédé sur son environnement.

Puisque les perturbations sont, par définition, non accessibles à la mesure, considérons un ensemble de données d'entrée-sortie discrètes sur un intervalle de temps $[1, M]$:

$$\mathcal{U} = \{u(t)\}_{t \in [1, M]} \quad (2.1)$$

$$\mathcal{Y} = \{y(t)\}_{t \in [1, M]} \quad (2.2)$$

$$\mathcal{Z} = \{\mathcal{U}, \mathcal{Y}\}. \quad (2.3)$$

Nous recherchons un modèle du procédé étudié sous forme d'une relation liant les entrées et sorties accessibles à l'instant t . Cette relation peut être représentée par :

$$y_m(t) = g(\mathcal{U}, \mathcal{Y}) \quad (2.4)$$

pour laquelle $y_m(t)$ est la sortie du modèle au t^e pas d'échantillonnage et g une fonction traduisant le lien implicite existant entre les données. Puisque nous ne considérons, dans la suite de ce manuscrit, que des modèles paramétriques, l'expression précédente peut être ajustée de la façon suivante :

$$g(\mathcal{U}, \mathcal{Y}) = g(\boldsymbol{\phi}(t), \boldsymbol{\eta}) \quad (2.5)$$

où $\boldsymbol{\eta}$ est un vecteur de paramètres de dimension finie et $\boldsymbol{\phi}(t)$ un vecteur de régression permettant de sélectionner les observations utiles à la description du modèle parmi les données \mathcal{U} et \mathcal{Y} . Cette relation, généralement non linéaire entre paramètres et vecteur de régression, peut être modifiée selon la paramétrisation du modèle choisi. Ainsi, dans le cas linéaire, le modèle peut être donné sous la forme suivante (cas des modèles ARX par exemple) [Lju99] :

$$y_m(t) = \boldsymbol{\phi}^T(t) \boldsymbol{\eta}. \quad (2.6)$$

On parle alors de régression linéaire puisque le modèle est linéaire en les paramètres. Cette modélisation étant parfois trop restrictive, il est possible de considérer le vecteur de régression comme dépendant des paramètres du système :

$$y_m(t) = \boldsymbol{\phi}^T(t, \boldsymbol{\eta}) \boldsymbol{\eta}. \quad (2.7)$$

Il s'agit alors d'une forme pseudo linéaire (cas des modèles ARMAX [Lju99]). Ces différences de notations nous obligent à étudier séparément les procédés modélisables par des régressions linéaires de ceux représentés par des modèles pseudo linéaires ou non linéaires. Dans la suite de ce chapitre ne seront exposés que des algorithmes développés pour des régressions linéaires. Ce choix, qui peut être considéré *empha priori* comme une restriction, se justifie par le fait que :

- de nombreuses méthodes initialement créées pour des modèles linéaires peuvent être utilisées ou aisément modifiées pour être appliquées à des régressions plus complexes (pseudo linéaires ou non linéaires) [LBB⁺01] ;
- ces techniques sont suffisantes pour appréhender le principe de l'identification récursive et sont, à notre connaissance, les seules à être employées en identification récursive des sous-espaces.

Afin de justifier leur utilisation³ en pratique, nous ferons l'hypothèse de nous placer autour d'un point de fonctionnement pour lequel le processus est représentable par un modèle linéaire. Nous supposons alors que les signaux utilisés ne font en aucun cas changer de mode de fonctionnement. De plus, les entrées seront choisies afin de ne pas contraindre le procédé étudié à travailler dans sa zone de non linéarité. Ces hypothèses étant précisées, le problème du traitement des systèmes non linéaires se simplifie généralement et devient un problème de linéarité par morceaux. Il est alors possible d'approcher le régresseur non linéaire par un régresseur linéaire. Une présentation intéressante de l'impact de l'utilisation de modèles linéaires pour la modélisation de procédés industriels est développée dans [SPDR03].

Il est essentiel de rappeler que l'un des objectifs de l'identification est de fournir un modèle de processus ayant un comportement dynamique le plus proche possible du système étudié. Il faut cependant remarquer que toute procédure d'identification est affectée par le bruit (bruit de mesure, erreur de modélisation...). Ainsi, la relation modèle / procédé est représentable par l'équation suivante :

$$y(t) = y_m(t) + n(t). \quad (2.8)$$

Le terme additif n exprime le fait que la sortie y n'est pas une fonction exacte des observations passées. Le but de l'identification consiste alors à construire l'erreur entre la sortie mesurée y et la sortie y_m prédite par le modèle puis d'utiliser cette erreur, à chaque instant, pour modifier les paramètres du modèle afin que y_m soit une bonne prédiction de la sortie du système y , c'est

³Rappelons que la plupart des systèmes physiques et industriels présentent des non linéarités.

à dire que $\{y(t) - y_m(t)\}$ soit de moyenne nulle et de variance minimale au sens d'un certain critère.

En supposant que le système soit modélisable par une régression linéaire, dans le cas idéal où n serait totalement déterministe, la connaissance du vecteur de paramètres $\boldsymbol{\eta}$ serait suffisante à la réalisation d'un modèle ayant les mêmes caractéristiques que le processus étudié. En effet, si ces paramètres étaient connus, il serait aisé de définir une valeur de la sortie du modèle calculée à partir des mesures disponibles à l'instant t :

$$y_m(t) = \boldsymbol{\phi}^T(t)\boldsymbol{\eta}. \quad (2.9)$$

Le problème rencontré en identification est que le vecteur $\boldsymbol{\eta}$ n'est pas connu. Nous définissons donc, pour chaque valeur du vecteur de paramètres estimés $\hat{\boldsymbol{\eta}}(t)$, une sortie estimée du modèle :

$$\hat{y}_m(t) = \boldsymbol{\phi}^T(t)\hat{\boldsymbol{\eta}}. \quad (2.10)$$

Nous cherchons alors un modèle $\{\hat{\boldsymbol{\eta}}(t)\}$ qui réaliserait les deux propriétés suivantes :

P1. : reproduire asymptotiquement le comportement d'entrée-sortie du système soit :

$$\lim_{t \rightarrow \infty} (y(t) - \boldsymbol{\phi}^T(t)\hat{\boldsymbol{\eta}}(t)) = 0 \text{ avec une probabilité de } 1 ; \quad (2.11)$$

P2. : converger fortement vers les vrais paramètres, soit :

$$\lim_{t \rightarrow \infty} (\hat{\boldsymbol{\eta}}(t) - \boldsymbol{\eta}(t)) = 0 \text{ avec une probabilité de } 1. \quad (2.12)$$

Une troisième contrainte, propre aux méthodes récursives d'identification, demande d'ajuster en permanence les paramètres du modèle de procédé à chaque acquisition ou lecture du fichier de données d'entrée-sortie :

P3. : mettre à jour les paramètres du modèle :

$$\hat{\boldsymbol{\eta}}(t) = \hat{\boldsymbol{\eta}}(t-1) + \mu_{\hat{\boldsymbol{\eta}}} g_{\hat{\boldsymbol{\eta}}}(\hat{\boldsymbol{\eta}}(t-1), \mathcal{M}(t-1), u(t), y(t)) \quad (2.13)$$

$$\mathcal{M}(t) = \mathcal{M}(t-1) + \mu_{\mathcal{M}} g_{\mathcal{M}}(\hat{\boldsymbol{\eta}}(t-1), \mathcal{M}(t-1), u(t), y(t)) \quad (2.14)$$

où $g_{\hat{\boldsymbol{\eta}}}$, $g_{\mathcal{M}}$, $\mu_{\hat{\boldsymbol{\eta}}}$ et $\mu_{\mathcal{M}}$ sont des fonctions et paramètres dépendants de la méthode d'identification récursive adoptée, \mathcal{M} un ensemble d'informations condensées dans une mémoire auxiliaire et $\{u(t), y(t)\}$ les mesures obtenues lors de la dernière acquisition.

De nombreuses techniques d'identification récursives répondent à ces trois propriétés [LS83, LBB⁺01]. Nous n'en présenterons que deux dans la suite de ce chapitre :

- la méthode des moindres carrés récursifs ;
- la méthode la variable instrumentale récursive.

Ce choix se justifie par le fait que ces dernières sont à la base de toutes les méthodes récursives classiques d'identification et qu'elles permettent ainsi de développer et de comprendre la grande majorité des algorithmes adaptatifs traditionnels proposés dans la littérature.

2.3 Méthodes classiques d'identification récurrente

L'objectif de cette section est de présenter les méthodes classiques d'identification récurrente permettant d'estimer en ligne un modèle du procédé qui reproduit asymptotiquement le comportement d'entrée-sortie du système. Différentes solutions ont été développées [LS83]. Elles diffèrent principalement selon la structure du modèle choisie *a priori*. En effet :

- Si le modèle utilisé est linéaire en les paramètres, de simples critères quadratiques peuvent être introduits. Leur minimisation est alors réalisable par des algorithmes récursifs de type moindres carrés.
- Si le modèle considéré est pseudo linéaire ou non linéaire, la tâche est bien plus complexe. Aucune solution analytique n'est envisageable. Des approximations, plus ou moins conséquentes, sont inévitables pour obtenir des algorithmes récursifs. Deux méthodes ont été plus particulièrement étudiées : le maximum de vraisemblance récursif [Söd73] et la méthode d'erreur de prédiction récurrente [Lju81]. Les approximations nécessaires à leur implémentation sont choisies de telle sorte que les algorithmes obtenus soient exacts lorsqu'ils sont appliqués à un modèle linéaire [LS83]. Ces techniques ne seront cependant pas présentées dans la suite de ce chapitre puisqu'elles n'apporteraient aucun bénéfice pour la compréhension des méthodes récurrentes des sous-espaces.

Deux solutions fondées sur la régression linéaire suivante :

$$y(t) = \boldsymbol{\phi}^T(t)\boldsymbol{\eta} + n(t) \quad (2.15)$$

vont être exposées dans la suite de cette section. Nous commencerons plus précisément cette présentation par l'analyse de la technique la plus ancienne : les moindres carrés [RAB⁺01]. Cette méthode est certainement le plus didactique des algorithmes adaptatifs d'estimation. Elle permet en effet de mettre en évidence les étapes fondamentales de toute procédure récurrente et de souligner les principales difficultés rencontrées en identification temps réel. Ce travail sera ensuite suivi de l'étude de la méthode de la variable instrumentale. Cette technique présente le sérieux avantage d'être capable d'éliminer le biais asymptotique existant lors de l'application des moindres carrés en présence d'un bruit non blanc. Cette propriété sera particulièrement utilisée en identification récurrente des sous-espaces.

2.3.1 Moindres carrés récursifs

Une solution à l'obtention du vecteur de paramètres est de sélectionner celui qui permet le meilleur ajustement de la sortie du modèle à la sortie du système. Dans le cas d'une régression linéaire, cette propriété peut être réalisée au sens des moindres carrés par un estimateur $\hat{\boldsymbol{\eta}}(t)$ qui minimise le critère quadratique :

$$J(\boldsymbol{\eta}, \mathcal{L}) = \sum_{t=1}^M \varepsilon(t|\boldsymbol{\eta})^2 = \boldsymbol{\varepsilon}^T(M)\boldsymbol{\varepsilon}(M) \quad (2.16)$$

pour lequel l'erreur d'équation $\varepsilon(t|\hat{\boldsymbol{\eta}})$ est définie par :

$$\varepsilon(t|\hat{\boldsymbol{\eta}}) = y(t) - \hat{y}_m(t|\hat{\boldsymbol{\eta}}) \quad (2.17)$$

et $\boldsymbol{\varepsilon}(M)$ le vecteur d'erreur de prédiction construit comme suit :

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon(1) \\ \vdots \\ \varepsilon(M) \end{bmatrix}. \quad (2.18)$$

Puisque $J(\boldsymbol{\eta}, \mathcal{L})$ est quadratique en $\boldsymbol{\eta}$ et donc strictement convexe, le minimum est obtenu en annulant sa dérivée. L'estimateur des moindres carrés peut donc se mettre sous la forme suivante :

$$\hat{\boldsymbol{\eta}}(M) = (\boldsymbol{\Phi}(M)\boldsymbol{\Phi}^T(M))^{-1} \boldsymbol{\Phi}(M)\mathbf{Y}(M) \quad (2.19)$$

où :

$$\boldsymbol{\Phi}^T(M) = \begin{bmatrix} \boldsymbol{\phi}(1) \\ \vdots \\ \boldsymbol{\phi}(M) \end{bmatrix} \quad (2.20)$$

$$\mathbf{Y}(M) = \begin{bmatrix} y(1) \\ \vdots \\ y(M) \end{bmatrix}. \quad (2.21)$$

La mise en œuvre en temps réel de cette équation est relativement lourde puisqu'elle requiert une sauvegarde de toutes les données et une inversion matricielle à chaque pas d'échantillonnage. Il serait intéressant de modifier l'équation (2.19) afin d'éviter cette inversion matricielle tout en tenant compte de la mise à jour des observations. Pour cela, remarquons que, lorsqu'une nouvelle acquisition est réalisée, les matrices $\boldsymbol{\Phi}^T(t)$ et $\mathbf{Y}(t)$ s'écrivent :

$$\boldsymbol{\Phi}^T(t) = \begin{bmatrix} \boldsymbol{\Phi}^T(t-1) \\ \boldsymbol{\phi}^T(t) \end{bmatrix} \quad (2.22)$$

$$\mathbf{Y}(t) = \begin{bmatrix} \mathbf{Y}(t-1) \\ y(t) \end{bmatrix}. \quad (2.23)$$

En introduisant ces deux équations dans l'expression (2.19), il s'en suit :

$$\hat{\boldsymbol{\eta}}(t) = (\boldsymbol{\Phi}(t-1)\boldsymbol{\Phi}^T(t-1) + \boldsymbol{\phi}(t)\boldsymbol{\phi}^T(t))^{-1} (\boldsymbol{\Phi}(t-1)\mathbf{Y}(t-1) + \boldsymbol{\phi}(t)y(t)). \quad (2.24)$$

Malheureusement, à ce stade, cette équation nécessite toujours une inversion matricielle. Afin d'éviter ce calcul coûteux numériquement, il est intéressant d'utiliser le lemme d'inversion matricielle :

Lemme 2.1. Soient \mathbf{A} , \mathbf{B} , \mathbf{C} et \mathbf{D} des matrices dont les dimensions sont telles que le produit \mathbf{BCD} et la somme $\mathbf{A} + \mathbf{BCD}$ existent. En supposant de plus que \mathbf{A} et \mathbf{C} soient régulières, l'identité usuelle suivante :

$$[\mathbf{A} + \mathbf{BCD}]^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}[\mathbf{DA}^{-1}\mathbf{B} + \mathbf{C}^{-1}]^{-1}\mathbf{DA}^{-1} \quad (2.25)$$

est toujours vérifiée.

En appliquant ce lemme à $\mathbf{L}(t) = (\Phi(t)\Phi^T(t))^{-1} = (\Phi(t-1)\Phi^T(t-1) + \phi(t)\phi^T(t))^{-1}$, il est aisé de construire la forme récursive de l'estimateur des moindres carrés :

$$\hat{\boldsymbol{\eta}}(t) = \hat{\boldsymbol{\eta}}(t-1) + \mathbf{K}(t)\boldsymbol{\varepsilon}_0(t) \quad (2.26a)$$

$$\boldsymbol{\varepsilon}_0(t) = y(t) - \boldsymbol{\phi}^T(t)\hat{\boldsymbol{\eta}}(t-1) \quad (2.26b)$$

$$\mathbf{K}(t) = \mathbf{L}(t)\boldsymbol{\phi}(t) = \frac{\mathbf{L}(t-1)\boldsymbol{\phi}(t)}{1 + \boldsymbol{\phi}^T(t)\mathbf{L}(t-1)\boldsymbol{\phi}(t)} \quad (2.26c)$$

$$\mathbf{L}(t) = \mathbf{L}(t-1) - \mathbf{K}(t)\boldsymbol{\phi}^T(t)\mathbf{L}(t-1) \quad (2.26d)$$

pour laquelle $\boldsymbol{\varepsilon}_0$ est l'erreur de prédiction *a priori*. Il est important de remarquer que l'erreur d'estimation a posteriori $\boldsymbol{\varepsilon} = y(t) - \boldsymbol{\phi}^T(t)\hat{\boldsymbol{\eta}}(t)$ vérifie :

$$\boldsymbol{\varepsilon}(t) \leq \boldsymbol{\varepsilon}_0(t). \quad (2.27)$$

Cette équation montre que le processus d'identification est affiné en temps réel au fur et à mesure que de nouvelles informations sont acquises.

Remarque 2.1. *Il est possible de récrire l'estimateur des moindres carrés sous la forme suivante :*

$$\hat{\boldsymbol{\eta}}(t) = \mathbf{L}(t) \left[\mathbf{L}(0)^{-1}\hat{\boldsymbol{\eta}}(0) + \sum_{i=1}^t \boldsymbol{\phi}(i)y(i) \right] \quad (2.28)$$

avec

$$\mathbf{L}(t)^{-1} = \mathbf{L}(0)^{-1} + \sum_{i=1}^t \boldsymbol{\phi}(i)\boldsymbol{\phi}^T(i) \quad (2.29)$$

qui permet de mettre en évidence que l'initialisation de l'algorithme d'adaptation paramétrique peut se faire comme suit :

$$\hat{\boldsymbol{\eta}}(0) = 0 \text{ et } \mathbf{L}(0)^{-1} = 0, \quad (2.30)$$

cette dernière condition pouvant être réalisée de la façon suivante :

$$\mathbf{L}(0) = \frac{1}{\delta} \mathbf{I} \quad (2.31)$$

avec $\delta = 0.001$ par exemple.

Remarque 2.2. *Dans certaines situations, il est intéressant de considérer que les données d'entrée-sortie sont des réalisations de variables aléatoires ayant une distribution donnée. Dans ce contexte, B. Widrow [WS85] propose d'utiliser le critère stochastique suivant :*

$$J_{stoch}(\boldsymbol{\eta}) = \mathbb{E} \left\| y - \boldsymbol{\phi}^T \boldsymbol{\eta} \right\| \quad (2.32)$$

où \mathbb{E} est l'espérance mathématique. Son minimum est atteint lorsque sa dérivée est nulle :

$$\frac{\partial J_{stoch}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} = \mathbb{E} \{ \boldsymbol{\phi} (y - \boldsymbol{\phi}^T \boldsymbol{\eta}) \} = 0. \quad (2.33)$$

L'algorithme récursif suivant peut alors être appliqué [WH60] :

$$\boldsymbol{\eta}(t) = \boldsymbol{\eta}(t-1) + \mu \mathbb{E} \{ \boldsymbol{\phi}(t) (y(t) - \boldsymbol{\phi}^T(t)\boldsymbol{\eta}(t)) \} \quad (2.34)$$

où μ est le pas de descente du gradient. Cet algorithme est appelé gradient stochastique. Il s'agit de la version stochastique des moindres carrés récursifs [Lju02].

A ce stade de l'étude, il est intéressant d'explorer les propriétés que possède l'estimateur $\hat{\boldsymbol{\eta}}$. Il est donc nécessaire de poursuivre cette analyse en examinant l'effet des perturbations sur les propriétés des moindres carrés récursifs.

2.3.1.1 Capacité de poursuite des moindres carrés récursifs : une approche bayésienne

Comme tout algorithme adaptatif, l'un des objectifs des moindres carrés récursifs est de pouvoir adapter les paramètres du modèle identifié à l'évolution du système étudié. Il est donc intéressant d'explicitier les caractéristiques d'adaptation de l'algorithme des moindres carrés vis-à-vis d'une modélisation aléatoire du système.

Lorsqu'on étudie la formule (2.26d), il est manifeste que l'algorithme des moindres carrés récursifs est un algorithme à gain d'adaptation décroissant. Ceci se voit clairement si l'on considère l'estimation d'un seul paramètre. En effet [LBB⁺01] :

$$L(t+1) = \frac{L(t)}{1 + \phi(t)^2 L(t)} < L(t). \quad (2.35)$$

Cet algorithme donne donc de moins en moins de poids aux nouvelles mesures. Il ne conviendra pas à l'estimation de paramètres variants dans le temps. Pour résoudre ce problème, des versions pondérées de l'algorithme classique des moindres carrés récursifs ont été développées [LS83]. Une façon simple d'introduire ces pondérations est de considérer de nouveau la définition de l'estimateur des moindres carrés :

$$\hat{\boldsymbol{\eta}}(t) = (\boldsymbol{\Phi}(t)\boldsymbol{\Phi}^T(t))^{-1} \boldsymbol{\Phi}(t)\mathbf{Y}(t) \quad (2.36)$$

et d'insérer des facteurs d'oubli λ_1 et λ_2 au sein des formules de mise à jour de $\boldsymbol{\Phi}$ et \mathbf{Y} :

$$\boldsymbol{\Phi}^T(t) = \begin{bmatrix} \sqrt{\lambda_1(t)}\boldsymbol{\Phi}^T(t-1) \\ \sqrt{\lambda_2(t)}\boldsymbol{\phi}^T(t) \end{bmatrix} \quad (2.37)$$

$$\mathbf{Y}(t) = \begin{bmatrix} \sqrt{\lambda_1(t)}\mathbf{Y}(t-1) \\ \sqrt{\lambda_2(t)}y(t) \end{bmatrix}. \quad (2.38)$$

L'utilisation de telles pondérations permet ainsi de donner des poids particuliers aux différentes observations. En introduisant ces relations au sein de l'équation (2.36), nous obtenons :

$$\hat{\boldsymbol{\eta}}(t) = (\lambda_1(t)\boldsymbol{\Phi}(t-1)\boldsymbol{\Phi}^T(t-1) + \lambda_2(t)\boldsymbol{\phi}(t)\boldsymbol{\phi}^T(t))^{-1} (\lambda_1(t)\boldsymbol{\Phi}(t-1)\mathbf{Y}(t-1) + \lambda_2(t)\boldsymbol{\phi}(t)y(t)). \quad (2.39)$$

En appliquant le lemme d'inversion matricielle (cf. equ. (2.25)) à $\mathbf{L}(t) = (\boldsymbol{\Phi}(t)\boldsymbol{\Phi}^T(t))^{-1}$, la version pondérée de l'algorithme récursif des moindres carrés s'écrit :

$$\hat{\boldsymbol{\eta}}(t) = \hat{\boldsymbol{\eta}}(t-1) + \mathbf{K}(t)\boldsymbol{\varepsilon}_0(t) \quad (2.40a)$$

$$\boldsymbol{\varepsilon}_0(t) = y(t) - \boldsymbol{\phi}^T(t)\hat{\boldsymbol{\eta}}(t-1) \quad (2.40b)$$

$$\mathbf{K}(t) = \frac{\mathbf{L}(t-1)\boldsymbol{\phi}(t)}{\frac{\lambda_1(t)}{\lambda_2(t)} + \boldsymbol{\phi}^T(t)\mathbf{L}(t-1)\boldsymbol{\phi}(t)} \quad (2.40c)$$

$$\mathbf{L}(t) = \frac{1}{\lambda_1(t)} (\mathbf{L}(t-1) - \mathbf{K}(t)\boldsymbol{\phi}^T(t)\mathbf{L}(t-1)). \quad (2.40d)$$

Preuve : Soit :

$$\mathbf{L}(t) = (\Phi(t)\Phi^T(t))^{-1} = (\lambda_1(t)\Phi(t-1)\Phi^T(t-1) + \lambda_2(t)\phi(t)\phi^T(t))^{-1}. \quad (2.41)$$

En appliquant le lemme d'inversion matricielle (cf. equ. (2.25)), nous obtenons :

$$\mathbf{L}(t) = \frac{1}{\lambda_1(t)} \left(\mathbf{L}(t-1) - \frac{\mathbf{L}(t-1)\phi(t)\phi^T(t)\mathbf{L}(t-1)}{\frac{\lambda_1(t)}{\lambda_2(t)} + \phi^T(t)\mathbf{L}(t-1)\phi(t)} \right). \quad (2.42)$$

Le vecteur des paramètres estimé s'écrit alors :

$$\begin{aligned} \hat{\boldsymbol{\eta}}(t) &= \mathbf{L}(t) (\lambda_1(t)\Phi(t-1)\mathbf{Y}(t-1) + \lambda_2(t)\phi(t)y(t)) \\ &= \mathbf{L}(t) ((\mathbf{L}^{-1}(t) - \lambda_2(t)\phi(t)\phi^T(t))\hat{\boldsymbol{\eta}}(t-1) + \lambda_2(t)\phi(t)y(t)) \\ &= \hat{\boldsymbol{\eta}}(t-1) + \lambda_2(t)\mathbf{L}(t)\phi(t) (y(t) - \phi^T(t)\hat{\boldsymbol{\eta}}(t-1)). \end{aligned} \quad (2.43)$$

Puisque :

$$\begin{aligned} \mathbf{K}(t) &= \lambda_2(t)\mathbf{L}(t)\phi(t) \\ &= \frac{\lambda_2(t)}{\lambda_1(t)} \left(\mathbf{L}(t-1)\phi(t) - \frac{\mathbf{L}(t-1)\phi(t)\phi^T(t)\mathbf{L}(t-1)\phi(t)}{\frac{\lambda_1(t)}{\lambda_2(t)} + \phi^T(t)\mathbf{L}(t-1)\phi(t)} \right) \\ &= \frac{\mathbf{L}(t-1)\phi(t)}{\frac{\lambda_1(t)}{\lambda_2(t)} + \phi^T(t)\mathbf{L}(t-1)\phi(t)}, \end{aligned} \quad (2.44)$$

l'algorithme final s'écrit :

$$\hat{\boldsymbol{\eta}}(t) = \hat{\boldsymbol{\eta}}(t-1) + \mathbf{K}(t)\boldsymbol{\varepsilon}_0(t) \quad (2.45a)$$

$$\boldsymbol{\varepsilon}_0(t) = y(t) - \phi^T(t)\hat{\boldsymbol{\eta}}(t-1) \quad (2.45b)$$

$$\mathbf{K}(t) = \frac{\mathbf{L}(t-1)\phi(t)}{\frac{\lambda_1(t)}{\lambda_2(t)} + \phi^T(t)\mathbf{L}(t-1)\phi(t)} \quad (2.45c)$$

$$\mathbf{L}(t) = \frac{1}{\lambda_1(t)} (\mathbf{L}(t-1) - \mathbf{K}(t)\phi^T(t)\mathbf{L}(t-1)). \quad (2.45d)$$

◀

La bornitude de $\mathbf{L}(t)$ est ainsi assurée afin d'éviter que le gain d'adaptation devienne asymptotiquement nul pourvu que les facteurs d'oubli $\lambda_1(t)$ et $\lambda_2(t)$ soient bornés, lentement variables et vérifient :

$$0 < \lambda_1(t) \leq 1 \text{ et } 0 \leq \lambda_2(t) \leq 1. \quad (2.46)$$

En pratique, la sélection des facteurs d'oubli a pour but de maintenir le gain d'adaptation \mathbf{K} à une valeur telle qu'il soit possible de suivre les paramètres. Malheureusement, comme nous allons le montrer dans la suite de ce paragraphe, l'utilisation de facteurs d'oubli faibles, idéaux pour une adaptation rapide, rend l'algorithme des moindres carrés récurrents particulièrement sensible au bruit du système. Afin de mettre en évidence ce compromis, considérons la relation suivante :

$$\hat{\boldsymbol{\eta}}(t) = \hat{\boldsymbol{\eta}}(t-1) + \lambda_2(t)\mathbf{L}(t)\phi(t) (y(t) - \phi^T(t)\hat{\boldsymbol{\eta}}(t-1)) \quad (2.47)$$

mise en évidence dans la preuve de l'algorithme (2.40) (cf. equ. (2.43)) et supposons que :

- Le vecteur des vrais paramètres $\boldsymbol{\eta}$ varie selon la loi suivante :

$$\boldsymbol{\eta}(t) = \boldsymbol{\eta}(t-1) + \boldsymbol{\omega}(t) \quad (2.48)$$

avec $\boldsymbol{\omega}(t)$ un bruit blanc gaussien de matrice de covariance $\mathbf{R}_{\boldsymbol{\omega}}$. Cette dernière écriture traduit la possible variation aléatoire du vecteur de paramètres selon une loi gaussienne.

– La sortie du système vérifie :

$$y(t) = \boldsymbol{\phi}^T(t) \boldsymbol{\eta}(t) + n(t) \quad (2.49)$$

où $\boldsymbol{\phi}$ et n sont des variables aléatoires indépendantes de moyenne nulle et de matrice de covariance respective \mathbf{R}_ϕ et \mathbf{R}_n .

Il est alors possible de calculer l'erreur d'estimation des paramètres $\tilde{\boldsymbol{\eta}}(t) = \boldsymbol{\eta}(t) - \hat{\boldsymbol{\eta}}(t)$:

$$\begin{aligned} \tilde{\boldsymbol{\eta}}(t) &= \boldsymbol{\eta}(t) - \hat{\boldsymbol{\eta}}(t) \\ &= \boldsymbol{\eta}(t-1) + \boldsymbol{\omega}(t) - \hat{\boldsymbol{\eta}}(t-1) - \lambda_2(t) \mathbf{L}(t) \boldsymbol{\phi}(t) (y(t) - \boldsymbol{\phi}^T(t) \hat{\boldsymbol{\eta}}(t-1)) \\ &= \tilde{\boldsymbol{\eta}}(t-1) + \boldsymbol{\omega}(t) - \lambda_2(t) \mathbf{L}(t) \boldsymbol{\phi}(t) (\boldsymbol{\phi}^T(t) \boldsymbol{\eta}(t) + n(t) - \boldsymbol{\phi}^T(t) \hat{\boldsymbol{\eta}}(t-1)) \\ &= \tilde{\boldsymbol{\eta}}(t-1) + \boldsymbol{\omega}(t) \\ &\quad - \lambda_2(t) \mathbf{L}(t) \boldsymbol{\phi}(t) (\boldsymbol{\phi}^T(t) (\boldsymbol{\eta}(t-1) + \boldsymbol{\omega}(t)) - \boldsymbol{\phi}^T(t) \hat{\boldsymbol{\eta}}(t-1) + n(t)) \\ &= [\mathbf{I} - \lambda_2(t) \mathbf{L}(t) \boldsymbol{\phi}(t) \boldsymbol{\phi}^T(t)] \tilde{\boldsymbol{\eta}}(t-1) \\ &\quad + [\mathbf{I} - \lambda_2(t) \mathbf{L}(t) \boldsymbol{\phi}(t) \boldsymbol{\phi}^T(t)] \boldsymbol{\omega}(t) - \lambda_2(t) \mathbf{L}(t) \boldsymbol{\phi}(t) n(t). \end{aligned} \quad (2.50)$$

La quantité qui nous intéresse plus particulièrement est la matrice de covariance de l'erreur d'estimation $\mathbf{R}_{\tilde{\boldsymbol{\eta}}}$. Puisque les variables considérées sont supposées indépendantes, cette matrice s'écrit :

$$\begin{aligned} \mathbf{R}_{\tilde{\boldsymbol{\eta}}}(t) &= [\mathbf{I} - \mathbb{E}\{\lambda_2(t) \mathbf{L}(t)\} \mathbf{R}_\phi(t)] \hat{\mathbf{R}}_{\tilde{\boldsymbol{\eta}}}(t-1) [\mathbf{I} - \mathbb{E}\{\lambda_2(t) \mathbf{L}(t)\} \mathbf{R}_\phi(t)]^T \\ &\quad + [\mathbf{I} - \mathbb{E}\{\lambda_2(t) \mathbf{L}(t)\} \mathbf{R}_\phi(t)] [\mathbf{I} - \mathbb{E}\{\lambda_2(t) \mathbf{L}(t)\} \mathbf{R}_\phi(t)]^T \mathbf{R}_\omega(t) \\ &\quad + \mathbb{E}\{\lambda_2(t) \mathbf{L}(t)\} \mathbf{R}_\phi(t) \mathbb{E}\{\lambda_2(t) \mathbf{L}(t)\}^T \mathbf{R}_n(t). \end{aligned} \quad (2.51)$$

Si nous nous plaçons dans le cas où le vecteur de régression est totalement déterministe et sachant que :

$$\mathbb{E}\{\mathbf{L}(t)\} = \mathbb{E}\{\boldsymbol{\phi} \boldsymbol{\phi}^T\}^{-1} = \mathbf{R}_\phi^{-1}, \quad (2.52)$$

l'équation (2.51) est équivalente à :

$$\mathbf{R}_{\tilde{\boldsymbol{\eta}}}(t) = (1 - \lambda_2(t))^2 \hat{\mathbf{R}}_{\tilde{\boldsymbol{\eta}}}(t-1) + (1 - \lambda_2(t))^2 \mathbf{R}_\omega(t) + \lambda_2^2(t) \mathbf{R}_\phi^{-1}(t) \mathbf{R}_n(t). \quad (2.53)$$

Ainsi, quand t tend vers l'infini, en négligeant le terme $\lambda_2^2 \mathbf{R}_{\tilde{\boldsymbol{\eta}}}$ [GL95] :

$$\lim_{t \rightarrow \infty} \mathbf{R}_{\tilde{\boldsymbol{\eta}}}(t) \approx \frac{1}{2} \left(\frac{(1 - \lambda_2)^2}{\lambda_2} \mathbf{R}_\omega + \lambda_2 \mathbf{R}_\phi^{-1} \mathbf{R}_n \right). \quad (2.54)$$

Cette dernière expression met en évidence un compromis intrinsèque dans le choix du facteur d'oubli : une faible valeur de λ_2 diminue l'effet du bruit $n(t)$ dans l'estimation des paramètres (par l'intermédiaire du terme $\lambda_2 \mathbf{R}_n$) mais augmente l'influence de l'erreur de poursuite (avec $\frac{(1-\lambda_2)^2}{\lambda_2} \mathbf{R}_\omega$) au cours de cette même estimation. Il est donc essentiel en pratique d'atteindre le bon équilibre entre capacité de poursuite et sensibilité aux bruits. De nombreux choix de valeurs des facteurs d'oubli $\lambda_1(t)$ et $\lambda_2(t)$ ont été proposés dans la littérature. Une description précise des principaux algorithmes pondérés est réalisée dans [LLM97, LBB⁺01]. Nous ne présentons qu'un bref récapitulatif des alternatives possibles pour $\lambda_1(t)$ et $\lambda_2(t)$ au sein du tableau 2.1, le lecteur intéressé pouvant facilement étudier les références ci-dessus.

| Nom | Facteur d'oubli | Utilisation |
|--------------------------|---|--|
| Gain décroissant | $\lambda_1(t) = 1$ et $\lambda_2(t) = 1$ | Systèmes stationnaires |
| Facteur d'oubli fixe | $\lambda_1(t) = \lambda_1 \in]0, 1[$ et $\lambda_2(t) = 1$ | Systèmes lentement variables |
| Facteur d'oubli variable | $\lambda_1(t) = \lambda_0 \lambda_1(t-1) + 1 - \lambda_0$ et $\lambda_2(t) = 1$ | Traitement des conditions initiales |
| Trace constante | $\lambda_1(t) = \lambda_2(t)$ tel que $\text{tr}(\mathbf{L}(t)) = \text{tr}(\mathbf{L}(t-1))$ | Systèmes non stationnaires. Gain borné |
| Gain constant | $\lambda_1(t) = 1$ et $\lambda_2(t) = 0$ | Convergence lente. Robustesse vis-à-vis des variations |

TAB. 2.1: Principaux choix pour $\lambda_1(t)$ et $\lambda_2(t)$ pour la poursuite de systèmes variant dans le temps.

Remarque 2.3. *Il est à noter que la représentation (2.48) et (2.49) est estimable de façon optimale au sens des moindres carrés dans le cas où les perturbations n et $\boldsymbol{\omega}$ sont des bruits blancs gaussiens. En effet, la solution de ce problème est le célèbre filtre de Kalman [KB61] dont l'algorithme récursif est décrit de la façon suivante :*

$$\hat{\boldsymbol{\eta}}(t) = \hat{\boldsymbol{\eta}}(t-1) + \mathbf{K}(t)\boldsymbol{\varepsilon}_0(t) \quad (2.55a)$$

$$\boldsymbol{\varepsilon}_0(t) = y(t) - \boldsymbol{\phi}^T(t)\hat{\boldsymbol{\eta}}(t-1) \quad (2.55b)$$

$$\mathbf{K}(t) = \frac{\mathbf{L}(t-1)\boldsymbol{\phi}(t)}{\hat{\mathbf{R}}_n(t) + \boldsymbol{\phi}^T(t-1)\mathbf{L}(t-1)\boldsymbol{\phi}(t)} \quad (2.55c)$$

$$\mathbf{L}(t) = \mathbf{L}(t-1) - \mathbf{K}(t)\boldsymbol{\phi}^T(t)\mathbf{L}(t-1) + \hat{\mathbf{R}}_{\boldsymbol{\omega}}(t). \quad (2.55d)$$

Il est alors possible de récrire l'algorithme des moindres carrés pondérés à partir des équations précédentes en considérant [Lju02] :

$$\hat{\mathbf{R}}_{\boldsymbol{\omega}}(t) = \left(\frac{1}{\lambda_1(t)} - 1 \right) \left[\mathbf{L}(t-1) - \frac{\mathbf{L}(t-1)\boldsymbol{\phi}(t)\boldsymbol{\phi}^T(t)\mathbf{L}(t-1)}{\frac{\lambda_1(t)}{\lambda_2(t)} + \boldsymbol{\phi}^T(t)\mathbf{L}(t-1)\boldsymbol{\phi}(t)} \right] \quad (2.56a)$$

$$\hat{\mathbf{R}}_n(t) = \frac{\lambda_1(t)}{\lambda_2(t)}. \quad (2.56b)$$

Cette dernière équation (2.56b) met en évidence que le choix optimal des facteurs d'oubli $\lambda_1(t)$ et $\lambda_2(t)$ nécessite la connaissance des propriétés du second ordre du bruit $n(t)$.

2.3.1.2 Biais asymptotique de l'estimateur en présence de perturbations

La méthode des moindres carrés a l'avantage d'une certaine simplicité. En pratique, elle présente toutefois l'inconvénient de donner fréquemment une estimation asymptotiquement biaisée du vecteur des paramètres. Afin de mettre en évidence ce biais d'estimation, considérons de nouveau la régression linéaire suivante :

$$y(t) = \boldsymbol{\phi}^T(t)\boldsymbol{\eta} + n(t). \quad (2.57)$$

Il est alors facile de vérifier que, matriciellement :

$$\mathbf{Y}(t) = \boldsymbol{\Phi}^T(t)\boldsymbol{\eta} + \mathbf{N}(t). \quad (2.58)$$

En introduisant cette équation au sein de la relation (2.19), nous obtenons :

$$\hat{\boldsymbol{\eta}}(M) = \boldsymbol{\eta} + (\boldsymbol{\Phi}(M)\boldsymbol{\Phi}^T(M))^{-1} \boldsymbol{\Phi}(M)\mathbf{N}(M). \quad (2.59)$$

L'estimation du vecteur des paramètres sera donc asymptotiquement non biaisée si et seulement si :

$$\mathbb{E}\{\boldsymbol{\phi}(t)n(t)\} = 0, \quad (2.60)$$

c'est à dire si $\boldsymbol{\phi}$ et n sont décorrélés. L'estimateur des moindres carrés est donc asymptotiquement non biaisé uniquement lorsque n est un bruit blanc gaussien. Pour remédier à cette restriction, plusieurs algorithmes ont été développés [LBB⁺01]. L'idée commune à ces différentes techniques est la suivante. Pour $\hat{\boldsymbol{\eta}} = \boldsymbol{\eta}$, l'erreur d'équation vaut :

$$\boldsymbol{\varepsilon}(t) = n(t). \quad (2.61)$$

Ainsi, pour avoir $\hat{\boldsymbol{\eta}}(M) = \boldsymbol{\eta}$ quand $M \rightarrow \infty$, il suffit que :

$$\mathbb{E}\{\boldsymbol{\phi}(t)\boldsymbol{\varepsilon}(t)\} = 0. \quad (2.62)$$

Une estimation asymptotiquement non biaisée peut alors être obtenue dans les cas suivants :

- l'erreur d'équation est un bruit blanc pour $\hat{\boldsymbol{\eta}} = \boldsymbol{\eta}$;
- le vecteur des observations et l'erreur d'équation sont non corrélés pour $\hat{\boldsymbol{\eta}} = \boldsymbol{\eta}$.

Nous pouvons donc distinguer deux types de méthodes d'estimation fournissant des paramètres asymptotiquement non biaisés :

- les méthodes basées sur le blanchiment de l'erreur d'équation ;
- les méthodes fondées sur la décorrélation observation / erreur d'équation.

Les techniques les plus connues sont les moindres carrés étendus [You68], les moindres carrés généralisés [Cla67], les méthodes d'erreur de sortie [Lan74] ou les méthodes de la variable instrumentale [Rei41]. Il n'y a malheureusement aucun algorithme d'estimation paramétrique capable de fournir des estimées asymptotiquement non biaisées pour tout type de modèle de bruit. Au contraire, pour chaque structure de modèle, il existe une ou plusieurs techniques aboutissant à des résultats consistants. Le choix de la méthode utilisée dépend donc du système, de la structure du modèle et de la nature du bruit. Nous développerons cependant plus précisément l'une d'entre elle : la méthode de la variable instrumentale [WP67]. Cette dernière est en effet très utilisée en identification récursive des sous-espaces.

2.3.2 Variable instrumentale récursive

Au sein de la section précédente, il a été mis en évidence que l'estimateur des moindres carrés (récurifs) fournissait des estimées asymptotiquement biaisées en présence de perturbations non blanches. Différentes solutions ont été développées durant ces cinquante dernières années pour résoudre ce problème [KS61, May67, Lan74, SS89, VV91, LLM97, Lju99, Yer00]. Parmi celles-ci, la méthode de la variable instrumentale a connu une attention particulière puisqu'elle peut être facilement considérée comme une simple modification des moindres carrés. Elle fut introduite par Reiersøl en 1941 pour résoudre certains problèmes d'économie [Rei41]. De nombreux articles et livres ont été écrits sur ce sujet [WP67, You70, Cai76, War77, SS83a, LS83, You84, SS89]. Nous n'en présenterons donc que les fondements, les lecteurs intéressés pouvant étudier les références précédentes.

Toutes les techniques basées sur l'utilisation de la variable instrumentale exploitent la propriété suivante : les signaux déterministes présentent des séquences de corrélation temporelle infinies tandis que les signaux stochastiques possèdent des séquences de corrélation temporelle bornées. En conséquence, les composantes de bruit au sein de la matrice d'intercorrélation formée à partir du vecteur de régression à l'instant t et d'une version retardée de ce dernier sont réduites de manière significative, à condition, bien sûr, que le signal et le bruit soient statistiquement indépendants et que le retard au sein du vecteur de régression décalé soit supérieur à la longueur de corrélation du bruit. L'idée générale des méthodes de la variable instrumentale consiste donc à créer un nouveau vecteur de régression ξ qui soit fortement corrélé avec les données non perturbées mais non corrélé avec le bruit. En effet, la condition d'excitation persistante [Lju99] :

$$\mathbb{E} \{ \xi(t) \phi(t) \} \text{ régulière} \quad (2.63)$$

nécessite que ϕ et ξ soient suffisamment corrélés, alors que la condition de consistance de l'estimateur :

$$\mathbb{E} \{ \xi(t) n(t) \} = 0 \quad (2.64)$$

requiert que ξ et n soient indépendantes et centrées. Or, nous savons que :

$$n = \varepsilon \quad (2.65)$$

pour $\hat{\eta} = \eta$. Nous recherchons donc une variable instrumentale telle que :

$$\mathbb{E} \{ \xi(t) \varepsilon(t) \} = 0. \quad (2.66)$$

Puisque nous n'avons accès qu'à un nombre fini de données, la relation précédente s'écrit :

$$\frac{1}{M} \sum_{t=1}^M \xi(t) \varepsilon(t) = \frac{1}{M} \sum_{t=1}^M \xi(t) (y(t) - \phi^T(t) \eta) = 0 \quad (2.67)$$

en remplaçant l'espérance mathématique par une somme finie. Deux algorithmes de minimisation doivent alors être envisagés selon le nombre d'instruments composant le vecteur ξ . En effet, lorsqu'il est possible de construire un vecteur ξ de dimension égale à celle de ϕ , en supposant que la condition d'excitation persistante (2.63) soit vérifiée, la matrice⁴ :

$$\sum_{t=1}^M \xi(t) \phi^T(t) = \Xi(t) \Phi^T(t) \quad (2.68)$$

est inversible. La technique de minimisation basée sur le lemme d'inversion matricielle proposée au sein du paragraphe précédent est alors applicable au problème suivant :

$$\hat{\eta}(t) = \underbrace{(\Xi(t) \Phi^T(t))^{-1}}_{\mathbf{L}_{IV}(t)} \Xi(t) \mathbf{Y}(t). \quad (2.69)$$

⁴La matrice Ξ est construite de manière similaire à Φ .

L'algorithme d'identification récursive qui en découle peut alors être simplement décrit par :

$$\hat{\boldsymbol{\eta}}(t) = \hat{\boldsymbol{\eta}}(t-1) + \mathbf{K}_{IV}(t)\boldsymbol{\varepsilon}_0(t) \quad (2.70a)$$

$$\boldsymbol{\varepsilon}_0(t) = y(t) - \boldsymbol{\phi}^T(t)\hat{\boldsymbol{\eta}}(t-1) \quad (2.70b)$$

$$\mathbf{K}_{IV}(t) = \frac{\mathbf{L}_{IV}(t-1)\boldsymbol{\xi}(t)}{1 + \boldsymbol{\phi}^T(t)\mathbf{L}_{IV}(t-1)\boldsymbol{\xi}(t)} \quad (2.70c)$$

$$\mathbf{L}_{IV}(t) = \mathbf{L}_{IV}(t-1) - \mathbf{K}_{IV}(t)\boldsymbol{\phi}^T(t)\mathbf{L}_{IV}(t-1). \quad (2.70d)$$

Remarque 2.4. *Il est important de remarquer que le fait d'avoir $\boldsymbol{\xi} \neq \boldsymbol{\phi}$ provoque une dissymétrie au niveau des coefficients de cette version récursive. Il peut alors se produire certains problèmes d'ordre numérique. Les propriétés asymptotiques de cet estimateur récursif sont cependant identiques à celles de la version hors ligne [LS83, You84, Lju99].*

Il est parfois difficile de créer une variable instrumentale ayant un nombre d'instruments strictement égal au nombre de paramètres à estimer. Il est souvent bien plus simple de construire un vecteur $\boldsymbol{\xi}$ tel que $\dim(\boldsymbol{\xi}) > \dim(\boldsymbol{\phi})$. Or, dans ce cas, le lemme d'inversion matriciel n'est plus directement applicable puisque la matrice rectangulaire $\boldsymbol{\Xi}(t)\boldsymbol{\Phi}^T(t)$ n'est pas inversible. Il est alors nécessaire d'appliquer une technique particulière de minimisation récursive, développée par B. Friedlander [Fri84], nommée « algorithme de la variable instrumentale étendue » :

$$\mathbf{K}_{EIV}(t) = \mathbf{L}_{EIV}(t-1)\boldsymbol{\Psi}(t) \left(\boldsymbol{\Lambda} + \boldsymbol{\Psi}^T(t)\mathbf{L}_{EIV}(t-1)\boldsymbol{\Psi}(t) \right)^{-1} \quad (2.71a)$$

$$\boldsymbol{\Psi} = \begin{bmatrix} \mathbf{R}_{\boldsymbol{\xi}\boldsymbol{\phi}}^T(t-1)\boldsymbol{\xi} & \boldsymbol{\phi} \end{bmatrix} \quad (2.71b)$$

$$\mathbf{g}(t) = \begin{bmatrix} \boldsymbol{\xi}^T(t)\mathbf{R}_{\boldsymbol{\xi}y}(t-1) \\ y(t) \end{bmatrix} \quad (2.71c)$$

$$\boldsymbol{\Lambda}(t) = \begin{bmatrix} -\boldsymbol{\xi}(t)\boldsymbol{\xi}^T(t) & 1 \\ 1 & 0 \end{bmatrix} \quad (2.71d)$$

$$\hat{\boldsymbol{\eta}}(t) = \hat{\boldsymbol{\eta}}(t-1) + \mathbf{K}_{EIV}(t) \left(\mathbf{g}(t) - \boldsymbol{\Psi}^T(t)\hat{\boldsymbol{\eta}}^T(t-1) \right) \quad (2.71e)$$

$$\mathbf{R}_{\boldsymbol{\xi}\boldsymbol{\phi}}(t) = \mathbf{R}_{\boldsymbol{\xi}\boldsymbol{\phi}}(t-1) + \boldsymbol{\xi}(t)\boldsymbol{\phi}^T(t) \quad (2.71f)$$

$$\mathbf{R}_{\boldsymbol{\xi}y}(t) = \mathbf{R}_{\boldsymbol{\xi}y}(t-1) + \boldsymbol{\xi}(t)y(t) \quad (2.71g)$$

$$\mathbf{L}_{EIV}(t) = \mathbf{L}_{EIV}(t-1) - \mathbf{K}_{EIV}(t)\boldsymbol{\Psi}^T(t)\mathbf{L}_{EIV}(t-1). \quad (2.71h)$$

Cet algorithme est plus complexe numériquement que la version classique (2.70). Cependant, dans certaines applications (en traitement du signal par exemple [SS89]), la qualité des estimées augmente avec le nombre d'instruments.

Un dernier problème concerne le choix des instruments composant le vecteur $\boldsymbol{\xi}$. En pratique, on distingue deux techniques usuelles pour générer ces instruments :

- la variable instrumentale à observations retardées pour laquelle :

$$\boldsymbol{\xi}^T(t) = [-y(t-r) \quad \dots \quad -y(t-n_1-r) \quad u(t) \quad \dots \quad u(t-n_2)] \quad (2.72)$$

où r désigne le retard d'observation et $\{n_1, n_2\}$ le degré du numérateur et du dénominateur de la fonction de transfert considérée ;

- la variable instrumentale avec modèle auxiliaire pour laquelle :

$$\boldsymbol{\xi}^T(t) = [-\mathbf{y}_{aux}(t) \quad \dots \quad -\mathbf{y}_{aux}(t - n_1) \quad u(t) \quad \dots \quad u(t - n_2)] \quad (2.73)$$

où $\mathbf{y}_{aux}(t) = \boldsymbol{\xi}^T(t - 1)\hat{\boldsymbol{\eta}}(t)$ désigne la sortie du modèle auxiliaire à l'instant t . Cette méthode nécessite une phase d'initialisation qui peut être réalisée, par exemple, à l'aide d'un algorithme tel que les moindres carrés récursifs.

Ces techniques, intéressantes d'un point de vue expérimental, ne fournissent malheureusement pas de variables instrumentales dites optimales [SS83a, SS83b]. L'implémentation d'une variable instrumentale optimale nécessite en effet de connaître les sorties non perturbées du procédé ainsi que la matrice de covariance du bruit, qui sont, dans la plupart des cas, inconnues et non mesurables. Un algorithme itératif a été proposé par P. Stoica et T. Söderström pour construire ce type de variable [SS83a]. Ce dernier est composé de quatre étapes dont deux doivent être répétées jusqu'à convergence de l'algorithme. Cette contrainte le rend difficilement applicable en ligne. Les instruments proposés précédemment (cf. equ. (2.72) et (2.73)) présentent donc un bon compromis consistance / temps de calcul.

2.4 Limitations des méthodes classiques et conclusion

Le principe de l'identification récursive consiste à calculer, à chaque instant, le vecteur des paramètres du modèle recherché en fonction des paramètres estimés à l'instant précédent et des nouvelles informations acquises sur le procédé. Elle constitue de ce fait la partie centrale des systèmes adaptatifs (commande ou traitement du signal) et peut être facilement utilisée comme première étape des techniques de surveillance et de diagnostic puisque, par un choix judicieux de certaines variables, il est possible de suivre les évolutions du procédé à surveiller. Les algorithmes récursifs exposés au sein de ce chapitre présentent malheureusement quelques inconvénients qui les rendent difficilement applicables dans certaines situations pratiques⁵.

- La première, qui est certainement la plus contraignante pour l'utilisateur, est que de nombreuses techniques récursives ont été proposées pour fonctionner exclusivement sur une structure de modèle de procédé et de bruit bien spécifique. Il est donc fondamental d'avoir une bonne connaissance *a priori* du système étudié pour correctement choisir l'algorithme récursif à appliquer.
- La grande majorité des méthodes récursives classiques d'identification ont été développées pour des systèmes monovariables. Il est théoriquement possible de les appliquer à des procédés multivariables. Cependant, lorsque le nombre d'entrées et de sorties devient conséquent, l'utilisation de représentations polynomiales n'est pas évidente.
- Comme toute technique basée sur des fonctions de transfert, les méthodes récursives classiques demandent de fixer *a priori* un grand nombre de paramètres *i.e.* les degrés des numérateurs et dénominateurs des fonctions de transfert des modèles de processus et de bruit ainsi que le retard relatif pour chaque couple d'entrée-sortie. Il est alors souvent nécessaire de surdimensionner les fonctions de transfert afin de ne négliger aucune bande de fréquences de la dynamique du système.

⁵Sur des systèmes présentant un grand nombre d'entrées et de sorties par exemple.

- Ces algorithmes ne permettent pas d'avoir un accès direct à des variables telles que l'état du système, variables particulièrement utiles en filtrage (Luenberger, Kalman) [Gus01] ou en commande optimale [DAC95].

Ces difficultés sont principalement liées à la structure du modèle utilisé par l'approche entrée-sortie : la fonction de transfert. Il semble donc important de développer de nouveaux algorithmes d'identification récursive fondés sur des écritures de modèle plus compactes. En identification hors ligne, l'approche d'état s'est avérée être une alternative intéressante aux représentations externes de type polynomial. Cette approche, qui évite d'utiliser des formes canoniques, présente le grand intérêt de ne nécessiter la connaissance que d'un unique paramètre : l'ordre du système. Ce constat a motivé le développement de nombreux algorithmes d'identification [Vib95], regroupés sous la dénomination de méthodes des sous-espaces, pour identifier directement un modèle d'état du procédé sans avoir recours à une représentation entrée-sortie intermédiaire. Puisque ces techniques ont démontré leur efficacité sur de nombreux cas pratiques [VVG94, WV96, AL97, FDV⁺98, BNSR98, BAB00, BR02], il est intéressant de les adapter à notre problème d'identification récursive. C'est ce que nous allons réaliser au sein des chapitres 4 et 5. Mais, avant cela, il est nécessaire d'exposer une synthèse des méthodes directes d'identification des sous-espaces en insistant plus particulièrement sur les techniques principalement utilisées en identification récursive des sous-espaces : les algorithmes MOESP [VD92a, VD92b, Ver93, Ver94]. Cette présentation nous permettra d'introduire les outils géométriques essentiels à la compréhension des méthodes fondées sur la notion de sous-espace ainsi que de poser les limites intrinsèques de ces techniques hors ligne pour leur application en temps réel.

Chapitre 3

Méthodes d'identification des sous-espaces : une vue d'ensemble des algorithmes MOESP

Sommaire

| | | |
|------------|--|-----------|
| 3.1 | Problématique | 23 |
| 3.2 | Bref historique | 25 |
| 3.3 | « MIMO Output Error State sPace model identification » | 29 |
| 3.3.1 | Algorithme Ordinary MOESP | 30 |
| 3.3.2 | Algorithmes PI et PO MOESP | 37 |
| 3.3.3 | Algorithme MOESP pour les modèles avec « erreurs en les variables » | 42 |
| 3.3.4 | Estimation des matrices d'état | 43 |
| 3.3.5 | Propriétés asymptotiques des algorithmes MOESP et choix des paramètres structuraux | 47 |
| 3.4 | Variable instrumentale et théorèmes unifiés | 53 |
| 3.5 | Conclusion | 59 |

3.1 Problématique

Au cours des deux dernières décennies, de nombreux auteurs [Lar83, De 88, MDVV89, VD91, VD92a, VD92b, Ver94, VD94, Vib95, VD96b] ont proposé plusieurs algorithmes d'identification alternatifs aux techniques traditionnelles d'erreur de prédiction et de maximum de vraisemblance [SS89, Lju99] : les méthodes des sous-espaces. Contrairement aux techniques « classiques » d'identification, les méthodes des sous-espaces ne nécessitent aucun algorithme d'optimisation non linéaire. Elles s'inspirent de la théorie de la réalisation [HK66] (cf. §3.2) et ont pour objectif de fournir un modèle d'état discret, linéaire et invariant du système étudié. Ce modèle est estimé directement à partir des données d'entrée-sortie acquises, sans calcul préalable d'une représentation externe du procédé, en ajustant un ou plusieurs sous-espaces vectoriels aux mesures d'entrée et de sortie. Plus précisément, le problème d'identification traité par les techniques d'identification des sous-espaces consiste à estimer les matrices d'état

déterministes \mathbf{A} , \mathbf{B} , \mathbf{C} et \mathbf{D} du modèle à temps discret du procédé dans le contexte stochastique suivant [VD92a] :

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\tilde{\mathbf{u}}(t) + \mathbf{w}(t) \quad (3.1a)$$

$$\tilde{\mathbf{y}}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\tilde{\mathbf{u}}(t) \quad (3.1b)$$

où $\tilde{\mathbf{u}}(t) \in \mathbb{R}^{n_u \times 1}$ et $\tilde{\mathbf{y}}(t) \in \mathbb{R}^{n_y \times 1}$ sont respectivement les vecteurs des entrées et des sorties non perturbées, $\mathbf{x}(t) \in \mathbb{R}^{n_x \times 1}$ le vecteur d'état et $\mathbf{w}(t) \in \mathbb{R}^{n_x \times 1}$ le bruit d'état. Ces signaux sont supposés être des réalisations (de longueur finie) de variables ergodiques et stochastiques [VD92a]. Les vecteurs de mesure d'entrée et de sortie sont caractérisés de la façon suivante :

$$\mathbf{u}(t) = \tilde{\mathbf{u}}(t) + \mathbf{v}(t) \quad (3.2a)$$

$$\mathbf{y}(t) = \tilde{\mathbf{y}}(t) + \mathbf{v}(t) \quad (3.2b)$$

avec $\mathbf{v}(t) \in \mathbb{R}^{n_u \times 1}$ et $\mathbf{v}(t) \in \mathbb{R}^{n_y \times 1}$ des bruits de mesure. La figure 3.1 illustre l'influence de ces perturbations sur le système. Les trois bruits considérés sont supposés être trois vecteurs de

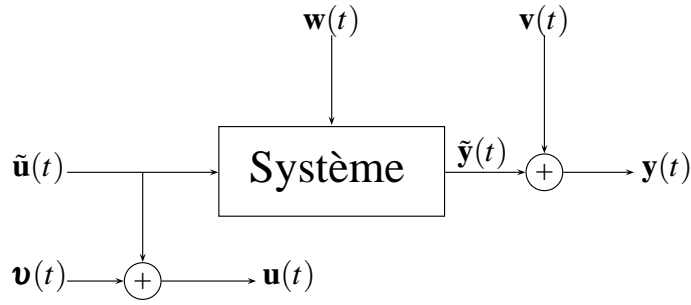


FIG. 3.1: Représentation du système étudié.

bruit blanc statistiquement indépendants des entrées passées non bruitées $\tilde{\mathbf{u}}$ tels que :

$$\mathbb{E} \left\{ \begin{bmatrix} \mathbf{w}(t) \\ \mathbf{v}(t) \\ \mathbf{v}(t) \end{bmatrix} \begin{bmatrix} \mathbf{w}(s) & \mathbf{v}(s) & \mathbf{v}(s) \end{bmatrix} \right\} = \begin{bmatrix} \mathbf{R}_{\mathbf{w}\mathbf{w}} & \mathbf{R}_{\mathbf{w}\mathbf{v}} & \mathbf{R}_{\mathbf{w}\mathbf{v}} \\ \mathbf{R}_{\mathbf{w}\mathbf{v}}^T & \mathbf{R}_{\mathbf{v}\mathbf{v}} & \mathbf{R}_{\mathbf{v}\mathbf{v}} \\ \mathbf{R}_{\mathbf{v}\mathbf{w}}^T & \mathbf{R}_{\mathbf{v}\mathbf{v}}^T & \mathbf{R}_{\mathbf{v}\mathbf{v}} \end{bmatrix} \delta_{t,s} \quad (3.3)$$

avec $\delta_{t,s}$ le symbole de Kronecker. Nous supposons que les bruits \mathbf{w} , \mathbf{v} et \mathbf{v} sont indépendants de la séquence d'entrée $\tilde{\mathbf{u}}$, que les bruits de mesure \mathbf{v} et \mathbf{v} sont indépendants du vecteur \mathbf{x} et que le bruit de procédé $\mathbf{w}(t)$ est indépendant de l'état initial. Nous faisons également l'hypothèse que la paire $(\mathbf{A}, [\mathbf{B} \mathbf{R}_{\mathbf{w}\mathbf{w}}^{1/2}])$ est commandable, que la paire (\mathbf{A}, \mathbf{C}) est observable et que le système est asymptotiquement stable à non minimum de phase.

De nombreux algorithmes (CVA [Lar90], N4SID [VD94], Ordinary MOESP [VD92a], PI MOESP [Ver93], PO MOESP [Ver94], PO_EIV MOESP [CV97]...) ont été développés pour estimer, de manière consistante¹, les matrices du modèle d'état (3.1) dans le contexte

¹Les estimées $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$ et $\hat{\mathbf{D}}$ sont consistantes s'il existe une matrice \mathbf{T} déterministe et non singulière telle que $\hat{\mathbf{A}} - \mathbf{T}\mathbf{A}\mathbf{T}^{-1}$, $\hat{\mathbf{B}} - \mathbf{T}\mathbf{B}$, $\hat{\mathbf{C}} - \mathbf{C}\mathbf{T}^{-1}$ et $\hat{\mathbf{D}} - \mathbf{D}$ convergent vers zéro avec une probabilité de 1. Cette notion sera plus particulièrement étudiée au sein du paragraphe 3.3.5.

stochastique défini précédemment. Chacun d'eux considère des hypothèses particulières sur les propriétés des entrées et des perturbations agissant sur le système. Les techniques regroupées sous l'acronyme MOESP sont plus particulièrement détaillées au sein de ce chapitre (cf. §3.3). Ce choix est réalisé en corrélation avec l'objectif de cette thèse : développer de nouvelles méthodes récursives d'identification des sous-espaces. En effet, ces dernières s'inspirent principalement de la démarche suivie par les algorithmes MOESP, à savoir :

- estimer la matrice d'observabilité du système de manière consistante ;
- estimer les matrices d'état du système à partir de l'estimée précédente et des données d'entrée-sortie, *sans estimation du vecteur d'état*.

Le plan d'étude de ce chapitre est le suivant : après une brève description (cf. §3.2) des techniques originelles des sous-espaces fondées sur la théorie de la réalisation, une présentation détaillée des algorithmes MOESP est proposée au sein de la section 3.3 afin d'introduire les notations, les outils et les notions de base nécessaires à la compréhension des méthodes récursives des sous-espaces décrites dans les chapitres 4 et 5. Cette analyse conduira à une description des principales techniques d'extraction des matrices d'état du système associées à la classe d'identification MOESP. Un bref exposé des propriétés asymptotiques des algorithmes MOESP et des principaux algorithmes d'estimation de l'ordre du système conclura le paragraphe 3.3. Une discussion sur le rôle central joué par la variable instrumentale parachèvera l'étude des méthodes hors ligne des sous-espaces.

3.2 Bref historique

Historiquement, les méthodes « modernes » (nommée directes² dans la suite de ce manuscrit) d'identification des sous-espaces reposent sur l'amélioration de techniques d'estimation fondées sur la théorie de la réalisation. Ces dernières, référencées ici comme méthodes indirectes, ont pour objectif de fournir une représentation d'état minimale³ du système à temps discret :

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \quad (3.4a)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \quad (3.4b)$$

connaissant une représentation externe particulière du procédé : un ensemble de réponses impulsionnelles $\{\hat{h}(t)\}$ sur un horizon donné [HK66, Kai80]. Considérons donc que nous ayons accès à $i + j - 1$ matrices $\bar{\mathbf{H}} \in \mathbb{R}^{n_y \times n_u}$ (ou paramètres de Markov) composées de réponses impulsionnelles du système telles que :

$$\bar{\mathbf{H}}(t) = \begin{bmatrix} \hat{h}_{11}(t) & \hat{h}_{12}(t) & \cdots & \hat{h}_{1n_u}(t) \\ \hat{h}_{21}(t) & \hat{h}_{22}(t) & \cdots & \hat{h}_{2n_u}(t) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{h}_{n_y1}(t) & \hat{h}_{n_y2}(t) & \cdots & \hat{h}_{n_y n_u}(t) \end{bmatrix} \quad (3.5)$$

²Nous utiliserons l'adjectif « directe » pour qualifier les méthodes des sous-espaces fournissant *directement* une représentation d'état du système, sans nécessiter de connaître *a priori* une représentation externe du procédé, par opposition à « indirecte » pour les techniques basées sur l'utilisation des réponses impulsionnelles [Vib95, Bas97].

³Une représentation d'état est dite minimale s'il n'existe aucune réalisation de degré inférieur accessible. Une telle forme est à la fois commandable et observable [Bro91].

où $\tilde{h}_{k\ell}(t)$ est le t^e échantillon de la réponse impulsionnelle entre la k^e sortie et la ℓ^e entrée. Il est alors possible de construire la matrice de Hankel suivante [Vib95] :

$$\mathbf{H}_{i,j} = \begin{bmatrix} \bar{\mathbf{H}}(1) & \bar{\mathbf{H}}(2) & \cdots & \bar{\mathbf{H}}(j) \\ \bar{\mathbf{H}}(2) & \bar{\mathbf{H}}(3) & \cdots & \bar{\mathbf{H}}(j+1) \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\mathbf{H}}(i) & \bar{\mathbf{H}}(i+1) & \cdots & \bar{\mathbf{H}}(i+j-1) \end{bmatrix} \in \mathbb{R}^{n_y i \times n_u j} \quad (3.6)$$

pour laquelle :

$$\bar{\mathbf{H}}(t) = \begin{cases} \mathbf{0} & \text{si } t < 0, \\ \mathbf{D} & \text{si } t = 0, \\ \mathbf{C}\mathbf{A}^{t-1}\mathbf{B} & \text{si } t > 0. \end{cases} \quad (3.7)$$

Les méthodes indirectes des sous-espaces sont alors basées sur la propriété suivante [HK66] :

$$\hat{\mathbf{H}}_{i,j} = \mathbf{\Gamma}_i \mathbf{U}_j, \quad (3.8)$$

factorisation matricielle correspondant au produit des matrices d'observabilité et de commandabilité étendues du procédé :

$$\mathbf{\Gamma}_i = \begin{bmatrix} \mathbf{C} \\ \mathbf{C}\mathbf{A} \\ \vdots \\ \mathbf{C}\mathbf{A}^{i-1} \end{bmatrix} \in \mathbb{R}^{n_y i \times n_x} \quad (3.9)$$

$$\mathbf{U}_j = [\mathbf{B} \quad \mathbf{A}\mathbf{B} \quad \cdots \quad \mathbf{A}^{j-1}\mathbf{B}] \in \mathbb{R}^{n_x \times n_u j}. \quad (3.10)$$

Cette écriture permet ainsi de déterminer l'ordre du système puisque, pour toute réalisation d'état minimale [HK66] :

$$\text{rang}(\mathbf{H}_{i,j}) = n_x. \quad (3.11)$$

De plus, en supposant que nous ayons accès à un couple d'estimées $\{\hat{\mathbf{\Gamma}}_i, \hat{\mathbf{U}}_j\}$ tel que :

$$\mathbf{H}_{i,j} = \hat{\mathbf{\Gamma}}_i \hat{\mathbf{U}}_j, \quad (3.12)$$

il est aisé d'estimer les matrices \mathbf{B} , \mathbf{C} et \mathbf{D} en remarquant que [ZM74] :

$$\mathbf{H}(0) = \mathbf{D}, \quad (3.13)$$

$$\mathbf{B} = \text{les } n_u \text{ premières colonnes de } \mathbf{U}_j, \quad (3.14)$$

$$\mathbf{C} = \text{les } n_y \text{ premières lignes de } \mathbf{\Gamma}_i. \quad (3.15)$$

L'estimation de \mathbf{A} est quant à elle réalisée en utilisant la propriété d'*A-invariance* de $\mathbf{\Gamma}_i$ [Kun78] :

$$\mathbf{\Gamma}_i^\downarrow = \mathbf{\Gamma}_i^\uparrow \mathbf{A} \quad (3.16)$$

où :

$$\Gamma_i^\uparrow = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{i-2} \end{bmatrix} = \Gamma_i(1 : n_y(i-1), :) \quad \text{et} \quad \Gamma_i^\downarrow = \begin{bmatrix} \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^{i-1} \end{bmatrix} = \Gamma_i(n_y + 1 : n_y i, :). \quad (3.17)$$

La matrice d'état est alors accessible en résolvant :

$$\hat{\mathbf{A}} = \left\{ \Gamma_i^\uparrow \right\}^\dagger \Gamma_i^\downarrow \quad (3.18)$$

où $\left\{ \Gamma_i^\uparrow \right\}^\dagger$ est la pseudo inverse de Moore Penrose [Rot01] de Γ_i^\uparrow . Cette estimation est légitime puisque, si la représentation est minimale, la matrice Γ_i^\uparrow est non singulière. Nous reviendrons plus en détail sur l'estimation de la matrice \mathbf{A} à partir de la propriété d'A-invariance de Γ_i dans la suite de ce chapitre (cf. §3.3.4).

Cette méthode d'identification nécessite de choisir les indices i et j tels que :

$$i > n_x \quad \text{et} \quad j > n_x \quad (3.19)$$

puisque, dans le cas contraire, l'estimation de l'ordre, et par conséquent des matrices Γ_i et \mathbf{U}_j , ne peut qu'échouer. Il est donc fondamental de surdimensionner la matrice de Hankel (3.6). Or, puisque les données d'entrée-sortie sont mesurées, la matrice de Hankel contient des réponses impulsionnelles bruitées. Le fait d'introduire des données bruitées dans une matrice surdimensionnée a pour conséquence d'augmenter le rang de la matrice ainsi construite. Il est alors nécessaire d'associer à l'estimation des matrices Γ_i et \mathbf{U}_j une étape de réduction de rang. Pour résoudre ce problème de compression de données, deux auteurs [ZM74, Kun78] ont successivement proposé d'appliquer à $\mathbf{H}_{i,j}$ un outil mathématique particulier : la décomposition en valeurs singulières (DVS) [Rot01] :

$$\mathbf{H}_{i,j} = \bar{\mathbf{U}} \bar{\mathbf{\Sigma}} \bar{\mathbf{V}}^T = \begin{bmatrix} \bar{\mathbf{U}}_s & \bar{\mathbf{U}}_b \end{bmatrix} \begin{bmatrix} \bar{\mathbf{\Sigma}}_s & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{\Sigma}}_b \end{bmatrix} \begin{bmatrix} \bar{\mathbf{V}}_s^T \\ \bar{\mathbf{V}}_b^T \end{bmatrix} \quad (3.20)$$

pour laquelle $\bar{\mathbf{U}}_s$ et $\bar{\mathbf{V}}_s$ sont les vecteurs singuliers gauche et droit correspondant aux n_x plus grandes valeurs singulières $\bar{\mathbf{\Sigma}}_s$ de $\bar{\mathbf{\Sigma}}$. $\bar{\mathbf{U}}_b$ et $\bar{\mathbf{V}}_b$ sont les vecteurs singuliers gauche et droit correspondant aux valeurs singulières complémentaires $\bar{\mathbf{\Sigma}}_b$. Le produit matriciel $\bar{\mathbf{U}}_s \bar{\mathbf{\Sigma}}_s \bar{\mathbf{V}}_s^T$ sera nommé par la suite « sous-espace signal », l'espace orthogonal $\bar{\mathbf{U}}_b \bar{\mathbf{\Sigma}}_b \bar{\mathbf{V}}_b^T$ « sous-espace bruit ». Il est évident que ce dernier est vide en l'absence de bruit puisque $\bar{\mathbf{\Sigma}}_b = \mathbf{0}$. La décomposition suivante est alors considérée :

$$\hat{\Gamma}_i = \bar{\mathbf{U}}_s \bar{\mathbf{\Sigma}}_s^{1/2} \quad (3.21)$$

$$\hat{\mathbf{U}}_j = \bar{\mathbf{\Sigma}}_s^{1/2} \bar{\mathbf{V}}_s^T. \quad (3.22)$$

Les matrices d'état du système sont, comme précédemment, extraites à partir de ces estimées (cf. equ. (3.13) à (3.16)).

Ces techniques, efficaces théoriquement, présentent le sévère désavantage de nécessiter de connaître *a priori* les paramètres de Markov du procédé. Or, obtenir des mesures ou des estimées fiables des réponses impulsionnelles d'un système n'est pas simple pratiquement. Même

si quelques auteurs ont développé des solutions ingénieuses pour estimer les paramètres de Markov à partir de représentations ou d'analyses particulières [Ise80, SS89, LS92, McK95, BGRS97, Lju99], un important progrès fut réalisé lorsque d'autres tentèrent d'estimer directement une représentation d'état du système à partir des mesures d'entrée-sortie, sans requérir une représentation externe. Trois familles d'algorithmes se distinguent parmi les méthodes directes des sous-espaces :

- « Canonical Variate Analysis » (CVA) ou « Canonical Correlation Analysis » (CCA) [Lar83, Lar90] ;
- « Numerical algorithm for SubSpace State Space IDentification » (N4SID) [VD93, VD94, VD96b] ;
- « Multivariable Output Error State sPace » (MOESP) [VD91, Ver93, Ver94].

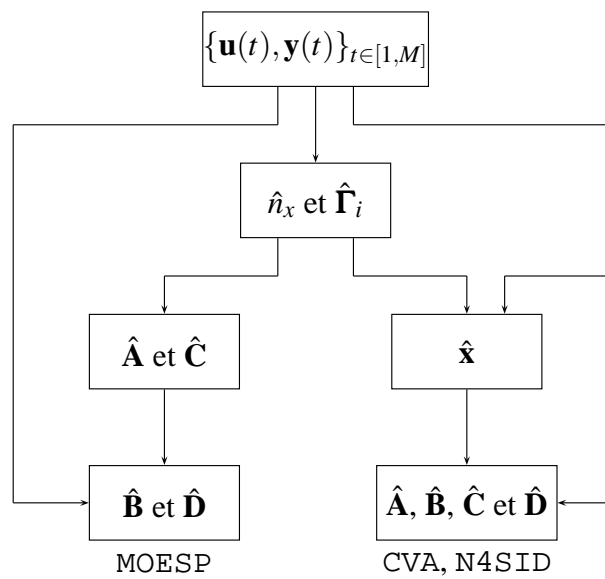


FIG. 3.2: Étapes d'identification des méthodes directes des sous-espaces.

Elles recherchent toutes trois à estimer un espace ligne ou colonne particulier en ajustant ce sous-espace aux données d'entrée-sortie à l'aide d'outils mathématiques robustes et de constatations géométriques simples. Elles se divisent en deux classes [VD95] (cf. fig. 3.2) :

- Les algorithmes MOESP recherchent une estimée consistante de la matrice d'observabilité étendue du système à partir des données d'entrée-sortie mesurées. Les matrices d'état du système sont alors extraites en deux temps, en utilisant la propriété d'A-invariance de la matrice d'observabilité étendue pour **A** et **C** puis une régression linéaire classique pour **B** et **D** [VD92a, VD92b, Ver93, Ver94, McK95].
- CVA et N4SID cherchent également à estimer la matrice d'observabilité étendue du procédé mais y ajoute une phase d'estimation de l'état du système à l'aide, respectivement, d'une analyse de corrélation [Lar83, Lar90] et d'une projection oblique [VD94, VD96b]. En ajoutant cette étape d'estimation, l'obtention des matrices d'état est alors plus aisément réalisée que dans le cas précédent puisque celles-ci sont calculées à partir du système surdimensionné suivant :

$$\begin{bmatrix} \hat{\mathbf{x}}(t+1) & \dots & \hat{\mathbf{x}}(t+M) \\ \mathbf{y}(t) & \dots & \mathbf{y}(t+M-1) \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}(t) & \dots & \hat{\mathbf{x}}(t+M-1) \\ \mathbf{u}(t) & \dots & \mathbf{u}(t+M-1) \end{bmatrix} \quad (3.23)$$

construit en réinjectant les estimées du vecteur d'état au sein de la représentation (3.1). Les matrices $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$ et $\hat{\mathbf{D}}$ sont ainsi obtenues en une seule phase à l'aide d'un critère quadratique minimisable via, par exemple, les moindres carrés [VD96b, Lju99].

Dans la suite de cette section, nous nous intéressons plus particulièrement aux algorithmes MOESP. Ce choix se justifie par le fait que les méthodes récursives des sous-espaces présentées dans les deux prochains chapitres ne cherchent en aucun cas à estimer le vecteur d'état mais adoptent la démarche proposée par M. Verhaegen : estimer uniquement la matrice d'observabilité étendue puis extraire, en deux temps, les matrices d'état du procédé.

3.3 « MIMO Output Error State sPACE model identification »

Les algorithmes d'identification MOESP ont pour principal objectif d'estimer, de manière consistante, la partie déterministe du modèle du système étudié dans un contexte stochastique, i.e. estimer les matrices \mathbf{A} , \mathbf{B} , \mathbf{C} , et \mathbf{D} du modèle sachant que le système peut être perturbé par des bruits d'état \mathbf{w} et de mesure \mathbf{v} en entrée et \mathbf{v} en sortie du procédé (cf. fig. 3.1) :

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\tilde{\mathbf{u}}(t) + \mathbf{w}(t) \quad (3.24a)$$

$$\tilde{\mathbf{y}}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\tilde{\mathbf{u}}(t) \quad (3.24b)$$

$$\mathbf{u}(t) = \tilde{\mathbf{u}}(t) + \mathbf{v}(t) \quad (3.24c)$$

$$\mathbf{y}(t) = \tilde{\mathbf{y}}(t) + \mathbf{v}(t). \quad (3.24d)$$

La classe d'identification MOESP a comme caractéristique [Ver94] d'être décomposable en deux étapes successives :

- une compression des matrices de Hankel d'entrée et de sortie à l'aide d'un ensemble de transformations orthogonales aboutissant à l'extraction de matrices particulières ;
- l'application d'une décomposition en valeurs singulières à certaines matrices de la factorisation précédente conduisant à une estimation consistante de la matrice d'observabilité étendue.

Les algorithmes MOESP diffèrent selon les hypothèses fixées sur les propriétés des entrées et des perturbations. Ainsi :

- Ordinary MOESP suppose que les bruits d'état \mathbf{w} et de mesure \mathbf{v} sont nuls et que le bruit de sortie \mathbf{v} est un bruit blanc gaussien (BB) de moyenne nulle et de variance finie indépendant de l'entrée non bruitée $\tilde{\mathbf{u}}$;
- Past Input (PI) MOESP permet de traiter deux types de système :
 - * les systèmes soumis à un bruit de sortie \mathbf{v} coloré (BC) de variance finie, les bruits d'état \mathbf{w} et de mesure \mathbf{v} étant nuls,
 - * les systèmes à bruit d'état nul perturbés par un bruit de sortie \mathbf{v} coloré (BC) de variance finie et excités par une entrée $\tilde{\mathbf{u}}$ blanche et gaussienne de moyenne nulle et de variance finie dont la mesure est perturbée par un bruit \mathbf{v} lui-même blanc et gaussien de moyenne nulle et de variance finie ;
- Past Output (PO) MOESP est utilisable dans deux situations :
 - * le bruit de mesure \mathbf{v} est nul et les bruits d'état \mathbf{w} et de sortie \mathbf{v} sont deux bruits blancs gaussiens de moyenne nulle et de variance finie,
 - * les bruits de mesure \mathbf{v} et \mathbf{v} , le bruit d'état \mathbf{w} et la séquence d'entrée $\tilde{\mathbf{u}}$ sont des bruits blancs gaussiens de moyenne nulle et de variance finie.

- Past Output Error-In-Variables (PO_EIV) MOESP fournit, quant à lui, des estimées consistantes lorsque \mathbf{v} , \mathbf{v} et \mathbf{w} sont trois bruits blancs gaussiens de moyenne nulle et de variance finie quelles que soient les caractéristiques de la séquence d'entrée $\tilde{\mathbf{u}}$.

Tous supposent bien évidemment que la séquence d'entrée vérifie la condition d'excitation persistante (CEP) [SS89, LLM97, Lju99] (cf. §3.3.5).

| Méthode | $\tilde{\mathbf{u}}$ | \mathbf{v} | \mathbf{w} | \mathbf{v} |
|----------------|----------------------|--------------|--------------|--------------|
| Ordinary MOESP | CEP | BB | $\mathbf{0}$ | $\mathbf{0}$ |
| PI MOESP | CEP | BC | $\mathbf{0}$ | $\mathbf{0}$ |
| | BB | BC | $\mathbf{0}$ | BB |
| PO MOESP | CEP | BB | BB | $\mathbf{0}$ |
| | BB | BB | BB | BB |
| PO_EIV MOESP | CEP | BB | BB | BB |

TAB. 3.1: Conditions d'application des méthodes MOESP. CEP : condition d'excitation persistante ; BB : bruit blanc ; BC : bruit coloré.

Ces différentes situations sont traitées au sein des paragraphes suivants. Leur objectif est de présenter les fondements de ces algorithmes en insistant sur les outils mathématiques utilisés pour résoudre la problématique posée, ces outils étant, pour la plupart, appliqués ou adaptés en identification récursive des sous-espaces.

3.3.1 Algorithme Ordinary MOESP

L'algorithme Ordinary MOESP [VD92a, Ver93] traite le problème de la modélisation par représentation d'état de systèmes à erreur de sortie de moyenne nulle et de matrice de covariance finie $\sigma^2\mathbf{I}$. Considérons donc la forme d'état suivante⁴ :

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \quad (3.25a)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) + \mathbf{v}(t) \quad (3.25b)$$

et introduisons le vecteur de sorties décalées suivant :

$$\mathbf{y}_f^+(t) = \begin{bmatrix} \mathbf{y}(t) \\ \vdots \\ \mathbf{y}(t+f-1) \end{bmatrix} \in \mathbb{R}^{n_y f \times 1}. \quad (3.26)$$

Il est alors facile de montrer que ce vecteur vérifie :

$$\mathbf{y}_f^+(t) = \mathbf{\Gamma}_f \mathbf{x}(t) + \mathbf{H}_f \mathbf{u}_f^+(t) + \mathbf{v}_f^+(t) \quad (3.27)$$

⁴Puisque $\mathbf{v} = \mathbf{0}$, $\tilde{\mathbf{u}}(t) = \mathbf{u}(t)$.

pour laquelle $\mathbf{u}_f^+ \in \mathbb{R}^{n_u f \times 1}$ et $\mathbf{v}_f^+ \in \mathbb{R}^{n_y f \times 1}$ sont construits de manière similaire à \mathbf{y}_f^+ . $\mathbf{\Gamma}_f \in \mathbb{R}^{n_y f \times n_x}$ est la matrice d'observabilité d'ordre f du système :

$$\mathbf{\Gamma}_f = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{f-1} \end{bmatrix}. \quad (3.28)$$

La matrice $\mathbf{H}_f \in \mathbb{R}^{n_y f \times n_u f}$ est une matrice de Toeplitz par blocs contenant les paramètres de Markov du procédé :

$$\mathbf{H}_f = \begin{bmatrix} \mathbf{D} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{CB} & \mathbf{D} & \cdots & \mathbf{0} \\ \mathbf{CAB} & \mathbf{CB} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{CA}^{f-2}\mathbf{B} & \mathbf{CA}^{f-3}\mathbf{B} & \cdots & \mathbf{D} \end{bmatrix}. \quad (3.29)$$

En supposant que nous ayons accès à un grand nombre de données d'entrée-sortie, par itérations successives sur l'indice temporel des vecteurs avancés, l'expression (3.27) peut être étendue à l'équation matricielle suivante :

$$\mathbf{Y}_{t,f,M}^+(\bar{t}) = \mathbf{\Gamma}_f \mathbf{X}_{t,M}^+(\bar{t}) + \mathbf{H}_f \mathbf{U}_{t,f,M}^+(\bar{t}) + \mathbf{V}_{t,f,M}^+(\bar{t}) \quad (3.30)$$

pour laquelle :

$$\begin{aligned} \mathbf{Y}_{t,f,M}^+(\bar{t}) &= [\mathbf{y}_f^+(t) \cdots \mathbf{y}_f^+(\underbrace{t+M-1}_{\bar{t}})] \\ &= \begin{bmatrix} \mathbf{y}(t) & \mathbf{y}(t+1) & \cdots & \mathbf{y}(\bar{t}) \\ \mathbf{y}(t+1) & \mathbf{y}(t+2) & \cdots & \mathbf{y}(\bar{t}+1) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{y}(t+f-1) & \mathbf{y}(t+f) & \cdots & \mathbf{y}(\bar{t}+f-1) \end{bmatrix} \in \mathbb{R}^{n_y f \times M}, \end{aligned} \quad (3.31)$$

$$\mathbf{U}_{t,f,M}^+(\bar{t}) = \begin{bmatrix} \mathbf{u}(t) & \mathbf{u}(t+1) & \cdots & \mathbf{u}(\bar{t}) \\ \mathbf{u}(t+1) & \mathbf{u}(t+2) & \cdots & \mathbf{u}(\bar{t}+1) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}(t+f-1) & \mathbf{u}(t+f) & \cdots & \mathbf{u}(\bar{t}+f-1) \end{bmatrix} \in \mathbb{R}^{n_u f \times M}, \quad (3.32)$$

$$\mathbf{V}_{t,f,M}^+(\bar{t}) = \begin{bmatrix} \mathbf{v}(t) & \mathbf{v}(t+1) & \cdots & \mathbf{v}(\bar{t}) \\ \mathbf{v}(t+1) & \mathbf{v}(t+2) & \cdots & \mathbf{v}(\bar{t}+1) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}(t+f-1) & \mathbf{v}(t+f) & \cdots & \mathbf{v}(\bar{t}+f-1) \end{bmatrix} \in \mathbb{R}^{n_y f \times M} \quad (3.33)$$

sont les matrices de Hankel de sortie, d'entrée et de bruit accessibles à l'instant $\bar{t} = t + M - 1$ et :

$$\mathbf{X}_{t,M}^+(\bar{t}) = [\mathbf{x}(t) \quad \mathbf{x}(t+1) \quad \cdots \quad \mathbf{x}(\bar{t})] \in \mathbb{R}^{n_x \times M} \quad (3.34)$$

la matrice contenant l'ensemble des vecteurs d'état sur l'horizon M .

Remarque 3.1. Afin d'alléger les équations, sauf mention contraire, nous utiliserons dans la suite de cette section l'expression condensée suivante :

$$\mathbf{Y}_f^+ = \mathbf{\Gamma}_f \mathbf{X}^+ + \mathbf{H}_f \mathbf{U}_f^+ + \mathbf{V}_f^+ \quad (3.35)$$

en posant :

$$\mathbf{Y}_{t,f,M}^+(\bar{t}) = \mathbf{Y}_f^+ \quad (3.36)$$

$$\mathbf{U}_{t,f,M}^+(\bar{t}) = \mathbf{U}_f^+ \quad (3.37)$$

$$\mathbf{V}_{t,f,M}^+(\bar{t}) = \mathbf{V}_f^+ \quad (3.38)$$

$$\mathbf{X}_{t,M}^+(\bar{t}) = \mathbf{X}^+. \quad (3.39)$$

Les indices t, f, M n'ont été utilisés que pour améliorer la compréhension de la construction des matrices de Hankel. La variable \bar{t} sera utilisée lors de la présentation des algorithmes récursifs d'identification.

L'objectif de l'algorithme Ordinary MOESP est de fournir une estimation consistante de $\mathbf{\Gamma}_f$ en supposant que le signal d'entrée appliqué au procédé vérifie la condition d'excitation persistante (cf. §3.3.5) *i.e.* :

$$\text{rang} \left(\lim_{M \rightarrow \infty} \mathbf{U}_f^+ \mathbf{U}_f^{+T} \right) = n_{uf}, \quad (3.40)$$

que la condition de rang suivante [VD92a] :

$$\text{rang} \left(\begin{bmatrix} \mathbf{U}_f^+ \\ \mathbf{X}^+ \end{bmatrix} \right) = n_{uf} + n_x \quad (3.41)$$

soit satisfaite et que le bruit de sortie soit de moyenne nulle et de matrice de covariance proportionnelle à l'identité. Or, pour atteindre ce but, il est nécessaire de résoudre trois problèmes mis en évidence par l'équation (3.35) :

- la matrice de Toeplitz par blocs \mathbf{H}_f est inconnue *a priori* puisque celle-ci est composée des paramètres de Markov du procédé. Il est donc nécessaire de supprimer le régime forcé $\mathbf{H}_f \mathbf{U}_f^+$;
- la matrice de Hankel du bruit de sortie doit être éliminée afin de garantir des estimées asymptotiquement non biaisées ;
- la matrice d'observabilité doit être estimée à partir des matrices résultant des traitements précédents.

Ces trois questions sont successivement abordées dans les trois paragraphes suivants.

3.3.1.1 Traitement de $\mathbf{H}_f \mathbf{U}_f^+$

Considérons le minimiseur du critère suivant⁵ [Vib95] :

$$\arg \min_{\mathbf{H}_f} \left\| \mathbf{Y}_f^+ - \mathbf{H}_f \mathbf{U}_f^+ \right\|_F^2 \quad (3.42)$$

⁵ $\|\cdot\|_F^2$ symbolise la norme de Frobenius [Rot01].

traduisant l'estimation de \mathbf{H}_f à partir des données d'entrée-sortie regroupées au sein des matrices de Hankel \mathbf{U}_f^+ et \mathbf{Y}_f^+ . La minimisation d'un tel critère conduit alors à la solution (optimale au sens des moindres carrés) $\hat{\mathbf{H}}_f$ qui vérifie :

$$\mathbf{Y}_f^+ - \hat{\mathbf{H}}_f \mathbf{U}_f^+ = \mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \quad (3.43)$$

où $\mathbf{\Pi}_{\mathbf{U}_f^{+\perp}}$ symbolise la projection orthogonale sur le noyau de \mathbf{U}_f^+ . Le calcul de cette projection orthogonale est réalisable efficacement en appliquant une factorisation QR [Rot01] au bloc matriciel $\begin{bmatrix} \mathbf{U}_f^{+T} & \mathbf{Y}_f^{+T} \end{bmatrix}^T$:

$$\begin{bmatrix} \mathbf{U}_f^+ \\ \mathbf{Y}_f^+ \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_{11} & \mathbf{0} \\ \bar{\mathbf{R}}_{21} & \bar{\mathbf{R}}_{22} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1 \\ \bar{\mathbf{Q}}_2 \end{bmatrix} \quad (3.44)$$

où $\bar{\mathbf{R}}_{11} \in \mathbb{R}^{n_{uf} \times n_{uf}}$ et $\bar{\mathbf{R}}_{22} \in \mathbb{R}^{n_{yf} \times n_{yf}}$ sont des matrices triangulaires inférieures, $\bar{\mathbf{R}}_{21} \in \mathbb{R}^{n_{yf} \times n_{uf}}$ est une matrice pleine et $\bar{\mathbf{Q}}_1 \in \mathbb{R}^{n_{uf} \times M}$, $\bar{\mathbf{Q}}_2 \in \mathbb{R}^{n_{yf} \times M}$ des matrices orthogonales. En effet, il est facile de montrer que (cf. §3.4) :

$$\mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} = \bar{\mathbf{R}}_{22} \bar{\mathbf{Q}}_2. \quad (3.45)$$

L'application d'une telle projection à l'équation (3.30) conduit à la solution suivante :

$$\begin{aligned} \mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} &= \bar{\mathbf{R}}_{22} \bar{\mathbf{Q}}_2 = \\ & \Gamma_f \mathbf{X}^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} + \underbrace{\mathbf{H}_f \mathbf{U}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} + \mathbf{V}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}}}_{=0} = \Gamma_f \mathbf{X}^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} + \mathbf{V}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}}. \end{aligned} \quad (3.46)$$

L'utilisation de cet outil mathématique permet ainsi de supprimer $\mathbf{H}_f \mathbf{U}_f^+$ dû à l'excitation des entrées du procédé et de conserver les composantes des vecteurs de sorties résultantes de l'état et du bruit.

3.3.1.2 Traitement de \mathbf{V}_f^+

Une fois le régime forcé supprimé, l'étape suivante consiste à isoler \mathbf{V}_f^+ de $\mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}}$. Pour cela, remarquons tout d'abord que :

$$\lim_{M \rightarrow \infty} \mathbf{V}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} = \mathbf{V}_f^+ \quad (3.47)$$

puisque le bruit de sortie et la séquence d'entrée sont décorrélés⁶. Ainsi, en supposant que le nombre de données accessibles soit suffisamment grand :

$$\mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} = \bar{\mathbf{R}}_{22} \bar{\mathbf{Q}}_2 = \Gamma_f \mathbf{X}^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} + \mathbf{V}_f^+. \quad (3.48)$$

L'outil mathématique de prédilection à appliquer pour séparer deux espaces imbriqués est la décomposition en valeurs singulières [Rot01]. En effet, par simple observation des valeurs

⁶Rappelons que l'identification est réalisée en boucle ouverte.

singulières, il est possible d'isoler les vecteurs singuliers correspondant aux sous-espaces recherchés. L'application d'une telle opération sur la matrice⁷ $\bar{\mathbf{R}}_{22}\bar{\mathbf{Q}}_2$ conduit à l'expression suivante :

$$\begin{aligned} \bar{\mathbf{R}}_{22}\bar{\mathbf{Q}}_2 &= \bar{\mathbf{U}} \begin{bmatrix} \bar{\sigma}_1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & & & & \vdots \\ \vdots & & \bar{\sigma}_{n_x} & & & \vdots \\ \vdots & & & \sigma & & \vdots \\ \vdots & & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & \sigma \end{bmatrix} \bar{\mathbf{V}}^T \\ &= [\bar{\mathbf{U}}_s \quad \bar{\mathbf{U}}_b] \begin{bmatrix} \bar{\boldsymbol{\Sigma}}_s & \mathbf{0} \\ \mathbf{0} & \bar{\boldsymbol{\Sigma}}_b \end{bmatrix} \begin{bmatrix} \bar{\mathbf{V}}_s^T \\ \bar{\mathbf{V}}_b^T \end{bmatrix} = \bar{\mathbf{U}}_s \bar{\boldsymbol{\Sigma}}_s \bar{\mathbf{V}}_s^T + \bar{\mathbf{U}}_b \bar{\boldsymbol{\Sigma}}_b \bar{\mathbf{V}}_b^T \end{aligned} \quad (3.49)$$

où $\bar{\boldsymbol{\Sigma}}_b$ contient les $n_y f - n_x$ plus faibles valeurs singulières (égales à σ). Puisque $\bar{\mathbf{U}}_b \bar{\boldsymbol{\Sigma}}_b \bar{\mathbf{V}}_b^T$ correspond au sous-espace bruit [VD92b], par analogie avec l'équation (3.48), le sous-espace conservé est $\bar{\mathbf{U}}_s \bar{\boldsymbol{\Sigma}}_s \bar{\mathbf{V}}_s^T$. Ce dernier correspond au sous-espace signal :

$$\bar{\mathbf{U}}_s \bar{\boldsymbol{\Sigma}}_s \bar{\mathbf{V}}_s^T = \boldsymbol{\Gamma}_f \mathbf{X}^+ \boldsymbol{\Pi}_{\mathbf{U}_f^{+\perp}}. \quad (3.50)$$

Il permet ainsi de supprimer la matrice des perturbations \mathbf{V}_f^+ et de ne conserver que les composantes des vecteurs de sortie engendrées par l'état du système.

Remarque 3.2. *L'inspection des valeurs singulières permet d'avoir accès à une estimée de l'ordre du système. En effet, lorsque la variance du bruit est raisonnable, le tracé des valeurs singulières dans un ordre décroissant présente théoriquement une chute visible entre la n_x^e et $(n_x + 1)^e$ valeur [Ver94, VD96b, Bau01] (cf. §3.3.5.2). Cette propriété est un avantage des méthodes des sous-espaces puisqu'elle permet d'estimer l'ordre du système au cours du processus d'identification. L'ordre n est donc plus un paramètre inconnu que l'utilisateur doit fixer a priori. Un simple majorant de ce dernier est nécessaire pour identifier le procédé.*

3.3.1.3 Extraction de $\boldsymbol{\Gamma}_f$

La dernière phase de l'algorithme consiste à extraire la matrice d'observabilité étendue du sous-espace signal obtenu précédemment. Elle réside plus précisément en l'estimation d'une base de $\boldsymbol{\Gamma}_f$, extrapolation se justifiant par le fait que la représentation d'état est plurielle et que toute estimation est réalisée à un changement de base près [Bro91]. Considérons donc l'équation (3.48) dans le cas déterministe⁸ :

$$\mathbf{Y}_f^+ \boldsymbol{\Pi}_{\mathbf{U}_f^{+\perp}} = \boldsymbol{\Gamma}_f \mathbf{X}^+ \boldsymbol{\Pi}_{\mathbf{U}_f^{+\perp}}. \quad (3.51)$$

En supposant que u soit suffisamment excitée (condition d'excitation persistante vérifiée) et que l'intersection entre $\text{Im}_{\text{lig}}(\mathbf{X}^+)$ et $\text{Im}_{\text{lig}}(\mathbf{U}_f^+)$ soit vide [MDVV89], il est montré que :

$$\text{Im}_{\text{col}} \left(\boldsymbol{\Gamma}_f \mathbf{X}^+ \boldsymbol{\Pi}_{\mathbf{U}_f^{+\perp}} \right) = \text{Im}_{\text{col}}(\boldsymbol{\Gamma}_f). \quad (3.52)$$

⁷Nous verrons dans le paragraphe suivant que nous pouvons restreindre la DVS à $\bar{\mathbf{R}}_{22}$.

⁸Nous avons présenté précédemment une méthode permettant de dissocier le bruit du sous-espace signal.

Or, d'après l'expression (3.51) :

$$\text{Im}_{\text{col}} \left(\Gamma_f \mathbf{X}^+ \Pi_{\mathbf{U}_f^{+\perp}} \right) = \text{Im}_{\text{col}} \left(\mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^{+\perp}} \right). \quad (3.53)$$

Ces deux dernières équations nous permettent d'écrire que :

$$\text{Im}_{\text{col}} \left(\mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^{+\perp}} \right) = \text{Im}_{\text{col}} \left(\Gamma_f \right). \quad (3.54)$$

De plus :

$$\mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^{+\perp}} = \bar{\mathbf{R}}_{22} \bar{\mathbf{Q}}_2. \quad (3.55)$$

Ainsi :

$$\text{Im}_{\text{col}} \left(\mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^{+\perp}} \right) = \text{Im}_{\text{col}} \left(\bar{\mathbf{R}}_{22} \right) \quad (3.56)$$

puisque $\bar{\mathbf{Q}}_2$ est une matrice orthogonale. Finalement :

$$\text{Im}_{\text{col}} \left(\bar{\mathbf{R}}_{22} \right) = \text{Im}_{\text{col}} \left(\Gamma_f \right). \quad (3.57)$$

Cette dernière équation justifie le fait que la décomposition en valeurs singulières ne soit appliquée qu'à la matrice $\bar{\mathbf{R}}_{22}$:

$$\bar{\mathbf{R}}_{22} = [\bar{\mathbf{U}}_s \quad \bar{\mathbf{U}}_b] \begin{bmatrix} \bar{\boldsymbol{\Sigma}}_s & \mathbf{0} \\ \mathbf{0} & \bar{\boldsymbol{\Sigma}}_b \end{bmatrix} \begin{bmatrix} \bar{\mathbf{V}}_s^T \\ \bar{\mathbf{V}}_b^T \end{bmatrix} = \bar{\mathbf{U}}_s \bar{\boldsymbol{\Sigma}}_s \bar{\mathbf{V}}_s^T + \bar{\mathbf{U}}_b \bar{\boldsymbol{\Sigma}}_b \bar{\mathbf{V}}_b^T. \quad (3.58)$$

Elle permet également de démontrer qu'une estimation de la matrice d'observabilité étendue est directement accessible à partir du sous-espace signal $\bar{\mathbf{U}}_s \bar{\boldsymbol{\Sigma}}_s \bar{\mathbf{V}}_s^T$ [Ver93] :

$$\hat{\Gamma}_f = \bar{\mathbf{U}}_s \text{ ou } \hat{\Gamma}_f = \bar{\mathbf{U}}_s \bar{\boldsymbol{\Sigma}}_s^{1/2}. \quad (3.59)$$

La décomposition en valeurs singulières de $\bar{\mathbf{R}}_{22}$ permet donc d'estimer l'ordre du système, de supprimer les effets des perturbations et d'extraire une base de la matrice d'observabilité étendue en une seule phase.

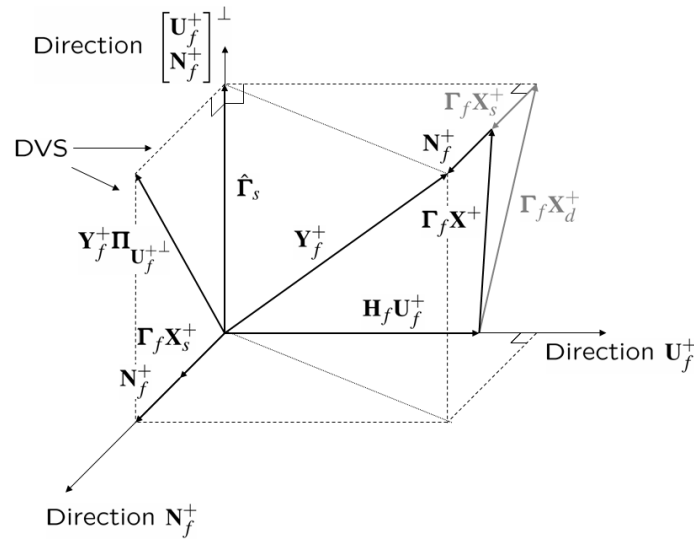


FIG. 3.3: Interprétation géométrique de l'algorithme Ordinary MOESP.

L'algorithme Ordinary MOESP est donc composé des étapes suivantes :

1. A partir des données d'entrée-sortie mesurées, construire les matrices de Hankel \mathbf{U}_f^+ et \mathbf{Y}_f^+ définies par (3.31) et (3.32) en fixant $M \gg f > n_x$.
2. Appliquer une factorisation QR au bloc matriciel $\begin{bmatrix} \mathbf{U}_f^{+T} & \mathbf{Y}_f^{+T} \end{bmatrix}^T$.
3. Extraire la matrice triangulaire $\bar{\mathbf{R}}_{22}$ (cf. equ. (3.44)).
4. Appliquer une décomposition en valeurs singulières à $\bar{\mathbf{R}}_{22}$ (cf. equ. (3.58)).
5. Inspecter les valeurs singulières, estimer l'ordre n_x du système et conserver les vecteurs singuliers gauches $\bar{\mathbf{U}}_s$ correspondant aux n_x plus grandes valeurs singulières.
6. Extraire les n_y premières lignes de $\bar{\mathbf{U}}_s$ pour construire la matrice \mathbf{C} .
7. Utiliser l'A-invariance de $\bar{\mathbf{U}}_s$ pour estimer \mathbf{A} .
8. Estimer \mathbf{B} et \mathbf{D} à l'aide d'une régression linéaire [VD92a].

Il est important de remarquer que l'étape d'annulation des effets des perturbations (et par conséquent d'extraction de la matrice d'observabilité) repose sur la possibilité de distinguer les valeurs singulières liées au bruit de celles associées à la dynamique du procédé. Malheureusement, cette opération n'est réalisable de manière consistante que si la matrice de covariance du bruit est proportionnelle à l'identité [Ver93, Vib95]. L'algorithme Ordinary MOESP fournit donc une estimation consistante de la partie déterministe du modèle de procédé si et seulement si le bruit de sortie est un bruit blanc de moyenne nulle et de variance finie. Dans tous les autres

cas, les estimées obtenues sont asymptotiquement biaisées. Pour remédier au manque de consistance de cette méthode en présence, par exemple, d'un bruit de sortie coloré, M. Verhaegen [Ver93] propose d'introduire une variable instrumentale au cours de la deuxième phase d'estimation afin d'annuler les effets des perturbations. Les techniques PI et PO MOESP traitent ce type de problématique.

3.3.2 Algorithmes PI et PO MOESP

Au sein de ce paragraphe, deux algorithmes d'identification des sous-espaces incorporant une variable instrumentale sont présentés. Ils diffèrent par le choix des instruments composant la variable instrumentale et par les hypothèses fixées sur les bruits agissant sur le procédé (cf. tab. 3.1) mais suivent le même formalisme mathématique. Ainsi, l'algorithme PI MOESP traite :

- les systèmes soumis à un bruit de sortie \mathbf{v} coloré (BC) de variance finie, les bruits d'état \mathbf{w} et de mesure \mathbf{v} étant nuls ;
- les systèmes à bruit d'état nul perturbés par un bruit de sortie \mathbf{v} coloré (BC) de variance finie, excités par une entrée $\tilde{\mathbf{u}}$ de type bruit blanc gaussien dont la mesure est perturbée par un bruit \mathbf{v} de moyenne nulle et de matrice de covariance proportionnelle à l'identité ;

alors que PO MOESP est utilisable dans deux situations :

- le bruit de mesure \mathbf{v} est nul et les bruits d'état \mathbf{w} et de sortie \mathbf{v} sont deux bruits blancs gaussiens ;
- les bruits de mesure \mathbf{v} et \mathbf{v} , le bruit d'état \mathbf{w} et la séquence d'entrée $\tilde{\mathbf{u}}$ sont des bruits blancs de moyenne nulle et de variance finie.

Quatre modèles sont donc à considérer :

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B} \underbrace{\tilde{\mathbf{u}}(t)}_{=\mathbf{u}(t) \text{ puisque } \mathbf{v}(t)=0} \quad (3.60a)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D} \underbrace{\tilde{\mathbf{u}}(t)}_{=\mathbf{u}(t)} + \mathbf{v}(t) \quad (3.60b)$$

où \mathbf{v} est un bruit coloré de variance finie,

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\{\mathbf{u}(t) - \mathbf{v}(t)\} \quad (3.61a)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\{\mathbf{u}(t) - \mathbf{v}(t)\} + \mathbf{v}(t) \quad (3.61b)$$

où $\mathbf{v}(t)$ est un bruit blanc gaussien et \mathbf{v} est un bruit coloré de variance finie,

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B} \underbrace{\tilde{\mathbf{u}}(t)}_{=\mathbf{u}(t) \text{ puisque } \mathbf{v}(t)=0} + \mathbf{w}(t) \quad (3.62a)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D} \underbrace{\tilde{\mathbf{u}}(t)}_{=\mathbf{u}(t)} + \mathbf{v}(t) \quad (3.62b)$$

où \mathbf{v} et \mathbf{w} sont deux bruits blancs gaussiens et

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\{\mathbf{u}(t) - \mathbf{v}(t)\} + \mathbf{w}(t) \quad (3.63a)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\{\mathbf{u}(t) - \mathbf{v}(t)\} + \mathbf{v}(t) \quad (3.63b)$$

où \mathbf{v} , \mathbf{w} , \mathbf{u} et $\tilde{\mathbf{u}}$ sont quatre bruits blancs gaussiens.

Ils vérifient respectivement les équations matricielles suivantes :

$$\mathbf{Y}_f^+ = \mathbf{\Gamma}_f \mathbf{X}^+ + \mathbf{H}_f \mathbf{U}_f^+ + \mathbf{V}_f^+ \quad (3.64a)$$

$$\mathbf{Y}_f^+ = \mathbf{\Gamma}_f \mathbf{X}^+ + \mathbf{H}_f \mathbf{U}_f^+ + \mathbf{V}_f^+ - \mathbf{H}_f \mathbf{Y}_f^+ \quad (3.64b)$$

$$\mathbf{Y}_f^+ = \mathbf{\Gamma}_f \mathbf{X}^+ + \mathbf{H}_f \mathbf{U}_f^+ + \mathbf{G}_f \mathbf{W}_f^+ + \mathbf{V}_f^+ \quad (3.64c)$$

$$\mathbf{Y}_f^+ = \mathbf{\Gamma}_f \mathbf{X}^+ + \mathbf{H}_f \mathbf{U}_f^+ + \mathbf{G}_f \mathbf{W}_f^+ - \mathbf{H}_f \mathbf{Y}_f^+ + \mathbf{V}_f^+ \quad (3.64d)$$

pour lesquelles $\mathbf{G}_f \in \mathbb{R}^{n_y f \times n_x f}$ est la matrice de Toeplitz par blocs liée au bruit d'état \mathbf{w} :

$$\mathbf{G}_f = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{C} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{CA} & \mathbf{C} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{CA}^{f-2} & \mathbf{CA}^{f-3} & \dots & \mathbf{0} \end{bmatrix}. \quad (3.65)$$

Il est alors aisé de regrouper ces quatre écritures en considérant l'expression suivante :

$$\mathbf{Y}_f^+ = \mathbf{\Gamma}_f \mathbf{X}^+ + \mathbf{H}_f \mathbf{U}_f^+ + \mathbf{N}_f^+ \quad (3.66)$$

en posant respectivement :

$$\mathbf{N}_f^+ = \mathbf{V}_f^+ \quad (3.67a)$$

$$\mathbf{N}_f^+ = \mathbf{V}_f^+ - \mathbf{H}_f \mathbf{Y}_f^+ \quad (3.67b)$$

$$\mathbf{N}_f^+ = \mathbf{G}_f \mathbf{W}_f^+ + \mathbf{V}_f^+ \quad (3.67c)$$

$$\mathbf{N}_f^+ = \mathbf{G}_f \mathbf{W}_f^+ - \mathbf{H}_f \mathbf{Y}_f^+ + \mathbf{V}_f^+. \quad (3.67d)$$

Puisque \mathbf{N}_f^+ ne possède plus, par construction, de matrice de covariance proportionnelle à l'identité :

$$\lim_{M \rightarrow \infty} \mathbf{N}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \neq \mathbf{N}_f^+. \quad (3.68)$$

L'application d'une simple projection orthogonale n'est alors plus suffisante pour estimer de manière consistante $\mathbf{\Gamma}_f$. L'astuce proposée par M. Verhaegen [Ver93] repose sur l'introduction d'une variable instrumentale au sein de la factorisation QR (3.44) d'Ordinary MOESP afin de projeter l'équation (3.66) dans le plan déterministe. Elle doit ainsi permettre d'annuler les effets des perturbations tout en conservant les informations utiles liées à la dynamique du système étudié. De plus, dans le cas particulier des méthodes d'identification des sous-espaces, la variable instrumentale $\boldsymbol{\xi}$ doit vérifier la propriété suivante :

$$\text{rang} \left(\mathbb{E} \left\{ \mathbf{x}(t) \boldsymbol{\xi}^T(t) \right\} \right) = n_x, \quad (3.69)$$

condition nécessaire à l'extraction de $\mathbf{\Gamma}_f$ à partir de $\mathbb{E} \left\{ \mathbf{y}(t) \boldsymbol{\xi}^T(t) \right\}$ [Vib95]. La solution naturelle serait d'utiliser le signal d'entrée comme instruments puisque ce dernier est décorrélié du bruit de sortie (en boucle ouverte) et suffisamment lié à l'état. Malheureusement, en identification des sous-espaces, la variable instrumentale doit également être orthogonale à la matrice de

Hankel des entrées afin d'isoler le terme $\mathbf{H}_f \mathbf{U}_f^+$. Cette condition est donc incompatible avec le choix de variable instrumentale énoncé précédemment. L'artifice, proposé par plusieurs auteurs [Ver93, VD94, OV94, CK95], consiste à scinder les données accessibles à la mesure en deux blocs afin de créer des matrices de Hankel passées et futures. Considérons donc le vecteur de sorties décalées passées suivant⁹ :

$$\mathbf{y}_p^-(t) = \begin{bmatrix} \mathbf{y}(t-p) \\ \vdots \\ \mathbf{y}(t-1) \end{bmatrix} \in \mathbb{R}^{n_y p \times 1}. \quad (3.70)$$

Il est alors facile de vérifier que :

$$\mathbf{y}_p^-(t) = \mathbf{\Gamma}_p \mathbf{x}(t-p) + \mathbf{H}_p \mathbf{u}_p^-(t) + \mathbf{n}_p^-(t) \quad (3.71)$$

où les matrices $\mathbf{\Gamma}_p$ et \mathbf{H}_p sont respectivement définies comme $\mathbf{\Gamma}_f$ et \mathbf{H}_f (cf. equ. (3.28) et (3.29)). Par itérations sur l'indice t , cette équation peut être étendue à l'équation matricielle suivante :

$$\mathbf{Y}_p^- = \mathbf{\Gamma}_p \mathbf{X}^- + \mathbf{H}_p \mathbf{U}_p^- + \mathbf{N}_p^- \quad (3.72)$$

avec¹⁰ :

$$\begin{aligned} \mathbf{Y}_p^- &= [\mathbf{y}_p^-(t) \cdots \mathbf{y}_p^-(\underbrace{t+M-1}_{\bar{t}})] \\ &= \begin{bmatrix} \mathbf{y}(t-p) & \mathbf{y}(t-p+1) & \cdots & \mathbf{y}(\bar{t}-p) \\ \mathbf{y}(t-p+1) & \mathbf{y}(t-p+2) & \cdots & \mathbf{y}(\bar{t}-p+1) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{y}(t-1) & \mathbf{y}(t) & \cdots & \mathbf{y}(\bar{t}-1) \end{bmatrix} \in \mathbb{R}^{n_y p \times M} \end{aligned} \quad (3.74)$$

et :

$$\mathbf{X}^- = [\mathbf{x}(t-p) \cdots \mathbf{x}(\bar{t}-p)]. \quad (3.75)$$

De même, en considérant le vecteur de sorties décalées futures suivant :

$$\mathbf{y}_f^+(t) = \begin{bmatrix} \mathbf{y}(t) \\ \vdots \\ \mathbf{y}(t+f-1) \end{bmatrix} \quad (3.76)$$

et la matrice de Hankel des sorties futures :

$$\mathbf{Y}_f^+ = [\mathbf{y}_f^+(t) \cdots \mathbf{y}_f^+(\bar{t})], \quad (3.77)$$

nous savons que :

$$\mathbf{y}_f^+ = \mathbf{\Gamma}_f \mathbf{x} + \mathbf{H}_f \mathbf{u}_f^+ + \mathbf{n}_f^+ \quad (3.78)$$

⁹La notion de données passées ou futures est relative à l'indice temporel t dans le cas des équations vérifiées par les vecteurs de données décalées. Elle sera relative à $\bar{t} = t + M - 1$ dans le cas des équations vérifiées par les matrices de Hankel.

¹⁰les matrices \mathbf{U}_p^- et \mathbf{N}_p^- sont construites de manière similaire à \mathbf{Y}_p^- .

et :

$$\mathbf{Y}_f^+ = \Gamma_f \mathbf{X}^+ + \mathbf{H}_f \mathbf{U}_f^+ + \mathbf{N}_f^+. \quad (3.79)$$

En posant $\Xi = \mathbf{U}_p^-$ comme variable instrumentale pour la classe d'identification PI MOESP et $\Xi = \begin{bmatrix} \mathbf{U}_p^{-T} & \mathbf{Y}_p^{-T} \end{bmatrix}^T$ pour PO MOESP [Ver93, Ver94], il est facile de constater que :

- Ξ est asymptotiquement décorrélée du bruit futur¹¹ :

$$\lim_{M \rightarrow \infty} \mathbf{N}_f^+ \Xi^T = \mathbf{0}; \quad (3.80)$$

- Ξ est suffisamment corrélée avec l'état :

$$\text{rang} \left(\lim_{M \rightarrow \infty} \mathbf{X}^+ \Xi^T \right) = n_x; \quad (3.81)$$

- Ξ est orthogonale à \mathbf{U}_f^+ :

$$\lim_{M \rightarrow \infty} \mathbf{U}_f^+ \Xi^T = \mathbf{0}. \quad (3.82)$$

Remarque 3.3. Il est important de noter que, si $\mathbf{v} \neq \mathbf{0}$, l'égalité (3.82) est uniquement vérifiée si la séquence d'entrée $\tilde{\mathbf{u}}$ est un bruit blanc gaussien [CV97]. Cette condition, très restrictive en pratique, sera étendue à des entrées non blanches dans le paragraphe 3.3.3.

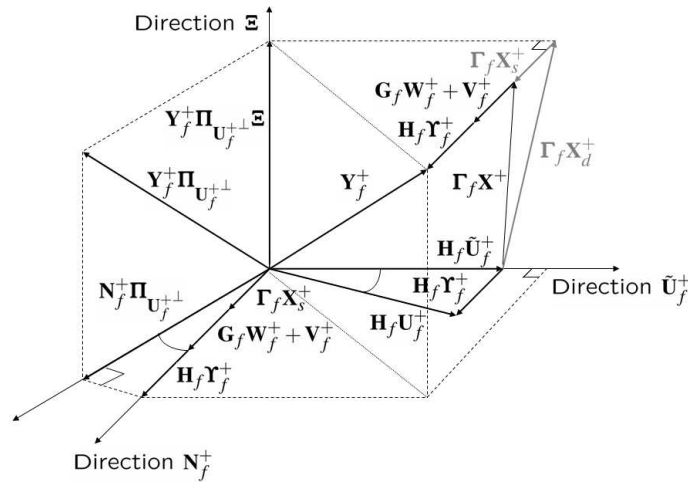


FIG. 3.4: Interprétation géométrique des algorithmes PI et PO MOESP.

La procédure d'estimation de la matrice d'observabilité étendue est alors identique à celle suivie pour l'algorithme Ordinary MOESP. En supposant que le signal d'entrée non bruité $\tilde{\mathbf{u}}$ vérifie la condition de rang suivante :

$$\text{rang} \left(\begin{bmatrix} \tilde{\mathbf{U}}_p^- \\ \tilde{\mathbf{U}}_f^+ \end{bmatrix} \right) = n_u(f + p), \quad (3.83)$$

¹¹L'application de l'algorithme PI MOESP suppose que $\mathbf{w} = \mathbf{0}$.

une factorisation QR du bloc matriciel $\begin{bmatrix} \mathbf{U}_f^{+T} & \mathbf{\Xi}^T & \mathbf{Y}_f^{+T} \end{bmatrix}^T$ est réalisée :

$$\begin{bmatrix} \mathbf{U}_f^+ \\ \mathbf{\Xi} \\ \mathbf{Y}_f^+ \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_{11} & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{21} & \bar{\mathbf{R}}_{22} & \mathbf{0} \\ \bar{\mathbf{R}}_{31} & \bar{\mathbf{R}}_{32} & \bar{\mathbf{R}}_{33} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1 \\ \bar{\mathbf{Q}}_2 \\ \bar{\mathbf{Q}}_3 \end{bmatrix} \quad (3.84)$$

où $\bar{\mathbf{R}}_{11} \in \mathbb{R}^{n_u f \times n_u f}$, $\bar{\mathbf{R}}_{21} \in \mathbb{R}^{kp \times n_u f}$, $\bar{\mathbf{R}}_{31} \in \mathbb{R}^{n_y f \times n_u f}$, $\bar{\mathbf{R}}_{22} \in \mathbb{R}^{kp \times kp}$, $\bar{\mathbf{R}}_{32} \in \mathbb{R}^{n_y f \times kp}$, $\bar{\mathbf{R}}_{33} \in \mathbb{R}^{n_y f \times n_y f}$, $\bar{\mathbf{Q}}_1 \in \mathbb{R}^{n_u f \times M}$, $\bar{\mathbf{Q}}_2 \in \mathbb{R}^{kp \times M}$ et $\bar{\mathbf{Q}}_3 \in \mathbb{R}^{n_y f \times M}$, k valant n_u pour l'algorithme PI MOESP, $n_u + n_y$ pour PO MOESP. Cette factorisation conduit au résultat suivant [Ver93, Ver94] :

$$\text{Im}_{\text{col}}(\bar{\mathbf{U}}_s) = \text{Im}_{\text{col}}(\mathbf{\Gamma}_f). \quad (3.85)$$

Puisque la variable instrumentale atténue les effets des perturbations sans les supprimer totalement, une décomposition en valeurs singulières du bloc matriciel $\bar{\mathbf{R}}_{32}$ est nécessaire pour lui appliquer une réduction de rang :

$$\bar{\mathbf{R}}_{32} = [\bar{\mathbf{U}}_s \quad \bar{\mathbf{U}}_b] \begin{bmatrix} \bar{\mathbf{\Sigma}}_s & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{\Sigma}}_b \end{bmatrix} \begin{bmatrix} \bar{\mathbf{V}}_s^T \\ \bar{\mathbf{V}}_b^T \end{bmatrix}. \quad (3.86)$$

Cette équation permet à la fois d'estimer l'ordre du système en parcourant les valeurs singulières calculées et d'extraire une base de la matrice d'observabilité du système puisque :

$$\text{Im}_{\text{col}}(\bar{\mathbf{U}}_s) = \text{Im}_{\text{col}}(\mathbf{\Gamma}_f). \quad (3.87)$$

Finalement, les algorithmes PI et PO MOESP se composent des étapes suivantes :

1. A partir des données d'entrée-sortie mesurées, construire les matrices de Hankel \mathbf{U}_p^- , \mathbf{U}_f^+ , \mathbf{Y}_p^- et \mathbf{Y}_f^+ en fixant $M \gg p > n_x$ et $M \gg f > n_x$.
2. Construire la variable instrumentale $\mathbf{\Xi} = \mathbf{U}_p^-$ pour l'algorithme PI MOESP, $\mathbf{\Xi} = \begin{bmatrix} \mathbf{U}_p^{-T} & \mathbf{Y}_p^{-T} \end{bmatrix}^T$ pour l'algorithme PO MOESP.
3. Appliquer une factorisation QR au bloc matriciel $\begin{bmatrix} \mathbf{U}_f^{+T} & \mathbf{\Xi}^T & \mathbf{Y}_f^{+T} \end{bmatrix}^T$ (cf. equ. (3.84)).
4. Extraire la matrice $\bar{\mathbf{R}}_{32}$.
5. Lui appliquer une décomposition en valeurs singulières (cf. equ. (3.86)).
6. Inspecter les valeurs singulières, estimer l'ordre n_x du système et conserver les vecteurs singuliers gauches $\bar{\mathbf{U}}_s$ correspondant aux n_x plus grandes valeurs singulières.
7. Extraire les n_y premières lignes de $\bar{\mathbf{U}}_s$ pour construire la matrice \mathbf{C} .
8. Utiliser l'A-invariance de $\bar{\mathbf{U}}_s$ pour estimer \mathbf{A} .
9. Estimer \mathbf{B} et \mathbf{D} à l'aide d'une régression linéaire [Ver93, Ver94].

Remarque 3.4. *La distinction entre les schémas d'identification PI MOESP et PO MOESP va au delà du simple choix d'une variable instrumentale. L'application de PI MOESP permet uniquement d'identifier la partie déterministe de la fonction de transfert du procédé tandis que PO MOESP estime le système combiné [Vib95].*

Une description plus détaillée de la phase d'estimation des matrices d'état à partir de Γ_f sera présentée dans le paragraphe 3.3.4.

3.3.3 Algorithme MOESP pour les modèles avec « erreurs en les variables »

On appelle modèle à « erreurs en les variables » toute représentation paramétrique de système tenant compte des bruits de mesure en entrée et en sortie du procédé. Il n'est pas rare que cette situation se présente en pratique [Söd81, And85, ZF89, ZF92]. Ces conditions d'identification ont été abordées de manière partielle dans le paragraphe précédent puisque les algorithmes PI MOESP et PO MOESP traitent cette problématique dans un cas particulier simple : l'entrée non perturbée et le bruit de mesure associé sont deux bruits blancs gaussiens de moyenne nulle et de variance finie. Or, cette hypothèse sur la séquence d'entrée $\tilde{\mathbf{u}}$ est assez restrictive en pratique. Il n'est en effet pas toujours possible d'appliquer ce type de signal au processus à identifier. C. T. Chou et M. Verhaegen furent les premiers à considérer ce problème dans le contexte des méthodes directes des sous-espaces [CV97]. Ils supposent que :

- les entrées et les sorties sont affectées par des bruits blancs additifs ;
- le système est perturbé par un bruit d'état blanc ;
- ces trois bruits sont partiellement ou totalement corrélés ;
- le vecteur d'entrée $\tilde{\mathbf{u}}$ est une séquence, non nécessairement blanche, vérifiant la condition d'excitation persistante.

Ils exploitent la variable instrumentale $\Xi = \begin{bmatrix} \mathbf{U}_p^{-T} & \mathbf{Y}_p^{-T} \end{bmatrix}^T$ et considère la factorisation QR suivante [CV97] :

$$\begin{bmatrix} \mathbf{U}_f^+ \mathbf{U}_p^{-T} & \mathbf{U}_f^+ \mathbf{Y}_p^{-T} \\ \mathbf{Y}_f^+ \mathbf{U}_p^{-T} & \mathbf{Y}_f^+ \mathbf{Y}_p^{-T} \end{bmatrix} = \begin{bmatrix} \mathbf{U}_f^+ \\ \mathbf{Y}_f^+ \end{bmatrix} \begin{bmatrix} \mathbf{U}_p^{-T} & \mathbf{Y}_p^{-T} \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_{11} & \mathbf{0} \\ \bar{\mathbf{R}}_{21} & \bar{\mathbf{R}}_{22} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1 \\ \bar{\mathbf{Q}}_2 \end{bmatrix}. \quad (3.88)$$

En appliquant une décomposition en valeurs singulières au bloc matriciel $\bar{\mathbf{R}}_{22}$:

$$\bar{\mathbf{R}}_{22} = \begin{bmatrix} \bar{\mathbf{U}}_s & \bar{\mathbf{U}}_b \end{bmatrix} \begin{bmatrix} \bar{\Sigma}_s & \mathbf{0} \\ \mathbf{0} & \bar{\Sigma}_b \end{bmatrix} \begin{bmatrix} \bar{\mathbf{V}}_s^T \\ \bar{\mathbf{V}}_b^T \end{bmatrix}, \quad (3.89)$$

une estimation consistante d'une base de la matrice d'observabilité étendue est accessible. L'estimation de \mathbf{A} et \mathbf{C} est réalisée, comme précédemment, en utilisant la propriété d'A-invariance de \mathbf{A} . L'obtention des matrices \mathbf{B} et \mathbf{D} est plus complexe que dans les cas précédents. Elle fait appel à un modèle d'erreur de sortie particulier contenant une variable instrumentale précise [CV97]. L'algorithme PO_EIV MOESP est donc constitué des étapes suivantes :

1. A partir des données d'entrée-sortie mesurées, construire les matrices de Hankel \mathbf{U}_p^- , \mathbf{U}_f^+ , \mathbf{Y}_p^- et \mathbf{Y}_f^+ en fixant $M \gg f > n_x$ et $M \gg p > n_x$.
2. Construire la variable instrumentale $\mathbf{\Xi} = \begin{bmatrix} \mathbf{U}_p^{-T} & \mathbf{Y}_p^{-T} \end{bmatrix}^T$.
3. Appliquer une factorisation QR au bloc matriciel $\begin{bmatrix} \mathbf{\Xi}^T \mathbf{U}_f^{+T} & \mathbf{\Xi}^T \mathbf{Y}_f^{+T} \end{bmatrix}^T$ (cf. equ. (3.88)).
4. Extraire la matrice triangulaire $\bar{\mathbf{R}}_{22}$.
5. Lui appliquer une décomposition en valeurs singulières.
6. Inspecter les valeurs singulières, estimer l'ordre n_x du système et conserver les vecteurs singuliers gauches $\bar{\mathbf{U}}_s$ correspondant aux n_x plus grandes valeurs singulières.
7. Extraire les n_y premières lignes de $\bar{\mathbf{U}}_s$ pour construire la matrice \mathbf{C} .
8. Utiliser l'A-invariance de $\bar{\mathbf{U}}_s$ pour estimer \mathbf{A} .
9. Estimer \mathbf{B} et \mathbf{D} à l'aide d'une régression linéaire [CV97].

Remarque 3.5. *Un des points forts de l'algorithme PO_EIV MOESP est qu'il peut facilement s'étendre à l'identification en boucle fermée. Le lecteur intéressé par l'identification des sous-espaces en boucle fermée pourra trouver quelques pistes de recherche dans [VD96a, LM96, CV97, TLS97, QL03, LQL04, CP04].*

3.3.4 Estimation des matrices d'état

Jusqu'ici, notre étude s'est principalement focalisée sur l'estimation de la matrice d'observabilité. Or, l'objectif des méthodes des sous-espaces est d'estimer de manière consistante une réalisation d'état du procédé, i.e. calculer les matrices d'état du modèle. Plusieurs options ont été proposées dans la littérature [Ver94, Vib95, McK95, VD96b, VWO97, Lov98]. Comme précédemment, nous restreignons notre présentation aux techniques développées pour la classe d'identification MOESP. Puisque ces dernières ne cherchent en aucun cas à estimer explicitement le vecteur d'état du système, l'estimation des matrices d'état du procédé comporte deux phases :

- l'estimation des matrices \mathbf{A} et \mathbf{C} à partir d'une estimée de la matrice d'observabilité $\mathbf{\Gamma}_f$;
- l'estimation des matrices \mathbf{B} et \mathbf{D} connaissant $\hat{\mathbf{A}}$ et $\hat{\mathbf{C}}$ et les données d'entrée-sortie.

Une description des principaux algorithmes d'estimation est proposée au sein des deux paragraphes suivants.

3.3.4.1 Estimation des matrices \mathbf{A} et \mathbf{C}

En supposant que nous ayons accès à une estimée consistante de la matrice d'observabilité, l'estimation de la matrice de sortie repose sur la constatation suivante :

$$\hat{\mathbf{C}} = \hat{\mathbf{\Gamma}}_f(1 : n_y, :). \quad (3.90)$$

La matrice \mathbf{C} est donc estimée en extrayant les n_y première lignes de $\hat{\mathbf{\Gamma}}_f$.

L'estimation de la matrice \mathbf{A} est quant à elle fondée sur la célèbre propriété d'A-invariance de $\mathbf{\Gamma}_f$ [Kun78] :

$$\mathbf{\Gamma}_f^\downarrow = \mathbf{\Gamma}_f^\uparrow \mathbf{A} \quad (3.91)$$

où :

$$\mathbf{\Gamma}_f^\uparrow = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{f-2} \end{bmatrix} = \mathbf{\Gamma}_f(1 : n_y(f-1), :) \quad \text{et} \quad \mathbf{\Gamma}_f^\downarrow = \begin{bmatrix} \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^{f-1} \end{bmatrix} = \mathbf{\Gamma}_f(n_y + 1 : n_y f, :). \quad (3.92)$$

Plusieurs algorithmes d'estimation ont été proposés. La première solution consiste simplement à résoudre, au sens des moindres carrés, le système surdimensionné :

$$\hat{\mathbf{A}} = \left\{ \hat{\mathbf{\Gamma}}_f^\uparrow \right\}^\dagger \hat{\mathbf{\Gamma}}_f^\downarrow \quad (3.93)$$

où $\left\{ \hat{\mathbf{\Gamma}}_f^\uparrow \right\}^\dagger$ est la pseudo inverse de Moore Penrose [Rot01] de $\hat{\mathbf{\Gamma}}_f^\uparrow$. Malheureusement, lorsque les matrices $\hat{\mathbf{\Gamma}}_f^\uparrow$ et $\hat{\mathbf{\Gamma}}_f^\downarrow$ sont bruitées, cette solution peut fournir une estimée asymptotiquement biaisée de \mathbf{A} . Une version alternative, fondée sur les moindres carrés totaux [VV91], a alors été exposée dans [VDCV94]. Elle considère le système matriciel :

$$\tilde{\mathbf{\Gamma}}_f^\downarrow + \Delta \hat{\mathbf{\Gamma}}_f^\downarrow = \left(\tilde{\mathbf{\Gamma}}_f^\uparrow + \Delta \hat{\mathbf{\Gamma}}_f^\uparrow \right) \mathbf{A} \quad (3.94)$$

où $\Delta \hat{\mathbf{\Gamma}}_f^\uparrow$ et $\Delta \hat{\mathbf{\Gamma}}_f^\downarrow$ symbolisent les erreurs ou bruits additifs contenus dans les matrices $\hat{\mathbf{\Gamma}}_f^\uparrow$ et $\hat{\mathbf{\Gamma}}_f^\downarrow$. Elle repose sur la décomposition en valeurs singulières suivante [OVK91] :

$$\begin{bmatrix} \left(\hat{\mathbf{\Gamma}}_f^\uparrow \right)^T \\ \left(\hat{\mathbf{\Gamma}}_f^\downarrow \right)^T \end{bmatrix} \begin{bmatrix} \hat{\mathbf{\Gamma}}_f^\uparrow & \hat{\mathbf{\Gamma}}_f^\downarrow \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{V}}_{11} & \bar{\mathbf{V}}_{12} \\ \bar{\mathbf{V}}_{21} & \bar{\mathbf{V}}_{22} \end{bmatrix} \bar{\mathbf{\Sigma}} \begin{bmatrix} \bar{\mathbf{V}}_{11}^T & \bar{\mathbf{V}}_{12}^T \\ \bar{\mathbf{V}}_{21}^T & \bar{\mathbf{V}}_{22}^T \end{bmatrix}, \quad (3.95)$$

décomposition conduisant à l'estimée :

$$\hat{\mathbf{A}} = -\bar{\mathbf{V}}_{12} \bar{\mathbf{V}}_{22}^\dagger. \quad (3.96)$$

Une troisième solution, inspirée d'un problème particulier de traitement d'antennes [VD94], a été développée par M. Lovera [Lov98, LFB98]. La justification de son développement repose sur une observation simple : de par la construction des matrices $\mathbf{\Gamma}_f^\uparrow$ et $\mathbf{\Gamma}_f^\downarrow$, les perturbations $\Delta \hat{\mathbf{\Gamma}}_f^\uparrow$ et $\Delta \hat{\mathbf{\Gamma}}_f^\downarrow$ sont fortement corrélées. Or, l'application des moindres carrés totaux à un tel

problème ne peut fournir qu'une estimée non optimale. Le remède, nommé moindres carrés totaux partiellement structurés, repose sur une prise en compte explicite de la structure des matrices de bruit. Cette technique cherche à estimer, dans un premier temps, la matrice de covariance des résidus en utilisant le principe « d'expansion du sous-espace des perturbations » [Vac94]. La matrice d'état du système est ensuite estimée à l'aide d'un moindre carré pondéré. Nous invitons les lecteurs intéressés à étudier les articles [Vac94] et [LFB98].

3.3.4.2 Estimation des matrices \mathbf{B} et \mathbf{D}

Les matrices \mathbf{A} et \mathbf{C} étant connues, \mathbf{B} et \mathbf{D} demeurent les inconnues du système d'état. Deux méthodologies sont proposées dans la littérature [Ver94, McK95]. La première est fondée sur l'observation suivante : la matrice de Toeplitz par blocs \mathbf{H}_f est linéaire en \mathbf{B} et \mathbf{D} [VD92a]. Son estimation permet donc d'obtenir $\hat{\mathbf{B}}$ et $\hat{\mathbf{D}}$ par une simple régression linéaire. La seconde repose sur l'utilisation de l'erreur de simulation du modèle :

$$\hat{\mathbf{y}}(t) = \hat{\mathbf{C}}\hat{\mathbf{A}}^t\mathbf{x}(0) + \sum_{k=0}^{t-1} \hat{\mathbf{C}}\hat{\mathbf{A}}^{t-k-1}\hat{\mathbf{B}}\mathbf{u}(k) + \hat{\mathbf{D}}\mathbf{u}(t) \quad (3.97)$$

employée comme fonction coût à minimiser [McK95]. Ces deux techniques sont présentées dans la suite de ce paragraphe.

A partir de la matrice \mathbf{H}_f

- Dans le cas déterministe, il est facile de montrer que, lorsque le nombre de données accessibles tend vers l'infini :

$$\bar{\mathbf{R}}_{21} = \mathbf{\Gamma}_f (\mathbf{X}^+ \bar{\mathbf{Q}}_1^T) + \mathbf{H}_f \bar{\mathbf{R}}_{11} \quad (3.98)$$

où $\bar{\mathbf{R}}_{11}$, $\bar{\mathbf{R}}_{21}$ et $\bar{\mathbf{Q}}_1$ sont les matrices résultantes de la décomposition QR du bloc matriciel $\begin{bmatrix} \mathbf{U}_f^{+T} & \mathbf{Y}_f^{+T} \end{bmatrix}^T$:

$$\begin{bmatrix} \mathbf{U}_f^+ \\ \mathbf{Y}_f^+ \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_{11} & \mathbf{0} \\ \bar{\mathbf{R}}_{21} & \bar{\mathbf{R}}_{22} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1 \\ \bar{\mathbf{Q}}_2 \end{bmatrix}. \quad (3.99)$$

- Dans le cas stochastique, si nous considérons la décomposition QR des algorithmes PI MOESP :

$$\begin{bmatrix} \mathbf{U}_f^+ \\ \mathbf{U}_p^- \\ \mathbf{Y}_f^+ \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_{11} & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{21} & \bar{\mathbf{R}}_{22} & \mathbf{0} \\ \bar{\mathbf{R}}_{31} & \bar{\mathbf{R}}_{32} & \bar{\mathbf{R}}_{33} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1 \\ \bar{\mathbf{Q}}_2 \\ \bar{\mathbf{Q}}_3 \end{bmatrix} \quad (3.100)$$

et PO MOESP :

$$\begin{bmatrix} \mathbf{U}_f^+ \\ \mathbf{U}_p^- \\ \mathbf{Y}_p^- \\ \mathbf{Y}_f^+ \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_{11} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{21} & \bar{\mathbf{R}}_{22} & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{31} & \bar{\mathbf{R}}_{32} & \bar{\mathbf{R}}_{33} & \mathbf{0} \\ \bar{\mathbf{R}}_{41} & \bar{\mathbf{R}}_{42} & \bar{\mathbf{R}}_{43} & \bar{\mathbf{R}}_{44} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1 \\ \bar{\mathbf{Q}}_2 \\ \bar{\mathbf{Q}}_3 \\ \bar{\mathbf{Q}}_4 \end{bmatrix}, \quad (3.101)$$

il est démontré qu'asymptotiquement [Ver93, Ver94] :

$$\bar{\mathbf{R}}_{31} = \mathbf{\Gamma}_f (\mathbf{X}^+ \bar{\mathbf{Q}}_1^T) + \mathbf{H}_f \bar{\mathbf{R}}_{11} \quad (3.102)$$

pour la classe d'identification PI MOESP et :

$$[\bar{\mathbf{R}}_{31} \quad \bar{\mathbf{R}}_{32} \quad \bar{\mathbf{R}}_{41}] = \mathbf{\Gamma}_f [\mathbf{X}^- \bar{\mathbf{Q}}_1^T \quad \mathbf{X}^- \bar{\mathbf{Q}}_2^T \quad \mathbf{X}^+ \bar{\mathbf{Q}}_1^T] + \mathbf{H}_f [\bar{\mathbf{R}}_{21} \quad \bar{\mathbf{R}}_{22} \quad \bar{\mathbf{R}}_{11}] \quad (3.103)$$

pour l'algorithme PO MOESP.

Ces trois écritures peuvent alors être regroupées via l'expression suivante :

$$\bar{\mathbf{R}}_2 = \mathbf{\Gamma}_f (\mathbf{X}^+ \bar{\mathbf{Q}}^T) + \mathbf{H}_f \bar{\mathbf{R}}_1. \quad (3.104)$$

Notre objectif est d'extraire de cette relation une équation linéaire en \mathbf{B} et \mathbf{D} . Puisque cette dernière est contenue dans \mathbf{H}_f , il est nécessaire d'isoler cette matrice de l'équation (3.104). La première étape consiste à supprimer le terme $\mathbf{\Gamma}_f (\mathbf{X}^+ \bar{\mathbf{Q}}^T)$. Étant donné que l'espace colonne de $\mathbf{\Gamma}_f$ et de son complément orthogonal sont respectivement égaux aux matrices $\bar{\mathbf{U}}_s$ et $\bar{\mathbf{U}}_b$ des décompositions (3.58) et (3.86), nous nous affranchissons de $\mathbf{\Gamma}_f (\mathbf{X}^+ \bar{\mathbf{Q}}^T)$ en prémultipliant l'équation (3.104) par $\bar{\mathbf{U}}_b^T$:

$$\bar{\mathbf{U}}_b^T \bar{\mathbf{R}}_2 = \bar{\mathbf{U}}_b^T \mathbf{H}_f \bar{\mathbf{R}}_1. \quad (3.105)$$

La seconde phase consiste à inverser la matrice $\bar{\mathbf{R}}_1$. En supposant que la condition d'excitation persistante soit vérifiée, la pseudo inverse de $\bar{\mathbf{R}}_1$ existe [VD92a, Ver93, Ver94]. En postmultipliant l'équation précédente par cette pseudo inverse, il s'en suit que :

$$\bar{\mathbf{U}}_b^T \bar{\mathbf{R}}_2 \{\bar{\mathbf{R}}_1\}^\dagger = \bar{\mathbf{U}}_b^T \mathbf{H}_f. \quad (3.106)$$

En posant :

$$\mathbf{K} = \bar{\mathbf{U}}_b^T \bar{\mathbf{R}}_2 \{\bar{\mathbf{R}}_1\}^\dagger, \quad (3.107)$$

le système suivant est construit [Ver94] en exploitant la structure de \mathbf{H}_f :

$$\begin{bmatrix} \mathbf{K}(:, 1 : n_y) \\ \vdots \\ \mathbf{K}(:, n_y(f-1) + 1 : n_y f) \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{U}}_b^T(:, 1 : n_y) & \cdots & \bar{\mathbf{U}}_b^T(:, n_y(f-1) + 1 : n_y f) \\ \vdots & & \\ \bar{\mathbf{U}}_b^T(:, n_y(f-1) + 1 : n_y f) & & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{U}}_s^\dagger \end{bmatrix} \begin{bmatrix} \mathbf{D} \\ \mathbf{B} \end{bmatrix}. \quad (3.108)$$

Ce système surdimensionné peut alors être résolu à partir de méthodes telles que les moindres carrés.

A partir de l'erreur de simulation

La sortie du modèle identifié est donnée par¹² :

$$\hat{\mathbf{y}}(t) = \sum_{k=0}^{t-1} \hat{\mathbf{C}} \hat{\mathbf{A}}^{t-k-1} \hat{\mathbf{B}} \mathbf{u}(k) + \hat{\mathbf{D}} \mathbf{u}(t). \quad (3.109)$$

¹²Par souci de simplification, $\mathbf{x}(0) = \mathbf{0}$. Dans le cas contraire, puisque le système est supposé stable, le terme $\hat{\mathbf{C}} \hat{\mathbf{A}}^t \mathbf{x}(0)$ disparaîtra lorsque $t \rightarrow \infty$.

En posant $\text{vec}(\mathbf{M})$ le vecteur composé des colonnes empilées de la matrice \mathbf{M} et en exploitant la propriété du produit de Kronecker \otimes suivante [Rot01] :

$$\text{vec}(\mathbf{M}_1\mathbf{M}_2\mathbf{M}_3) = \mathbf{M}_3^T \otimes \mathbf{M}_1 \text{vec}(\mathbf{M}_2), \quad (3.110)$$

la relation (3.109) peut être réécrite comme suit [McK95] :

$$\hat{\mathbf{y}}(t) = [\mathbf{u}^T(t) \otimes \mathbf{I}_{n_y}] \text{vec}(\hat{\mathbf{D}}) + \left(\sum_{k=0}^{t-1} \mathbf{u}^T(k) \otimes \hat{\mathbf{C}}\hat{\mathbf{A}}^{t-k-1} \right) \text{vec}(\hat{\mathbf{B}}). \quad (3.111)$$

Les matrices \mathbf{B} et \mathbf{D} sont alors estimées en minimisant le critère suivant :

$$\hat{\mathbf{B}}, \hat{\mathbf{D}} = \arg \min_{\mathbf{B}, \mathbf{D}} \sum_t \left[\mathbf{y}(t) - [\mathbf{u}^T(t) \otimes \mathbf{I}_{n_y} \quad \sum_{k=0}^{t-1} \mathbf{u}^T(k) \otimes \hat{\mathbf{C}}\hat{\mathbf{A}}^{t-k-1}] \begin{bmatrix} \text{vec}(\mathbf{D}) \\ \text{vec}(\mathbf{B}) \end{bmatrix} \right]. \quad (3.112)$$

Remarque 3.6. *Il est intéressant de noter que le second membre du régresseur vérifie la relation de récurrence suivante [Gus96] :*

$$\left(\sum_{k=0}^t \mathbf{u}^T(k) \otimes \hat{\mathbf{C}}\hat{\mathbf{A}}^{t-k} \right) = \left(\sum_{k=0}^{t-1} \mathbf{u}^T(k) \otimes \hat{\mathbf{C}}\hat{\mathbf{A}}^{t-k-1} \right) (\mathbf{I}_{n_u} \otimes \hat{\mathbf{A}}) + (\mathbf{u}^T(t-1) \otimes \hat{\mathbf{C}}). \quad (3.113)$$

Cette observation est essentielle pour l'estimation en ligne des matrices \mathbf{B} et \mathbf{D} .

3.3.5 Propriétés asymptotiques des algorithmes MOESP et choix des paramètres structuraux

Bien que les méthodes des sous-espaces présentent une simplicité conceptuelle évidente et des caractéristiques numériques intéressantes, les propriétés statistiques de ces dernières ne sont, à ce jour, pas encore totalement établies. Cette difficulté repose principalement sur le fait que les relations liant les estimées aux données sont, la plupart du temps, non linéaires. Des exigences particulières sur le rang de certaines matrices ont été énoncées au sein des paragraphes 3.3.1 et 3.3.2 pour assurer la consistance des estimées calculées à partir des algorithmes Ordinary et PI/PO MOESP. Ces contraintes sont malencontreusement difficiles à vérifier en pratique. Des travaux de recherche ont donc été conduits pour proposer des conditions sur les paramètres de construction et les données d'entrée-sortie suffisantes à l'obtention de modèles fiables. Plusieurs auteurs [VOWL93, DPS95, PSD96, JW98, BDS99, BJ00, Bau03b, Bau04, CP04] ont plus particulièrement démontré que les méthodes hors ligne des sous-espaces fournissaient des estimées consistantes sous certaines conditions d'excitation et que la normalité asymptotique de ces dernières était également vérifiée. Une présentation détaillée de certaines propriétés statistiques des algorithmes MOESP est proposée dans la suite de ce paragraphe. Celle-ci permettra au lecteur de connaître des conditions suffisantes pour assurer la fiabilité du modèle identifié, conditions qu'il pourra imposer lors de l'élaboration de la procédure d'identification. Cette analyse sera suivie d'une description des principales méthodes hors ligne d'estimation d'ordre du modèle du procédé.

3.3.5.1 Analyse statistique des algorithmes MOESP

Avant de présenter les principales propriétés statistiques des algorithmes MOESP, il est nécessaire de définir ce que nous entendons par condition d'excitation persistante, consistance et normalité asymptotique, ces notions étant utilisées tout au long de cette analyse [SS89, Bau04].

Définition 3.1. \mathbf{u} est un signal d'excitation persistante d'ordre n si et seulement si :

1. la limite suivante existe :

$$\mathbf{R}_u(\tau) = \lim_{M \rightarrow \infty} \sum_{t=1}^M \mathbf{u}(t + \tau) \mathbf{u}^T(t) ; \quad (3.114)$$

2. la matrice

$$\begin{bmatrix} \mathbf{R}_u(0) & \mathbf{R}_u(1) & \cdots & \mathbf{R}_u(n-1) \\ \mathbf{R}_u(1) & \mathbf{R}_u(0) & \cdots & \mathbf{R}_u(n-2) \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{R}_u(n-1) & \mathbf{R}_u(n-2) & \cdots & \mathbf{R}_u(0) \end{bmatrix} \quad (3.115)$$

est non singulière.

Cette définition peut également être réécrite comme suit [VD92a] :

Définition 3.2. $\mathbf{u} \in \mathbb{R}^{n_u \times 1}$ est un signal d'excitation persistante d'ordre n si et seulement si :

$$\text{rang} \left(\lim_{M \rightarrow \infty} \mathbf{U}_f^+ \mathbf{U}_f^{+T} \right) = n_u f \quad (3.116)$$

où \mathbf{U}_f^+ est la matrice de Hankel d'ordre f de \mathbf{u} .

Définition 3.3. Les estimées $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$ et $\hat{\mathbf{D}}$ sont consistantes s'il existe une matrice \mathbf{T} déterministe et non singulière telle que $\hat{\mathbf{A}} - \mathbf{TAT}^{-1}$, $\hat{\mathbf{B}} - \mathbf{TB}$, $\hat{\mathbf{C}} - \mathbf{CT}^{-1}$ et $\hat{\mathbf{D}} - \mathbf{D}$ convergent vers zéro avec une probabilité de 1.

Définition 3.4. Les estimées $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$ et $\hat{\mathbf{D}}$ sont asymptotiquement normales si et seulement si leur distribution asymptotique suit une loi normale.

Les propriétés statistiques majeures de la classe d'identification MOESP sont les suivantes :

- l'algorithme Ordinary MOESP ne peut fournir d'estimées consistantes des matrices d'état du procédé que si la condition d'excitation persistante est vérifiée et si le bruit de sortie \mathbf{v} est un bruit blanc gaussien de moyenne nulle et de variance finie [Ver93, Vib95] ;
- les algorithmes PI et PO MOESP conduisent à des estimées consistantes lorsque la condition d'excitation persistante est assurée et que [Ver94, CV97] :
 - le bruit d'état \mathbf{w} est nul et le bruit de sortie \mathbf{v} est un bruit coloré quelconque ;
 - le bruit de sortie \mathbf{v} et le bruit de procédé \mathbf{w} sont deux bruits blancs gaussiens.

Les démonstrations proposées par M. Verhaegen demandent de garantir que certaines matrices soient de rang plein. Ainsi :

- Ordinary MOESP nécessite que :

$$\text{rang} \left(\mathbf{U}_f^+ \right) = n_u f. \quad (3.117)$$

– PI et PO MOESP requièrent de vérifier que :

$$\text{rang} \left(\begin{bmatrix} \mathbf{U}_p^- \\ \mathbf{U}_f^+ \end{bmatrix} \right) = n_u(p+f) \text{ et } \text{rang}(\mathbf{X}^+\mathbf{\Xi}) = n_x. \quad (3.118)$$

Malheureusement, ces conditions de rang ne peuvent être vérifiées qu'*a posteriori*. Il serait plus intéressant de proposer à l'utilisateur des contraintes, ne serait-ce que suffisantes, sur les données d'entrée et sur le système pour assurer la consistance des estimées.

Ordinary MOESP Comme cela a été présenté au sein du paragraphe 3.3.1, la classe d'identification Ordinary MOESP repose sur la projection orthogonale $\mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{++}}$. Il est alors assez facile de constater que la matrice de covariance des résidus $\mathbf{R}_\epsilon^{\text{Ord}} = \mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{++}} \mathbf{Y}_f^{+T}$ tend vers [JW98] :

$$\mathbf{R}_\epsilon^{\text{Ord}} = \mathbf{\Gamma}_f \left(\mathbf{R}_x - \mathbf{R}_{\mathbf{xu}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{\mathbf{xu}_f^+}^T \right) \mathbf{\Gamma}_f^T + \mathbf{R}_{\mathbf{v}_f^+} \quad (3.119)$$

lorsque $M \rightarrow \infty$. L'estimation consistante de $\mathbf{\Gamma}_f$ est alors garantie si et seulement si :

$$\mathbf{R}_{\mathbf{v}_f^+} = \sigma^2 \mathbf{I} \quad (3.120)$$

$$\text{rang} \left(\mathbf{R}_x - \mathbf{R}_{\mathbf{xu}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{\mathbf{xu}_f^+}^T \right) = n_x. \quad (3.121)$$

Or, puisque $\mathbf{R}_x - \mathbf{R}_{\mathbf{xu}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{\mathbf{xu}_f^+}^T$ est symétrique, une condition suffisante à la propriété (3.121) est de s'assurer que :

$$\mathbf{R}_x - \mathbf{R}_{\mathbf{xu}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{\mathbf{xu}_f^+}^T > 0. \quad (3.122)$$

En appliquant le théorème suivant [RAB⁺01] :

Théorème 3.1. Soit $\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3$ et \mathbf{M}_4 quatre matrices de dimension respectives $p \times p$, $p \times q$, $q \times p$ et $q \times q$ telles que :

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_1 & \mathbf{M}_2 \\ \mathbf{M}_3 & \mathbf{M}_4 \end{bmatrix} \in \mathbb{R}^{(p+q) \times (p+q)}. \quad (3.123)$$

Supposons de plus que la matrice \mathbf{M}_4 soit inversible. Le complément de Schur de la matrice \mathbf{M}_4 est alors défini par :

$$\mathbf{S}_{\mathbf{M}_4} = \mathbf{M}_1 - \mathbf{M}_2 \mathbf{M}_4^{-1} \mathbf{M}_3. \quad (3.124)$$

Il vérifie la propriété suivante suivante :

$$\mathbf{M} > 0 \Rightarrow \mathbf{S}_{\mathbf{M}_4} > 0. \quad (3.125)$$

De plus :

$$\text{rang}(\mathbf{M}) = \text{rang}(\mathbf{M}_4) + \text{rang}(\mathbf{S}_{\mathbf{M}_4}). \quad (3.126)$$

une condition suffisante pour vérifier la relation (3.122) est que :

$$\begin{bmatrix} \mathbf{R}_x & \mathbf{R}_{\mathbf{xu}_f^+} \\ \mathbf{R}_{\mathbf{xu}_f^+}^T & \mathbf{R}_{\mathbf{u}_f^+} \end{bmatrix} = \mathbb{E} \left\{ \begin{bmatrix} \mathbf{x} \\ \mathbf{u}_f^+ \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{u}_f^+ \end{bmatrix}^T \right\} > 0 \quad (3.127)$$

puisque $\mathbf{R}_x - \mathbf{R}_{x\mathbf{u}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{x\mathbf{u}_f^+}^T$ est le complément de Schur de $\mathbf{R}_{\mathbf{u}_f^+}$. M. Jansson et B. Wahlberg [JW98] montrent que cette condition est équivalente à :

$$\mathbb{E} \left\{ \frac{q^{n_x}}{\det(q\mathbf{I} - \mathbf{A})} \begin{bmatrix} \mathbf{u}(t - n_x) \\ \vdots \\ \mathbf{u}(t + f - 1) \\ \vdots \\ \mathbf{e}(t - n_x) \\ \vdots \\ \mathbf{e}(t - 1) \end{bmatrix} \frac{q^{n_x}}{\det(q\mathbf{I} - \mathbf{A})} \begin{bmatrix} \mathbf{u}(t - n_x) \\ \vdots \\ \mathbf{u}(t + f - 1) \\ \vdots \\ \mathbf{e}(t - n_x) \\ \vdots \\ \mathbf{e}(t - 1) \end{bmatrix}^T \right\} > 0 \quad (3.128)$$

où q est l'opérateur avance et \mathbf{e} un bruit blanc gaussien. En effet, en décomposant le vecteur d'état \mathbf{x} en un terme déterministe \mathbf{x}_d et un élément stochastique \mathbf{x}_s respectivement liés au signal d'entrée \mathbf{u} par la relation¹³ :

$$\mathbf{x}_d(t) = \frac{\text{Com}^T(p\mathbf{I} - \mathbf{A})}{\det(p\mathbf{I} - \mathbf{A})} \mathbf{u}(t) \quad (3.129)$$

et au vecteur d'innovation \mathbf{e} comme suit :

$$\mathbf{x}_s(t) = \frac{\text{Com}^T(p\mathbf{I} - \mathbf{A})}{\det(p\mathbf{I} - \mathbf{A})} \mathbf{e}(t), \quad (3.130)$$

le bloc matriciel $\begin{bmatrix} \mathbf{x}^T(t) & \mathbf{u}_f^{+T}(t) \end{bmatrix}^T$ vérifie [JW98] :

$$\begin{bmatrix} \mathbf{x}(t) \\ \mathbf{u}_f^+(t) \end{bmatrix} = \mathcal{S} \frac{q^{n_x}}{\det(q\mathbf{I} - \mathbf{A})} \begin{bmatrix} \mathbf{u}(t - n_x) \\ \vdots \\ \mathbf{u}(t + f - 1) \\ \vdots \\ \mathbf{e}(t - n_x) \\ \vdots \\ \mathbf{e}(t - 1) \end{bmatrix} \quad (3.131)$$

où \mathcal{S} est une matrice de rang plein. Puisque \mathbf{u} et \mathbf{e} sont décorrélés, une condition suffisante à la consistance des estimées est d'assurer que le signal d'entrée est excité de manière persistante d'ordre au moins égal à $f + n_x$. Le théorème suivant est alors démontré [JW98] :

Théorème 3.2. *l'algorithme Ordinary MOESP fournit des estimées consistantes des matrices d'état du système si :*

1. *la matrice de covariance du bruit est proportionnelle à l'identité ;*
2. *le signal d'entrée est excité de manière persistante d'ordre $f + n_x$.*

¹³Com^T est la transposée de la comatrice [Rot01].

PI et PO MOESP La consistance des estimées obtenues par une approche de type variable instrumentale est plus complexe à démontrer. Plusieurs paramètres sont en effet à prendre en compte. Ces derniers sont mis en évidence en considérant, comme précédemment, la matrice des résidus $\mathbf{R}_\varepsilon^{\text{PI/PO}}$ définie par¹⁴ :

$$\mathbf{R}_\varepsilon^{\text{PI/PO}} = \mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T \left(\mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T \right)^{-1} \mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{Y}_f^+. \quad (3.132)$$

Cette matrice vérifie asymptotiquement que (cf. annexe A.1) :

$$\begin{aligned} \mathbf{R}_\varepsilon^{\text{PI/PO}} &= \mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T \left(\mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T \right)^{-1} \mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{Y}_f^+ \\ &= \mathbf{\Gamma}_f \left(\mathbf{R}_{\mathbf{x}\xi} - \mathbf{R}_{\mathbf{xu}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{\mathbf{u}_f^+ \xi} \right) \left(\mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T \right)^{-1} \left(\mathbf{R}_{\mathbf{x}\xi} - \mathbf{R}_{\mathbf{xu}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{\mathbf{u}_f^+ \xi} \right)^T \mathbf{\Gamma}_f^T, \end{aligned} \quad (3.133)$$

L'estimation de $\mathbf{\Gamma}_f$ est donc consistante si et seulement si les blocs matriciels $\mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T$ et $\mathbf{R}_{\mathbf{x}\xi} - \mathbf{R}_{\mathbf{xu}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{\mathbf{u}_f^+ \xi}$ sont de rang plein.

Commençons tout d'abord par préciser les conditions assurant la non singularité de la matrice $\mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T$. Puisque cette dernière est symétrique et peut facilement se récrire comme suit :

$$\mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T = \mathbf{\Xi} \mathbf{\Xi}^T - \mathbf{\Xi} \mathbf{U}_f^{+T} \left(\mathbf{U}_f^+ \mathbf{U}_f^{+T} \right) \mathbf{U}_f^+ \mathbf{\Xi}^T \quad (3.134)$$

par définition de la projection orthogonale, une condition suffisante à la non singularité de $\mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T$ (cf. th. 3.1) est de s'assurer que :

$$\lim_{M \rightarrow \infty} \begin{bmatrix} \mathbf{\Xi} \\ \mathbf{U}_f^+ \end{bmatrix} \begin{bmatrix} \mathbf{\Xi} \\ \mathbf{U}_f^+ \end{bmatrix}^T > 0. \quad (3.135)$$

Cette condition est équivalente à celle présentée en complément C5.1 du livre de T. Söderström et P. Stoica [SS89]. Elle est vérifiée lorsque le signal d'entrée est excité de manière persistante d'ordre au moins égal $f + p$.

Une condition suffisante à la non singularité du terme $\mathbf{R}_{\mathbf{x}\xi} - \mathbf{R}_{\mathbf{xu}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{\mathbf{u}_f^+ \xi}$ est obtenue en considérant la matrice suivante :

$$\begin{bmatrix} \mathbf{R}_{\mathbf{x}\xi} & \mathbf{R}_{\mathbf{xu}_f^+} \\ \mathbf{R}_{\mathbf{u}_f^+ \xi} & \mathbf{R}_{\mathbf{u}_f^+} \end{bmatrix} \quad (3.136)$$

et le complément de Schur de $\mathbf{R}_{\mathbf{u}_f^+}$:

$$\mathbf{S}_{\mathbf{R}_{\mathbf{u}_f^+}} = \mathbf{R}_{\mathbf{x}\xi} - \mathbf{R}_{\mathbf{xu}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{\mathbf{u}_f^+ \xi}. \quad (3.137)$$

¹⁴Nous montrerons dans le paragraphe 3.4 que $\bar{\mathbf{R}}_{32} \bar{\mathbf{Q}}_2 = \mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T \left(\mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T \right)^{-1} \mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^+}$.

En effet, puisque (cf. th. 3.1) :

$$\text{rang} \left(\begin{bmatrix} \mathbf{R}_{\mathbf{x}\xi} & \mathbf{R}_{\mathbf{x}\mathbf{u}_f^+} \\ \mathbf{R}_{\mathbf{u}_f^+\xi} & \mathbf{R}_{\mathbf{u}_f^+} \end{bmatrix} \right) = \text{rang} \left(\mathbf{S}_{\mathbf{R}_{\mathbf{u}_f^+}} \right) + \text{rang} \left(\mathbf{R}_{\mathbf{u}_f^+} \right) \quad (3.138)$$

et que $\mathbf{R}_{\mathbf{u}_f^+}$ est une matrice symétrique définie positive de rang $n_{\mathbf{u}_f}$, une condition suffisante pour assurer que $\mathbf{R}_{\mathbf{x}\xi} - \mathbf{R}_{\mathbf{x}\mathbf{u}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{\mathbf{u}_f^+\xi}$ soit de rang plein est que :

$$\text{rang} \left(\begin{bmatrix} \mathbf{R}_{\mathbf{x}\xi} & \mathbf{R}_{\mathbf{x}\mathbf{u}_f^+} \\ \mathbf{R}_{\mathbf{u}_f^+\xi} & \mathbf{R}_{\mathbf{u}_f^+} \end{bmatrix} \right) = \text{rang} \left(\mathbb{E} \left\{ \begin{bmatrix} \mathbf{x} \\ \mathbf{u}_f^+ \end{bmatrix} \begin{bmatrix} \xi \\ \mathbf{u}_f^+ \end{bmatrix}^T \right\} \right) = n_x + n_{\mathbf{u}_f}. \quad (3.139)$$

Plusieurs signaux d'entrée vérifiant cette dernière condition ont été proposés dans la littérature [Jan97, JW98]. N. L. C. Chiu ([Chi97] th. 5.15) démontre que celle-ci est assurée si :

1. $f \geq \partial$ où ∂ est le degré de l'*A-annihilateur* défini comme le polynôme P de degré minimum tel que $P(\mathbf{A}) = 0$;
2. $p \geq 3\partial$;
3. le signal d'entrée est excité de manière persistante d'ordre $f + p + \partial$.

Quatre conditions suffisantes peuvent donc être énoncées pour assurer la consistance de l'estimée de $\mathbf{\Gamma}_f$. Nous pouvons cependant remarquer que la dernière condition suffisante à la non singularité de $\mathbf{R}_{\mathbf{x}\xi} - \mathbf{R}_{\mathbf{x}\mathbf{u}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{\mathbf{u}_f^+\xi}$ contient celle proposée pour imposer la non singularité de $\mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T$. L'utilisation de signaux et d'indices f et p vérifiant les trois conditions avancées par N. L. C. Chiu [Chi97] permet donc d'obtenir une estimée consistante de la matrice d'observabilité du système.

Remarque 3.7. *Il est nécessaire de préciser que ces règles sont difficilement utilisables en pratique. En effet, le calcul du paramètre ∂ nécessite de connaître la matrice d'état \mathbf{A} qui n'est seulement connue qu'au terme de l'identification du procédé. On peut cependant remarquer que, si nous fixons f et p « suffisamment » grands, aucun problème de consistance n'apparaîtra.*

Remarque 3.8. *Les matrices de résidus $\mathbf{R}_{\boldsymbol{\varepsilon}}^{\text{Ord}}$ et $\mathbf{R}_{\boldsymbol{\varepsilon}}^{\text{PI/PO}}$ peuvent être utilisées pour estimer la matrice d'observabilité du système [Liu92, Vib95]. En effet, la relation (3.119) (resp. (3.133)) indique clairement que les vecteurs propres de $\mathbf{R}_{\boldsymbol{\varepsilon}}^{\text{Ord}}$ (resp. $\mathbf{R}_{\boldsymbol{\varepsilon}}^{\text{PI/PO}}$) coïncident avec les vecteurs singuliers gauche de $\mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^+}$ (resp. $\mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T \mathbf{P}_d$). Le sous-espace estimé à partir du bloc matriciel $\mathbf{R}_{\boldsymbol{\varepsilon}}^{\text{Ord}}$ (resp. $\mathbf{R}_{\boldsymbol{\varepsilon}}^{\text{PI/PO}}$) est donc identique à celui obtenu via $\mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^+}$ (resp. $\mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T \mathbf{P}_d$).*

D'autres propriétés asymptotiques des schémas d'identification MOESP ont été présentées dans la littérature [BJ00, Bau03b, Bau04, CP04] (normalité asymptotique des estimées, variance de la distribution des estimées...) Elles demandent malheureusement de fortes conditions sur les données d'entrée. Elles ne seront donc pas détaillées dans ce manuscrit.

3.3.5.2 Choix des paramètres structuraux

Bien que l'estimation de l'ordre soit un point crucial en identification des sous-espaces, très peu d'articles traitent de cette problématique. La majorité des techniques développées [Ver94, VD94, Pet95, Bau98, Bau99, Bau01] reposent sur l'inspection des valeurs singulières obtenues lors de la décomposition en valeurs singulières de blocs matriciels particuliers¹⁵. Malheureusement, la détermination d'un creux ou d'un coude au sein même du tracé de ces estimées n'est pas toujours une chose évidente. Les approches les plus efficaces reposent sur une imitation du critère d'information AIC [Aka74]. Elles proposent d'examiner les valeurs singulières en ajoutant à ces dernières un facteur pénalisant lié au nombre de paramètres à calculer et à la complexité du modèle. Deux critères ont été proposés pour le schéma d'identification MOESP [Pet95, Bau99] :

$$\text{NIC}(n_x) = \sum_{i=n_x+1}^{\bar{M}} \hat{\sigma}_i^2 + \frac{d(n_x)J(M)}{M} \quad (3.140)$$

$$\text{SVC}(n_x) = \hat{\sigma}_{n_x+1}^2 + \frac{d(n_x)J(M)}{M} \quad (3.141)$$

où $d(n_x) = 2n_y n_x + n_y n_u + n_u n_x$ est le nombre de paramètres du système d'état d'ordre n_x , $J(M)$ le terme de pénalité et $\bar{M} = \min(n_y f, (n_y + n_u)p)$ le nombre de valeurs singulières estimées. Il est montré dans l'article [Bau99] que :

Théorème 3.3. *Le critère SVC fournit une estimation consistante de n_x si :*

- $f \geq n_x + 1$ et $p \geq n_x$ avec $\max(f, p) = o((\log(M))^a)$, $a < \infty$;
- $J(M) > 0$, $J(M)/M \rightarrow 0$ et $J(M)/(\log(\log(M))) \rightarrow \infty$.

Aucune démonstration n'est proposée pour le critère NIC. Il semble cependant naturel d'étendre le théorème précédent à cette fonction coût. La seule différence entre NIC et SVC repose en effet sur le choix de la norme utilisée (norme de Frobenius pour NIC et norme 2 pour SVC).

En pratique [Mer01, Bau03a], choisir $f = p = 2\hat{n}_x^{AIC}$, où \hat{n}_x^{AIC} est l'ordre estimé par le critère AIC, permet de remplir les hypothèses précédentes. De bons résultats sont obtenus lorsque $J(M) = fp \log(M)$ est adopté [Mer01, Bau03a]. De nombreuses simulations [Mer01, Bau01] indiquent que le critère SVC est plus robuste au choix des indices f et p que NIC. Malheureusement, aucune combinaison optimale n'a pour l'instant été trouvée. La principale difficulté repose en effet sur l'imbrication des indices structuraux f , p et n_x .

3.4 Variable instrumentale et théorèmes unifiés

Dans les paragraphes 3.3.2 et 3.3.3, deux variables instrumentales ont été proposées afin d'estimer, de manière consistante, une base de la matrice d'observabilité du système dans un contexte stochastique. Leur introduction au sein de factorisations QR particulières (cf. equ.

¹⁵Les matrices utilisées par la classe d'identification MOESP sont respectivement :

- $\bar{\mathbf{R}}_{22}$ de la décomposition QR d'Ordinary MOESP dans le cas déterministe ;
- $\bar{\mathbf{R}}_{32}$ de la décomposition QR de PI/PO MOESP dans le cas stochastique.

(3.84) et (3.88)) a ainsi permis de supprimer l'effet des perturbations tout en éliminant l'influence du régime forcé. Ces écritures, intéressantes d'un point de vue numérique, présentent l'inconvénient de ne pas mettre en évidence l'action des variables instrumentales au sein des algorithmes MOESP. L'objectif de cette section consiste donc à écrire explicitement ces factorisations en terme de projections orthogonales et de variables instrumentales afin de démontrer que les algorithmes MOESP partagent une écriture commune, écriture que nous relierons au théorème unifié proposé par P. Van Overschee et B. De Moor [VD95].

L'algorithme Ordinary MOESP utilise la factorisation QR suivante (cf §3.3.1) :

$$\begin{bmatrix} \mathbf{U}_f^+ \\ \mathbf{Y}_f^+ \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_{11} & \mathbf{0} \\ \bar{\mathbf{R}}_{21} & \bar{\mathbf{R}}_{22} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1 \\ \bar{\mathbf{Q}}_2 \end{bmatrix}. \quad (3.142)$$

Il est alors facile de vérifier que :

$$\bar{\mathbf{R}}_{22} \bar{\mathbf{Q}}_2 = \mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^{+\perp}} \quad (3.143)$$

puisque :

$$\mathbf{U}_f^+ = \bar{\mathbf{R}}_{11} \bar{\mathbf{Q}}_1 \quad (3.144)$$

$$\mathbf{Y}_f^+ = \bar{\mathbf{R}}_{21} \bar{\mathbf{Q}}_1 + \bar{\mathbf{R}}_{22} \bar{\mathbf{Q}}_2 \quad (3.145)$$

et que [GV96] :

$$\Pi_{\mathbf{U}_f^{+\perp}} = \mathbf{I} - \mathbf{U}_f^{+T} \left(\mathbf{U}_f^+ \mathbf{U}_f^{+T} \right)^{-1} \mathbf{U}_f^+. \quad (3.146)$$

En effet :

$$\begin{aligned} \mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^{+\perp}} &= \mathbf{Y}_f^+ - \mathbf{Y}_f^+ \mathbf{U}_f^{+T} \left(\mathbf{U}_f^+ \mathbf{U}_f^{+T} \right)^{-1} \mathbf{U}_f^+ \\ &= \bar{\mathbf{R}}_{21} \bar{\mathbf{Q}}_1 + \bar{\mathbf{R}}_{22} \bar{\mathbf{Q}}_2 \\ &\quad - \left(\bar{\mathbf{R}}_{21} \bar{\mathbf{Q}}_1 + \bar{\mathbf{R}}_{22} \bar{\mathbf{Q}}_2 \right) \bar{\mathbf{Q}}_1^T \bar{\mathbf{R}}_{11}^T \underbrace{\left(\bar{\mathbf{R}}_{11} \bar{\mathbf{Q}}_1 \bar{\mathbf{Q}}_1^T \bar{\mathbf{R}}_{11}^T \right)^{-1}}_{\mathbf{I}} \bar{\mathbf{R}}_{11} \bar{\mathbf{Q}}_1 \\ &= \bar{\mathbf{R}}_{21} \bar{\mathbf{Q}}_1 + \bar{\mathbf{R}}_{22} \bar{\mathbf{Q}}_2 - \bar{\mathbf{R}}_{21} \underbrace{\bar{\mathbf{R}}_{11}^T \bar{\mathbf{R}}_{11}^{-T}}_{\mathbf{I}} \underbrace{\bar{\mathbf{R}}_{11}^{-1} \bar{\mathbf{R}}_{11}}_{\mathbf{I}} \bar{\mathbf{Q}}_1 = \bar{\mathbf{R}}_{22} \bar{\mathbf{Q}}_2 \end{aligned} \quad (3.147)$$

puisque $\bar{\mathbf{Q}}_1$ et $\bar{\mathbf{Q}}_2$ sont des matrices orthogonales et que $\bar{\mathbf{R}}_{11}$ est inversible.

Les algorithmes PI et PO MOESP sont fondés sur la décomposition suivante (cf. §3.3.2) :

$$\begin{bmatrix} \mathbf{U}_f^+ \\ \bar{\mathbf{\Xi}} \\ \mathbf{Y}_f^+ \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_{11} & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{21} & \bar{\mathbf{R}}_{22} & \mathbf{0} \\ \bar{\mathbf{R}}_{31} & \bar{\mathbf{R}}_{32} & \bar{\mathbf{R}}_{33} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1 \\ \bar{\mathbf{Q}}_2 \\ \bar{\mathbf{Q}}_3 \end{bmatrix} \quad (3.148)$$

où $\bar{\mathbf{\Xi}} = \mathbf{U}_p^-$ pour l'algorithme PI MOESP et $\bar{\mathbf{\Xi}} = \left[\mathbf{U}_p^{-T} \mathbf{Y}_p^{-T} \right]^T$ pour PO MOESP. La décomposition du facteur $\bar{\mathbf{R}}_{32} \bar{\mathbf{Q}}_2$ en produit de projections et de variables instrumentales est ici plus difficile. Considérons donc l'astuce suivante¹⁶ :

$$\bar{\mathbf{R}}_{32} \bar{\mathbf{Q}}_2 = \bar{\mathbf{R}}_{32} \bar{\mathbf{Q}}_2 + \bar{\mathbf{R}}_{33} \bar{\mathbf{Q}}_3 - \bar{\mathbf{R}}_{33} \bar{\mathbf{Q}}_3 \quad (3.149)$$

¹⁶Cette démonstration s'inspire de celle proposée par M. Viberg en 1995 [Vib95] pour l'algorithme PO MOESP. Elle la généralise en combinant PI et PO MOESP.

qui consiste à ajouter puis soustraire le terme $\bar{\mathbf{R}}_{33}\bar{\mathbf{Q}}_3$. Il est alors facile de voir que cette expression traduit l'égalité suivante :

$$\bar{\mathbf{R}}_{32}\bar{\mathbf{Q}}_2 = \mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^{\perp}} - \mathbf{Y}_f^+ \Pi_{\Psi^{\perp}} = \mathbf{Y}_f^+ \left(\Pi_{\mathbf{U}_f^{\perp}} - \Pi_{\Psi^{\perp}} \right) = \mathbf{Y}_f^+ \left(\Pi_{\Psi} - \Pi_{\mathbf{U}_f^+} \right) \quad (3.150)$$

où :

$$\Psi = \begin{bmatrix} \mathbf{U}_f^+ \\ \Xi \end{bmatrix} \quad (3.151)$$

puisque :

$$\bar{\mathbf{R}}_{32}\bar{\mathbf{Q}}_2 + \bar{\mathbf{R}}_{33}\bar{\mathbf{Q}}_3 = \mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^{\perp}} \quad (3.152)$$

$$\bar{\mathbf{R}}_{33}\bar{\mathbf{Q}}_3 = \mathbf{Y}_f^+ \Pi_{\Psi^{\perp}}. \quad (3.153)$$

Or, en utilisant le lemme d'inversion matricielle par bloc [Bro91], il est possible de montrer que (cf. annexe A.2) :

$$\Pi_{\Psi} = \Pi_{\mathbf{U}_f^+} + \Pi_{\mathbf{U}_f^{\perp}} \Xi^T \left(\Xi \Pi_{\mathbf{U}_f^{\perp}} \Xi^T \right)^{-1} \Xi \Pi_{\mathbf{U}_f^{\perp}}. \quad (3.154)$$

Ainsi, en combinant cette expression avec l'équation (3.150), $\bar{\mathbf{R}}_{32}\bar{\mathbf{Q}}_2$ vérifie :

$$\bar{\mathbf{R}}_{32}\bar{\mathbf{Q}}_2 = \mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^{\perp}} \Xi^T \left(\Xi \Pi_{\mathbf{U}_f^{\perp}} \Xi^T \right)^{-1} \Xi \Pi_{\mathbf{U}_f^{\perp}}. \quad (3.155)$$

L'algorithme PO_EIV MOESP considère la décomposition QR suivante (cf. §3.3.3) :

$$\begin{bmatrix} \mathbf{U}_f^+ \mathbf{U}_p^{-T} & \mathbf{U}_f^+ \mathbf{Y}_p^{-T} \\ \mathbf{Y}_f^+ \mathbf{U}_p^{-T} & \mathbf{Y}_f^+ \mathbf{Y}_p^{-T} \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_{11} & \mathbf{0} \\ \bar{\mathbf{R}}_{21} & \bar{\mathbf{R}}_{22} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1 \\ \bar{\mathbf{Q}}_2 \end{bmatrix} \quad (3.156)$$

qu'il est aisé de récrire comme suit :

$$\begin{bmatrix} \mathbf{U}_f^+ \\ \mathbf{Y}_f^+ \end{bmatrix} \Xi^T = \begin{bmatrix} \bar{\mathbf{R}}_{11} & \mathbf{0} \\ \bar{\mathbf{R}}_{21} & \bar{\mathbf{R}}_{22} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1 \\ \bar{\mathbf{Q}}_2 \end{bmatrix}. \quad (3.157)$$

Ainsi :

$$\mathbf{U}_f^+ \Xi^T = \bar{\mathbf{R}}_{11} \bar{\mathbf{Q}}_1 \quad (3.158)$$

$$\mathbf{Y}_f^+ \Xi^T = \bar{\mathbf{R}}_{21} \bar{\mathbf{Q}}_1 + \bar{\mathbf{R}}_{22} \bar{\mathbf{Q}}_2. \quad (3.159)$$

En appliquant une démonstration similaire à celle développée pour Ordinary MOESP, il est facile de montrer que :

$$\bar{\mathbf{R}}_{22} \bar{\mathbf{Q}}_2 = \mathbf{Y}_f^+ \Xi^T \Pi_{(\mathbf{U}_f^+ \Xi^T)^{\perp}}. \quad (3.160)$$

| | \mathbf{P}_g | \mathbf{O} | \mathbf{P}_d |
|----------------|----------------|--|---|
| Ordinary MOESP | \mathbf{I} | $\mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^{+\perp}}$ | \mathbf{I} |
| PI/PO MOESP | \mathbf{I} | $\mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^{+\perp}} \Xi^T$ | $\left(\Xi \Pi_{\mathbf{U}_f^{+\perp}} \Xi^T \right)^{-1} \Xi \Pi_{\mathbf{U}_f^{+\perp}}$ |
| PO.EIV MOESP | \mathbf{I} | $\mathbf{Y}_f^+ \Xi^T \Pi_{\left(\mathbf{U}_f^+ \Xi^T \right)^\perp}$ | \mathbf{I} |

TAB. 3.2: Écriture unifiée des algorithmes MOESP.

Les équations (3.143), (3.155) et (3.160) peuvent alors être regroupées sous la forme compacte suivante (cf. tab. 3.2) :

$$\mathbf{P}_g \mathbf{O} \mathbf{P}_d. \quad (3.161)$$

L'estimation de la matrice d'observabilité étendue est ensuite réalisable en extrayant les n_x principaux vecteurs singuliers gauche de la décomposition en valeurs singulières suivante :

$$\mathbf{P}_g \mathbf{O} \mathbf{P}_d = [\bar{\mathbf{U}}_s \quad \bar{\mathbf{U}}_b] \begin{bmatrix} \bar{\Sigma}_s & \mathbf{0} \\ \mathbf{0} & \bar{\Sigma}_b \end{bmatrix} \begin{bmatrix} \bar{\mathbf{V}}_s^T \\ \bar{\mathbf{V}}_b^T \end{bmatrix}. \quad (3.162)$$

La décomposition (3.161) est par ailleurs généralisable aux méthodes d'identification des sous-espaces N4SID et CVA [Vib95, VD95, VWO97] :

L'algorithme N4SID estime la matrice d'observabilité du système en appliquant une DVS au résultat de la projection oblique des sorties futures sur les observations passées selon les directions des entrées futures [VD94] :

$$\mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^+}^{\Xi} \quad (3.163)$$

où $\Pi_{\mathbf{U}_f^+}^{\Xi}$ symbolise la projection oblique sur les données d'entrée futures \mathbf{U}_f^+ parallèlement à la variable instrumentale $\Xi = \left[\mathbf{U}_p^{-T} \quad \mathbf{Y}_p^{-T} \right]^T$. Puisque (cf. [VD96b] p. 22) :

$$\Pi_{\mathbf{U}_f^+}^{\Xi} = \Pi_{\mathbf{U}_f^{+\perp}} \left(\Xi \Pi_{\mathbf{U}_f^{+\perp}} \right)^\dagger \Xi, \quad (3.164)$$

la projection (3.163) s'écrit alors :

$$\mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^+}^{\Xi} = \mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^{+\perp}} \Xi^T \left(\Xi \Pi_{\mathbf{U}_f^{+\perp}} \Xi^T \right)^{-1} \Xi. \quad (3.165)$$

L'algorithme CVA est quant à lui fondé sur les notions d'angles et de directions principales entre espaces vectoriels. De manière simplifiée, il considère les angles et directions principales entre $\Xi \Pi_{\mathbf{U}_f^{+\perp}}$ et $\mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^{+\perp}}$ avec $\Xi = \left[\mathbf{U}_p^{-T} \quad \mathbf{Y}_p^{-T} \right]^T$. La décomposition en valeurs

singulières du produit matriciel :

$$\mathcal{R}^{-1/2} \left\{ \mathbf{Y}_f^+ \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}}, \mathbf{Y}_f^+ \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}} \right\} \left[\mathbf{Y}_f^+ \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}} \left(\boldsymbol{\Xi} \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}} \right)^T \right] \mathcal{R}^{-1/2} \left\{ \boldsymbol{\Xi} \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}}, \boldsymbol{\Xi} \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}} \right\} \quad (3.166)$$

est alors réalisée afin d'extraire une base de la matrice d'observabilité étendue [VD95], $\mathcal{R}_{\{\mathbf{M}, \mathbf{N}\}}$ étant la covariance entre les matrices $\mathbf{M} \in \mathbb{R}^{m \times j}$ et $\mathbf{N} \in \mathbb{R}^{n \times j}$:

$$\mathcal{R}_{\{\mathbf{M}, \mathbf{N}\}} = \lim_{j \rightarrow \infty} [\mathbf{M} \mathbf{N}^T]. \quad (3.167)$$

Cette corrélation canonique peut alors être réécrite comme suit :

$$\mathcal{R}^{-1/2} \left\{ \mathbf{Y}_f^+ \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}}, \mathbf{Y}_f^+ \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}} \right\} \left[\mathbf{Y}_f^+ \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}} \boldsymbol{\Xi}^T \right] \mathcal{R}^{-1/2} \left\{ \boldsymbol{\Xi} \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}}, \boldsymbol{\Xi} \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}} \right\}. \quad (3.168)$$

Ces écritures, associées au tableau 3.2, démontrent bien que les algorithmes de sous-espaces partagent une même et unique écriture (cf. tab. 3.3) :

$$\mathbf{P}_g \mathbf{O} \mathbf{P}_d. \quad (3.169)$$

| | \mathbf{P}_g | \mathbf{O} | \mathbf{P}_d |
|----------------|--|---|--|
| Ordinary MOESP | \mathbf{I} | $\mathbf{Y}_f^+ \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}}$ | \mathbf{I} |
| PI/PO MOESP | \mathbf{I} | $\mathbf{Y}_f^+ \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}} \boldsymbol{\Xi}^T$ | $\left(\boldsymbol{\Xi} \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}} \boldsymbol{\Xi}^T \right)^{-1} \boldsymbol{\Xi} \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}}$ |
| PO_EIV MOESP | \mathbf{I} | $\mathbf{Y}_f^+ \boldsymbol{\Xi}^T \boldsymbol{\Pi}_{\left(\mathbf{U}_f^+ \boldsymbol{\Xi}^T \right)^\perp}$ | \mathbf{I} |
| N4SID | \mathbf{I} | $\mathbf{Y}_f^+ \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}} \boldsymbol{\Xi}^T$ | $\left(\boldsymbol{\Xi} \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}} \boldsymbol{\Xi}^T \right)^{-1} \boldsymbol{\Xi}$ |
| CVA | $\mathcal{R}^{-1/2} \left\{ \mathbf{Y}_f^+ \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}}, \mathbf{Y}_f^+ \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}} \right\}$ | $\mathbf{Y}_f^+ \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}} \boldsymbol{\Xi}^T$ | $\mathcal{R}^{-1/2} \left\{ \boldsymbol{\Xi} \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}}, \boldsymbol{\Xi} \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}} \right\}$ |

TAB. 3.3: Écriture unifiée des algorithmes 4SID.

Elle permet de mettre en évidence deux remarques fondamentales en identification directe des sous-espaces :

- Les algorithmes PI MOESP, PO MOESP, CVA et N4SID possède la même matrice \mathbf{O} :

$$\mathbf{O} = \mathbf{Y}_f^+ \boldsymbol{\Pi}_{\mathbf{U}_f^{\perp\perp}} \boldsymbol{\Xi}^T. \quad (3.170)$$

Cette propriété, démontrée pour la première fois par B. De Moor et P. Van Overschee dans leur célèbre théorème unifié [VD95], permet de bien comprendre les actions respectives de la projection orthogonale $\Pi_{\mathbf{U}_f^{\perp}}$ et de la variable instrumentale Ξ . En effet, puisque :

$$\mathbf{Y}_f^+ = \Gamma_f \mathbf{X}^+ + \mathbf{H}_f \mathbf{U}_f^+ + \mathbf{N}_f^+, \quad (3.171)$$

nous vérifions que :

$$\mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^{\perp}} = \Gamma_f \mathbf{X}^+ \Pi_{\mathbf{U}_f^{\perp}} + \underbrace{\mathbf{H}_f \mathbf{U}_f^+ \Pi_{\mathbf{U}_f^{\perp}} + \mathbf{N}_f^+ \Pi_{\mathbf{U}_f^{\perp}}}_{=0} = \Gamma_f \mathbf{X}^+ \Pi_{\mathbf{U}_f^{\perp}} + \mathbf{N}_f^+ \Pi_{\mathbf{U}_f^{\perp}} \quad (3.172)$$

correspond simplement à la suppression du régime forcé. En postmultipliant cette expression par Ξ^T , en supposant que les entrées futures¹⁷ et passées ainsi que les sorties passées soient décorréliées du bruit futur, il est facile de vérifier que :

$$\mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^{\perp}} \Xi^T = \Gamma_f \mathbf{X}^+ \Pi_{\mathbf{U}_f^{\perp}} \Xi^T + \underbrace{\mathbf{N}_f^+ \Pi_{\mathbf{U}_f^{\perp}} \Xi^T}_{=0} = \Gamma_f \mathbf{X}^+ \Pi_{\mathbf{U}_f^{\perp}} \Xi^T \quad (3.173)$$

puisque [VWO97] :

$$\begin{aligned} \lim_{M \rightarrow \infty} \mathbf{N}_f^+ \Pi_{\mathbf{U}_f^{\perp}} \mathbf{U}_p^{-T} &= \\ \mathbb{E} \left\{ \mathbf{n}_f^+ \mathbf{u}_p^{-T} \right\} - \mathbb{E} \left\{ \mathbf{n}_f^+ \mathbf{u}_f^{+T} \right\} \mathbb{E} \left\{ \left(\mathbf{u}_f^+ \mathbf{u}_f^{+T} \right)^{-1} \right\} \mathbb{E} \left\{ \mathbf{u}_f^+ \mathbf{u}_p^{-T} \right\} &= \mathbf{0} \end{aligned} \quad (3.174)$$

$$\begin{aligned} \lim_{M \rightarrow \infty} \mathbf{N}_f^+ \Pi_{\mathbf{U}_f^{\perp}} \mathbf{Y}_p^{-T} &= \\ \mathbb{E} \left\{ \mathbf{n}_f^+ \mathbf{y}_p^{-T} \right\} - \mathbb{E} \left\{ \mathbf{n}_f^+ \mathbf{u}_f^{+T} \right\} \mathbb{E} \left\{ \left(\mathbf{u}_f^+ \mathbf{u}_f^{+T} \right)^{-1} \right\} \mathbb{E} \left\{ \mathbf{u}_f^+ \mathbf{y}_p^{-T} \right\} &= \mathbf{0}. \end{aligned} \quad (3.175)$$

- Les égalités précédentes sont malheureusement fausses lorsque le bruit de mesure \mathbf{v} , associé à un vecteur d'entrée $\tilde{\mathbf{u}}$ non blanc, est non nul. La solution consiste alors à inverser les étapes d'application de la variable instrumentale et de la projection orthogonale :

$$\mathbf{O} = \mathbf{Y}_f^+ \Xi^T \Pi_{\left(\mathbf{U}_f^+ \Xi^T \right)^{\perp}}. \quad (3.176)$$

En effet, puisque Ξ est choisie de telle sorte que :

$$\lim_{M \rightarrow \infty} \mathbf{N}_f^+ \Xi^T = \mathbf{0}, \quad (3.177)$$

il est évident qu'asymptotiquement :

$$\mathbf{Y}_f^+ \Xi^T = \Gamma_f \mathbf{X}^+ \Xi^T + \mathbf{H}_f \mathbf{U}_f^+ \Xi^T + \underbrace{\mathbf{N}_f^+ \Xi^T}_{=0} = \Gamma_f \mathbf{X}^+ \Xi^T + \mathbf{H}_f \mathbf{U}_f^+ \Xi^T. \quad (3.178)$$

¹⁷Cette hypothèse n'est pas vérifiée lorsque le bruit de mesure \mathbf{v} , associé à un vecteur d'entrée $\tilde{\mathbf{u}}$ non blanc, est non nul.

La suppression du régime forcé $\mathbf{H}_f \mathbf{U}_f^+$ pondéré par $\mathbf{\Xi}$ est alors réalisable en considérant la projection orthogonale suivante :

$$\mathbf{\Pi}_{(\mathbf{U}_f^+ \mathbf{\Xi}^T)^\perp} \quad (3.179)$$

puisque :

$$\begin{aligned} & \mathbf{Y}_f^+ \mathbf{\Xi}^T \mathbf{\Pi}_{(\mathbf{U}_f^+ \mathbf{\Xi}^T)^\perp} \\ &= \mathbf{\Gamma}_f \mathbf{X}^+ \mathbf{\Xi}^T \mathbf{\Pi}_{(\mathbf{U}_f^+ \mathbf{\Xi}^T)^\perp} + \underbrace{\mathbf{H}_f \mathbf{U}_f^+ \mathbf{\Xi}^T \mathbf{\Pi}_{(\mathbf{U}_f^+ \mathbf{\Xi}^T)^\perp}}_{=0} = \mathbf{\Gamma}_f \mathbf{X}^+ \mathbf{\Xi}^T \mathbf{\Pi}_{(\mathbf{U}_f^+ \mathbf{\Xi}^T)^\perp}. \end{aligned} \quad (3.180)$$

D'un point de vue algébrique, la différence entre les équations (3.170) et (3.176) réside dans une simple inversion des étapes de corrélation et de projection. Cette distinction semble être, à première vue, mineure. Elle est pourtant fondamentale puisqu'elle différencie les algorithmes applicables pour des systèmes à « erreurs en les variables » de ceux qui ne le sont pas [Gus01].

3.5 Conclusion

Une présentation détaillée des algorithmes de la classe d'identification MOESP a été exposée au sein de ce chapitre. Cette description a permis de mettre en évidence les principaux points forts de ces techniques des sous-espaces :

- elles permettent d'estimer directement une forme d'état du système à identifier à partir des données d'entrée-sortie sans utiliser de représentation externe intermédiaire ;
- elles ne nécessitent aucune phase initiale de choix de structure ou de forme canonique ;
- les problèmes d'optimisation non linéaire habituellement rencontrés lors de l'application de méthodes d'erreur de prédiction sont évités puisque les algorithmes présentés utilisent des outils mathématiques robustes tels que la décomposition en valeurs singulières ou la factorisation QR.

Bien que les propriétés asymptotiques (performances statistiques, biais, variance...) des méthodes d'identification des sous-espaces ne soient que partiellement connues et que leur optimalité n'ait pas encore été démontrée [Bau03b], ces techniques peuvent être présentées comme l'une des meilleures alternatives aux algorithmes classiques d'erreur de prédiction. Ce constat, associé au fait que de nombreux systèmes modernes demandent d'avoir accès à des modèles adaptatifs pour leur asservissement, justifie *a priori* le développement d'algorithmes récursifs des sous-espaces. Cette problématique est étudiée depuis la fin des années 80 [MDVV89, VD91]. L'inconvénient majeur des méthodes hors ligne des sous-espaces est la charge calculatoire de la décomposition en valeurs singulières. Il a donc été nécessaire de développer des algorithmes d'identification évitant d'utiliser cet outil mathématique tout en conservant les avantages de l'approche des sous-espaces. Plusieurs solutions performantes ont été développées depuis le milieu des années 90 [Gus96, LV98, OK99, Lov03]. Ces dernières ont comme point commun d'être fondées sur l'adaptation au problème d'identification d'un critère particulier utilisé dans le cadre du traitement d'antennes : le critère de Yang [Yan95b]. Le chapitre 4 a pour objectif de présenter, dans un contexte unifié, ces différents algorithmes

d'identification récursive des sous-espaces. Nous tenterons plus particulièrement de mettre en évidence les avantages et les inconvénients d'une telle approche.

Chapitre 4

Méthodes récursives d'identification des sous-espaces fondées sur le critère de Yang

Sommaire

| | | |
|------------|---|-----------|
| 4.1 | Motivations | 61 |
| 4.2 | Problématique et méthodes du traitement d'antennes | 63 |
| 4.2.1 | Modèle d'antenne et modélisation des signaux | 64 |
| 4.2.2 | MULTiple SIGNAL Classification (MUSIC) | 67 |
| 4.2.3 | Critère de Yang et algorithmes dérivés | 68 |
| 4.2.4 | Méthode du propagateur | 76 |
| 4.2.5 | Conclusion | 78 |
| 4.3 | Adaptation des algorithmes fondés sur le critère de Yang au problème d'identification récursive des sous-espaces | 78 |
| 4.3.1 | Algorithmes développés par T. Gustafsson | 79 |
| 4.3.2 | Algorithmes développés par M. Lovera | 82 |
| 4.3.3 | Adaptation des algorithmes COPAST et COPASTd | 85 |
| 4.3.4 | Algorithmes développés par H. Oku | 88 |
| 4.3.5 | Algorithme développé par M. Lovera (suite) | 90 |
| 4.3.6 | Synthèse | 91 |
| 4.4 | Conclusion | 92 |

4.1 Motivations

Les méthodes d'identification des sous-espaces sont désormais présentées comme une alternative intéressante aux techniques classiques fondées sur l'erreur de prédiction [ZVDL94, BGRS97, AVVD98, BNSR98, FDV⁺98, FDV00, JSL01, DPV⁺02]. Elles permettent, en effet, d'obtenir un modèle d'état du système à l'aide d'outils d'algèbre linéaire robustes. Elles ne nécessitent que peu de paramètres structuraux puisqu'elles fournissent des modèles d'état pleins et emploient des critères de décision pour estimer l'ordre du système au cours du proces-

sus d'identification¹. Toutes ces propriétés ont rapidement amené les chercheurs à tenter de les appliquer à des procédés requérant une modélisation en ligne [MDVV89]. Malheureusement, la plupart des techniques d'identification des sous-espaces hors ligne (4SID) utilisent des outils tels que la décomposition en valeurs propres ou singulières qui ne sont pas appropriés aux applications temps réel de par leur complexité numérique. Or, dans de nombreux cas pratiques, il est nécessaire de mettre à jour le modèle du procédé, à chaque nouvelle acquisition, avec un coût calculatoire faible. Le challenge consiste donc à développer de nouvelles méthodes évitant l'application de ces lourds outils mathématiques. Historiquement, deux approches permettant de contourner cette difficulté peuvent être distinguées :

- Dans un premier temps, des adaptations directes des techniques 4SID ont été proposées [VV91, XK94, CXK94]. Ces ajustements sont principalement basés sur une compression des données d'entrée-sortie auxquelles sont appliquées des méthodes d'extraction fondées sur l'algorithme de Lanczos [Lan50, GV96]. Ces algorithmes ont l'inconvénient de nécessiter des perturbations temporellement et spatialement blanches, ce qui est bien évidemment restrictif en pratique.
- La seconde approche [Gus97a, LGV00, OK02], qui sera détaillée dans la suite de ce chapitre, repose sur l'utilisation d'un critère particulier [Yan95b] : le critère « Projection Approximation Subspace Tracking » (PAST) [Yan95b, Yan96]. Développée à l'origine en traitement d'antennes pour la détection des directions d'arrivée (DDA), cette fonction coût a été adaptée au problème d'identification récursive des sous-espaces par T. Gustafsson [Gus97a]. Il a enrichi les algorithmes initiaux de minimisation en introduisant une variable instrumentale afin de traiter les perturbations, même colorées, pouvant agir sur le procédé. Ces méthodes d'identification ont ensuite été étendues aux schémas d'identification Ordinary, PI et PO MOESP ainsi qu'au traitement des systèmes non linéaires de type Wiener [Lov98, LGV00]. De nouvelles contributions ont été proposées plus récemment [OK02, Lov03]. Celles-ci utilisent directement le critère originel d'ordre 4 de B. Yang [Yan95b].

Les techniques composant cette seconde classe de méthodes d'identification se sont avérées être très efficaces en pratique, en particulier dans le domaine de la surveillance et du suivi de procédé [ONVV01, Lov03, ML03, Oku03, PGMR03, MLL04]. Elles présentent en effet une rapidité calculatoire intéressante pour des applications en ligne. Elles utilisent de plus des outils d'algèbre linéaire robustes et possèdent l'avantage indéniable de traiter le problème des perturbations au cours de la récursion. Ces méthodes reposent avant tout sur l'adaptation, au problème d'identification, d'algorithmes particuliers du traitement d'antennes [Yan95b]. Il est donc essentiel de préciser les fondements de cet axe de recherche ainsi que les algorithmes de détection des directions d'arrivée utiles à la compréhension de cette problématique. Cette étude permettra ainsi de développer l'embase nécessaire à la caractérisation des méthodes de poursuite proposées par B. Yang [Yan95b], techniques qui sont à l'origine des algorithmes d'identification récursive des sous-espaces.

Ce chapitre est composé de deux parties :

- Après une brève description des notations et de la problématique du traitement d'antennes, les principales méthodes d'estimation des directions d'arrivée adaptées en identification récursive des sous-espaces sont exposées au sein de la section 4.2.

¹Les éléments propres utilisés par ces critères ne seront pas accessibles en identification récursive des sous-espaces. L'ordre du système ne pourra donc pas être estimé pendant la récursion.

- Le paragraphe 4.3 permet, quant à lui, de présenter les adaptations des algorithmes de poursuite basés sur le critère de Yang et ses dérivées au problème d'identification en insistant plus particulièrement sur les contributions de chaque auteur.

4.2 Problématique et méthodes du traitement d'antennes

Le traitement d'antennes est une discipline particulière du traitement du signal dont l'objectif majeur est d'extraire les caractéristiques spatio-temporelles de sources émettrices à partir des observations mesurées simultanément en sortie d'un ensemble de capteurs appelé antenne [Jan97] (cf. fig. 4.1). Ces caractéristiques peuvent être très variées (polarisation, fréquence, mo-

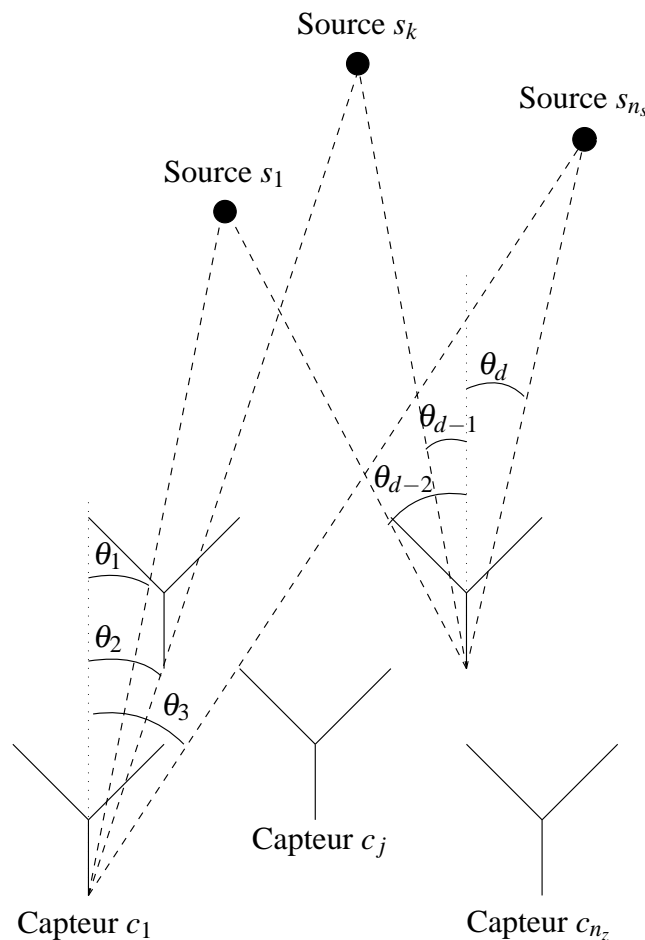


FIG. 4.1: Représentation du problème de traitement d'antennes.

dulation...). L'une des problématiques les plus étudiées consiste à localiser les sources émettrices en détectant les directions d'arrivée des ondes. De nombreux algorithmes ont été développés [JD93, VDS93, KV96, Nai00] et appliqués avec succès dans de nombreux domaines (imagerie médicale, sismologie, communication numérique...) [Far92, FK94, Böh95]. Les méthodes de DDA qui nous intéressent plus particulièrement sont les techniques des sous-espaces. Comme pour l'identification, leur principe repose sur la décomposition de l'espace d'observation en deux sous-espaces orthogonaux : le sous-espace signal et le sous-espace bruit. Ces méthodes ont

connu un essor significatif dans les années 80 avec le développement de l'algorithme MUSIC (Multiple Signal Classification) [BK83]. Nous en présenterons le principe dans la suite de cette section. Mais tout d'abord, afin de poser totalement les fondations de cette étude, il est essentiel d'établir le modèle des observations utilisé dans la plupart des travaux de DDA et retenu dans ce manuscrit pour la présentation des principales techniques de détection.

4.2.1 Modèle d'antenne et modélisation des signaux

L'objectif de ce paragraphe n'est pas de présenter de manière exhaustive la physique de la propagation des ondes dans un milieu tel que l'air mais d'introduire les principales notations et hypothèses nécessaires à la compréhension des méthodes des sous-espaces utilisées en DDA. Considérons donc de nouveau la figure 4.1 et supposons que le signal émis par la source k soit un signal bande étroite de forme :

$$s_k(t) = S_k(t)e^{i\omega_0 t} \quad (4.1)$$

pour lequel S_k est l'amplitude complexe du signal émis et ω_0 sa fréquence centrale. Soit $\mathcal{H}_j(\omega, \theta)$ la réponse du j^{e} capteur à une onde émise depuis la direction θ . En supposant que $\mathcal{H}_j(\omega, \theta) \approx \mathcal{H}_j(\omega_0, \theta)$ dans la bande de fréquence considérée et puisque toute source émet un signal bande étroite², la sortie du j^{e} capteur vérifie :

$$z_j(t) = \mathcal{H}_j(\omega, \theta)s_k(t - \tau_j) + n_j(t) \approx \mathcal{H}_j(\omega_0, \theta)s_k(t)e^{-i\omega_0\tau_j} + n_j(t) \quad (4.2)$$

où τ_j représente le retard de propagation de l'onde relativement à un point de référence fixé *a priori* et n_j le bruit additif contenant les bruits thermiques du capteur, les bruits de mesure, les erreurs de modèle...

Définition 4.1. *Un modèle d'antenne souvent utilisé en DDA est l'antenne linéaire uniforme (cf. fig. 4.2). Cette dernière est composée de capteurs alignés et équidistants. Les sources et l'antenne sont supposées être dans un même plan. Les sources sont également supposées être suffisamment éloignées de l'antenne de telle sorte que les fronts d'ondes incidents sur chaque capteur soient plans. Ce type d'antenne permet d'écrire simplement le retard relatif τ_j du j^{e} capteur, pour une direction d'arrivée θ_k , en fixant comme référence le premier capteur, puisque :*

$$\tau_j(\theta_k) = \frac{(j-1)d \sin(\theta_k)}{v} = (j-1)\tau(\theta_k) \quad (4.3)$$

avec :

$$\tau(\theta_k) = \frac{d \sin(\theta_k)}{v} \quad (4.4)$$

où d est la distance intercapteur et v la célérité de l'onde dans le milieu considéré. Sauf mention contraire, ce modèle d'antenne sera utilisé de manière exclusive dans la suite de ce manuscrit.

Lorsque n_s sources émettent au même instant, la contribution cumulée de ces n_s signaux au niveau du j^{e} capteur s'écrit :

$$z_j(t) = \sum_{\ell=1}^{n_s} \mathcal{H}_j(\omega_0, \theta_\ell) s_\ell(t) e^{-i\omega_0(j-1)\tau(\theta_\ell)} + n_j(t). \quad (4.5)$$

²Cette hypothèse permet d'écrire que $s_k(t - \tau) = s_k(t)e^{-i\omega_0\tau}$.

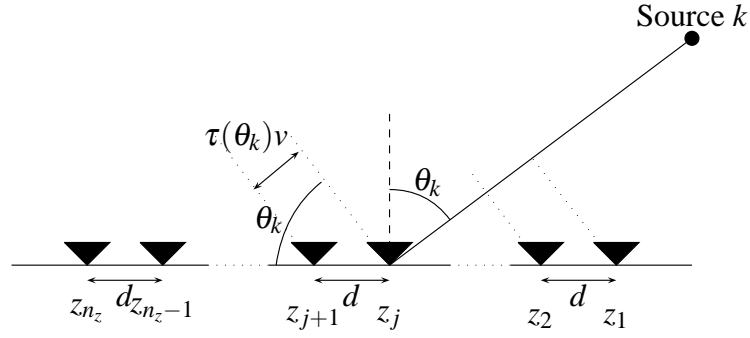


FIG. 4.2: Modèle d'antenne linéaire uniforme.

En supposant que l'antenne considérée soit composée de n_s capteurs ($n_s < n_z$) :

$$\underbrace{\mathbf{z}(t)}_{\in \mathbb{C}^{n_z \times 1}} = \begin{bmatrix} z_1(t) \\ z_2(t) \\ \vdots \\ z_{n_z}(t) \end{bmatrix} = \sum_{\ell=1}^{n_s} \begin{bmatrix} \mathcal{H}_1(\omega_0, \theta_\ell) \\ \mathcal{H}_2(\omega_0, \theta_\ell) e^{-i\omega_0 \tau(\theta_\ell)} \\ \vdots \\ \mathcal{H}_{n_z}(\omega_0, \theta_\ell) e^{-i\omega_0 (n_z-1) \tau(\theta_\ell)} \end{bmatrix} s_\ell(t) + \begin{bmatrix} n_1(t) \\ n_2(t) \\ \vdots \\ n_{n_z}(t) \end{bmatrix} \quad (4.6)$$

$$= \sum_{\ell=1}^{n_s} \boldsymbol{\gamma}(\theta_\ell) s_\ell(t) + \mathbf{n}(t) \quad (4.7)$$

$$= [\boldsymbol{\gamma}(\theta_1) \quad \boldsymbol{\gamma}(\theta_2) \quad \cdots \quad \boldsymbol{\gamma}(\theta_{n_s})] \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_{n_s}(t) \end{bmatrix} + \mathbf{n}(t) \quad (4.8)$$

$$= \underbrace{\boldsymbol{\Gamma}(\boldsymbol{\theta})}_{\in \mathbb{C}^{n_z \times n_s}} \underbrace{\mathbf{s}(t)}_{\in \mathbb{C}^{n_s \times 1}} + \underbrace{\mathbf{n}(t)}_{\in \mathbb{C}^{n_z \times 1}}. \quad (4.9)$$

Le vecteur $\boldsymbol{\gamma}(\theta)$ est nommé vecteur source ou vecteur directionnel. La matrice $\boldsymbol{\Gamma}(\boldsymbol{\theta})$ est quant à elle baptisée matrice de direction. Il est important de remarquer que l'équation (4.9) correspond à l'acquisition simultanée des signaux mesurés en sortie des n_z capteurs à l'instant t . En supposant que nous ayons accès à M instants de mesure, les $M \times n_z$ données peuvent alors se récrire sous la forme compacte suivante :

$$\mathbf{Z}_M = [\mathbf{z}(1) \quad \mathbf{z}(2) \quad \cdots \quad \mathbf{z}(M)] = \boldsymbol{\Gamma}(\boldsymbol{\theta}) \mathbf{S}_M + \mathbf{N}_M \quad (4.10)$$

où \mathbf{S}_M et \mathbf{N}_M sont construits de manière similaire à \mathbf{Z}_M . Le problème d'estimation des directions d'arrivée s'énonce alors comme suit : à partir des données accessibles à la mesure \mathbf{Z}_M , estimer le nombre de sources n_s et les directions d'arrivée $\boldsymbol{\theta}$ des signaux émis en supposant que la matrice de covariance du bruit soit proportionnelle à la matrice identité. Nous nous intéresserons plus précisément à l'extraction des directions d'arrivée $\boldsymbol{\theta}$ sachant qu'une littérature conséquente [Wax91, Fuc92, Wax92, ZW93, XRK94, Yan95a, MV96] traite le problème d'estimation du nombre de sources émettrices et que les méthodes récursives d'identification des sous-espaces s'inspirent exclusivement des techniques de détection des directions d'arrivée. Trois approches peuvent alors être considérées [KV96] :

- Les méthodes de formation de voie [Muc84, VB88, CYH02] reposent sur l'analyse spectrale de la somme pondérée des signaux mesurés en sortie des capteurs de l'antenne, pondération permettant de mettre en évidence une direction donnée par rapport aux autres directions. Ces méthodes peuvent être considérées comme une extension de l'analyse de Fourier au problème spatio-temporel du traitement d'antennes. Elles fournissent malheureusement des modèles non paramétriques et sont fortement dépendantes de la géométrie de l'antenne utilisée [KV96]. Cette limitation a conduit les chercheurs à réfléchir à de nouvelles méthodes d'estimation.
- Les méthodes de maximum de vraisemblance [VS94, KV96] se décomposent en deux familles selon les hypothèses fixées sur la nature des signaux étudiés. Lorsque les signaux reçus sont modélisés par des quantités déterministes, les méthodes de maximum de vraisemblance déterministes sont considérées. Lorsqu'au contraire ces mêmes signaux sont supposés être des processus aléatoires gaussiens, on parle de méthodes de maximum de vraisemblance stochastique. Ces dernières possèdent une précision bien supérieure aux versions déterministes [OVK92] puisque l'estimateur du maximum de vraisemblance stochastique atteint asymptotiquement la borne de Cramér Rao³ [Cra46] lorsque le nombre d'observations accessibles tend vers l'infini. Ces techniques, intéressantes d'un point de vue statistique, présentent néanmoins l'inconvénient d'aboutir à des critères non linéaires de grande dimension nécessitant l'application de méthodes itératives de descente de type Newton [OVSN93].
- Les méthodes des sous-espaces [RPK86, Cad90, ESS94, Rab96, BYR98, Yu00, Uts02] sont quant à elles fondées sur la décomposition de l'espace d'observation en deux sous-espaces complémentaires : le sous-espace signal et le sous-espace bruit. Les premières études faisaient appel, de manière implicite, à la décomposition spectrale de la matrice de covariance $\mathbf{R}_z(t) = \mathbb{E} \{ \mathbf{z}(t) \mathbf{z}^H(t) \}$. C'est à partir du moment où celle-ci a été utilisée explicitement qu'un grand nombre de techniques utilisant ses propriétés ont été développées [Pis73]. Elles ont particulièrement été mises sur le devant de la scène avec l'introduction de l'algorithme MUSIC [BK79, Sch81, BK83, Sch86]. La popularité de ces méthodes est liée au fait qu'elles possèdent des capacités de résolution⁴ bien supérieures aux techniques classiques de formation de voie tout en étant de complexité numérique plus faible que les algorithmes de maximum de vraisemblance.

Dans la suite de ce chapitre, nous nous intéresserons plus précisément aux méthodes des sous-espaces. Elles présentent en effet un bon compromis précision / complexité numérique et sont à l'origine des techniques d'identification récursive des sous-espaces. Un grand nombre de méthodes ont été développées et appliquées avec succès au problème de DDA [Kar86, VO91, VOK91, DeG92, SS92, SES93, ESS94, Rab96, Yu00]. Nous n'en présenterons cependant que trois : l'algorithme MUSIC (MUltiple Signal Classification) [BK79, Sch81], les algorithmes fondés sur le critère de Yang [Yan95b, Yan96] et ceux basés sur le propagateur [MD91]. Ce choix se justifie de la façon suivante :

- MUSIC est l'algorithme qui a permis le développement des méthodes des sous-espaces pour la DDA. Il est fondamental et didactique puisqu'il présente explicitement le principe de détection par décomposition de l'espace d'observation ;

³La borne de Cramér Rao est un minorant de la variance de l'estimateur pour tout estimateur asymptotiquement non biaisé.

⁴Elles sont capables de séparer des sources spatialement proches.

- le critère de Yang est la fonction coût dont découlent toutes les techniques modernes d'identification récursive des sous-espaces développées jusqu'en 2003 [Gus97a, Lov98, LGV00, OK02, Lov03];
 - le propagateur est le point de départ de nouveaux algorithmes récursifs d'identification des sous-espaces [MLV03b, MLV03a, ML03, MLV04, MLL04].
- Ces méthodes sont présentées dans les trois paragraphes suivants.

4.2.2 Multiple Signal Classification (MUSIC)

Considérons l'équation (4.9). La matrice de covariance des sorties de l'antenne vérifie alors :

$$\mathbf{R}_z = \mathbb{E} \{ \mathbf{z}(t) \mathbf{z}^H(t) \} = \underbrace{\mathbf{\Gamma}(\boldsymbol{\theta}) \mathbb{E} \{ \mathbf{s}(t) \mathbf{s}^H(t) \} \mathbf{\Gamma}^H(\boldsymbol{\theta})}_{\mathbf{R}_s} + \sigma^2 \mathbf{I} \quad (4.11)$$

en supposant que le bruit de mesure soit blanc et que ce dernier soit non corrélé aux signaux sources. En lui appliquant une décomposition en valeurs propres :

$$\mathbf{R}_z = \mathbf{U}_s \mathbf{\Lambda}_s \mathbf{U}_s^H + \mathbf{U}_b \mathbf{\Lambda}_b \mathbf{U}_b^H \quad (4.12)$$

où $\mathbf{\Lambda}_s = \text{diag}(\underline{\lambda}_1, \dots, \underline{\lambda}_{\bar{n}_s})$ et $\mathbf{\Lambda}_b = \text{diag}(\underline{\lambda}_{\bar{n}_s+1}, \dots, \underline{\lambda}_{n_z})$ avec \bar{n}_s le rang de \mathbf{R}_s , il est montré que :

$$\text{Im}_{\text{col}}(\mathbf{U}_s) \subseteq \text{Im}_{\text{col}}(\mathbf{\Gamma}(\boldsymbol{\theta})), \quad (4.13)$$

l'égalité ayant lieu lorsque $\bar{n}_s = n_s$. Le sous-espace $\mathbf{U}_s \in \mathbb{C}^{n_z \times \bar{n}_s}$ est nommé sous-espace signal. Son complémentaire $\mathbf{U}_b \in \mathbb{C}^{n_z \times (n_z - \bar{n}_s)}$ est appelé sous-espace bruit. Lorsque $\bar{n}_s = n_s$, l'équation (4.13) permet d'écrire que :

$$\text{Ker}(\mathbf{U}_b^H) = \text{Im}_{\text{col}}(\mathbf{\Gamma}(\boldsymbol{\theta})) \quad (4.14)$$

où $\text{Ker}(\mathbf{U}_b^H)$ est le noyau de \mathbf{U}_b^H . L'équation (4.14) est à la base de l'algorithme MUSIC. En effet, puisque le sous-espace bruit est orthogonal à $\mathbf{\Gamma}(\boldsymbol{\theta})$, l'égalité suivante est assurée :

$$\mathbf{U}_b^H \boldsymbol{\gamma}(\boldsymbol{\theta}) = \mathbf{0} \quad (4.15)$$

pour tout $\boldsymbol{\theta}$ choisi parmi les vraies directions d'arrivée (et uniquement pour ces directions). En supposant qu'un grand nombre de données soient accessibles à la mesure, une estimée de la matrice de covariance \mathbf{R}_z est calculée. Une décomposition en éléments propres est alors appliquée à $\hat{\mathbf{R}}_z$ afin de séparer l'espace d'observation en sous-espace signal et sous espace bruit (cf. equ. (4.12)) et extraire $\hat{\mathbf{U}}_b$. En supposant que les sources émettrices ne soient pas trop proches spatialement ou que les données reçues ne soient pas trop corrélées, le tracé du spectre normalisé suivant :

$$Sp(\boldsymbol{\theta}) = \frac{\boldsymbol{\gamma}(\boldsymbol{\theta})^H \boldsymbol{\gamma}(\boldsymbol{\theta})}{\boldsymbol{\gamma}^H(\boldsymbol{\theta}) \hat{\mathbf{U}}_b \hat{\mathbf{U}}_b^H \boldsymbol{\gamma}(\boldsymbol{\theta})} \quad (4.16)$$

fournit des estimées statistiquement consistantes des angles d'arrivée $\boldsymbol{\theta}$ [KV96]. Ce dernier présente en effet des pics significatifs au voisinage des directions d'émission effectives. Le relevé de ces derniers permet d'estimer les directions d'arrivée.

Plusieurs modifications de l'algorithme MUSIC originel ont été proposées afin d'améliorer ses propriétés statistiques lorsque le rapport signal sur bruit est faible ou lorsque les données sont fortement corrélées. Le lecteur intéressé pourra étudier les références [KT83, KB86, LVT89, SN90].

4.2.3 Critère de Yang et algorithmes dérivés

Dans de nombreux cas pratiques, il est intéressant, voire nécessaire, de pouvoir suivre une ou plusieurs sources se déplaçant. Or, dans ce genre de situation, la division de l'espace d'observation en deux sous-espaces complémentaires via une décomposition en valeurs propres ou singulières n'est pas envisageable de par son coût calculatoire. Pour résoudre cette problématique, plusieurs techniques de poursuite des sous-espaces ont été développées. Leur principe consiste à suivre les éléments propres de la matrice de covariance des observations mise à jour, à chaque acquisition, de la façon suivante :

$$\mathbf{R}_z(t) = \alpha \mathbf{R}_z(t-1) + \beta \mathbf{z}(t) \mathbf{z}^H(t) \quad (4.17)$$

où α et β sont des facteurs de pondération permettant d'influer sur l'importance des nouvelles informations vis-à-vis des anciennes. Les auteurs de l'article [CG90] proposent un classement de ces méthodes de poursuite des sous-espaces en considérant trois familles :

- les méthodes proposant des modifications des algorithmes de décomposition en valeurs propres ou en valeurs singulières afin que ces derniers soient applicables sur certains processus adaptatifs [Fos86, DAD94, DDLY95] ;
- Les techniques fondées sur des algorithmes de mise à jour par addition d'une matrice de rang 1 [Kar86, Sch89, DR90] ;
- celles considérant les décompositions en éléments propres comme un problème d'optimisation, sans ou avec contrainte, minimisable de manière adaptative à l'aide de méthodes itératives ou récursives [Kar84, YSA89, Yan95b].

Les algorithmes fondés sur le critère de Yang [Yan95b, Yan96] font partie de cette troisième famille. Ils reposent plus précisément sur la constatation suivante : sachant que la décomposition en valeurs propres (DVP) de la matrice de covariance \mathbf{R}_z s'écrit :

$$\mathbf{R}_z = \underline{\mathbf{U}}_s \underline{\mathbf{\Lambda}}_s \underline{\mathbf{U}}_s^H + \underline{\mathbf{U}}_b \underline{\mathbf{\Lambda}}_b \underline{\mathbf{U}}_b^H, \quad (4.18)$$

il est aisé de montrer que :

$$\begin{aligned} \underline{\mathbf{U}}_s \underline{\mathbf{U}}_s^H \mathbf{R}_z &= \underline{\mathbf{U}}_s \underline{\mathbf{U}}_s^H (\underline{\mathbf{U}}_s \underline{\mathbf{\Lambda}}_s \underline{\mathbf{U}}_s^H + \underline{\mathbf{U}}_b \underline{\mathbf{\Lambda}}_b \underline{\mathbf{U}}_b^H) \\ &= \underline{\mathbf{U}}_s \underbrace{\underline{\mathbf{U}}_s^H \underline{\mathbf{U}}_s}_{\mathbf{I}_{n_z}} \underline{\mathbf{\Lambda}}_s \underline{\mathbf{U}}_s^H + \underbrace{\underline{\mathbf{U}}_s \underline{\mathbf{U}}_s^H \underline{\mathbf{U}}_b}_{\mathbf{0}} \underline{\mathbf{\Lambda}}_b \underline{\mathbf{U}}_b^H = \underline{\mathbf{U}}_s \underline{\mathbf{\Lambda}}_s \underline{\mathbf{U}}_s^H. \end{aligned} \quad (4.19)$$

Le sous-espace bruit est donc accessible à partir de la relation suivante :

$$\mathbf{R}_z - \underline{\mathbf{U}}_s \underline{\mathbf{U}}_s^H \mathbf{R}_z = \underline{\mathbf{U}}_b \underline{\mathbf{\Lambda}}_b \underline{\mathbf{U}}_b^H. \quad (4.20)$$

De cette observation découle le critère sans contrainte [Yan95b] :

$$V(\mathbf{\Omega}) = \mathbb{E} \|\mathbf{z} - \mathbf{\Omega} \mathbf{\Omega}^H \mathbf{z}\|^2 \quad (4.21)$$

pour lequel $\mathbf{\Omega} \in \mathbb{C}^{n_z \times n_s}$ est supposée de rang plein ($= n_s$). Le théorème suivant est alors vérifié [Yan95b] :

Théorème 4.1. $\mathbf{\Omega}_{opt}$ est le minimum global du critère de Yang $V(\mathbf{\Omega})$ si et seulement si $\mathbf{\Omega}_{opt} = \underline{\mathbf{U}}_s \mathbf{T}$ où $\underline{\mathbf{U}}_s \in \mathbb{C}^{n_z \times n_s}$ contient les n_s vecteurs propres dominants de \mathbf{R}_z et $\mathbf{T} \in \mathbb{C}^{n_s \times n_s}$ une matrice unitaire arbitraire. Tous les autres points stationnaires sont des points selle. Pour chacun d'eux, $V(\mathbf{\Omega})$ vaut la somme des valeurs propres dont les vecteurs propres correspondant ne sont pas inclus dans $\underline{\mathbf{U}}_s$.

Ce théorème permet ainsi de reformuler le problème d'estimation des n_s vecteurs propres dominants d'une matrice symétrique définie positive en un problème d'optimisation sans contrainte. La minimisation du critère $V(\mathbf{\Omega})$ conduit en effet à estimer les n_s vecteurs propres dominants de \mathbf{R}_z , vecteurs propres qui correspondent au sous-espace engendré par les colonnes de $\mathbf{\Gamma}(\boldsymbol{\theta})$. Elle permet donc d'estimer la matrice directionnelle de l'antenne et, par extrapolation, les directions d'arrivée.

Remarque 4.1. *Il est intéressant de noter que la minimisation du critère $V(\mathbf{\Omega})$ fournit une base orthonormale⁵ du sous-espace signal sans nécessiter d'opération de réorthonormalisation.*

Plusieurs algorithmes adaptatifs de minimisation du critère $V(\mathbf{\Omega})$ ont été proposés dans la littérature. Leur objectif est de calculer $\mathbf{\Omega}(t)$ en fonction des estimées accessibles à l'instant $t - 1$ et des nouvelles observations $\mathbf{z}(t)$. Nous en présentons sept⁶ au sein des paragraphes suivants, ce choix se justifiant par le fait que ces derniers sont (ou peuvent être) utilisés en identification récursive des sous-espaces.

4.2.3.1 Méthode de descente du gradient

La première technique consiste à appliquer au critère de Yang $V(\mathbf{\Omega})$ un algorithme de descente de gradient. En effet, puisque⁷ :

$$V(\mathbf{\Omega}) = \mathbb{E} \|\mathbf{z} - \mathbf{\Omega}\mathbf{\Omega}^H \mathbf{z}\|^2 = \text{Tr}(\mathbf{R}_z) - 2\text{Tr}(\mathbf{\Omega}^H \mathbf{R}_z \mathbf{\Omega}) + \text{Tr}(\mathbf{\Omega}^H \mathbf{R}_z \mathbf{\Omega} \mathbf{\Omega}^H \mathbf{\Omega}), \quad (4.22)$$

le gradient de $V(\mathbf{\Omega})$ relatif à $\mathbf{\Omega}$ vaut [Yan95b] :

$$\nabla V = (-2\mathbf{R}_z + \mathbf{R}_z \mathbf{\Omega} \mathbf{\Omega}^H + \mathbf{\Omega} \mathbf{\Omega}^H \mathbf{R}_z) \mathbf{\Omega}. \quad (4.23)$$

La mise à jour du sous-espace signal peut alors être écrite comme suit :

$$\begin{aligned} \mathbf{\Omega}(t) &= \mathbf{\Omega}(t-1) \\ &\quad - \mu (-2\mathbf{R}_z(t) + \mathbf{R}_z(t) \mathbf{\Omega}(t-1) \mathbf{\Omega}^H(t-1) + \mathbf{\Omega}(t-1) \mathbf{\Omega}^H(t-1) \mathbf{R}_z(t)) \mathbf{\Omega}(t-1) \end{aligned} \quad (4.24)$$

où $\mu > 0$ est le pas de descente. Une façon simple de mettre à jour la matrice de covariance \mathbf{R}_z consiste à utiliser la somme pondérée exponentiellement suivante :

$$\mathbf{R}_z(t) = \sum_{k=1}^t \lambda^{t-k} \mathbf{z}(k) \mathbf{z}^H(k) = \lambda \mathbf{R}_z(t-1) + \mathbf{z}(t) \mathbf{z}^H(t) \quad (4.25)$$

où λ est un facteur d'oubli permettant d'atténuer l'effet des données passées. Un tel algorithme, nommé $\text{Yang}_{\text{grad}}$ dans la suite de ce manuscrit, possède une complexité calculatoire en $\mathcal{O}(n_s n_z^2)$.

⁵ $\mathbf{\Omega}$ ne contient pas les vecteurs propres du sous-espace signal mais une base orthonormale arbitraire de ce dernier puisque toute estimation est réalisée à un changement de base \mathbf{T} près.

⁶ $\text{Yang}_{\text{grad}}$, PAST, PASTd, COPAST1, COPAST2, COPASTd et PT.

⁷Rappelons que $\text{Tr}(\mathbf{ABC}) = \text{Tr}(\mathbf{BCA}) = \text{Tr}(\mathbf{CAB})$ [Bro91].

4.2.3.2 Algorithmes PAST et PASTd

Bien que l'algorithme précédent ait montré de bons résultats en simulation [Yan95b], il repose malheureusement sur un critère d'ordre 4 en $\mathbf{\Omega}$ (cf. equ. (4.21)) qui nécessite d'appliquer des techniques de minimisation itératives. Un algorithme récursif a donc été proposé par B. Yang [Yan95b] pour pallier ce défaut. Il considère de nouveau le critère (4.21) et remplace l'espérance mathématique par une somme pondérée :

$$V(\mathbf{\Omega}(t)) = \sum_{k=1}^t \lambda^{t-k} \|\mathbf{z}(k) - \mathbf{\Omega}(t)\mathbf{\Omega}^H(t)\mathbf{z}(k)\|^2. \quad (4.26)$$

La clef de l'algorithme de PAST (Projection Approximation Subspace Tracking) est d'approximer $\mathbf{\Omega}^H(t)\mathbf{z}(k)$ par :

$$\mathbf{h}(k) = \mathbf{\Omega}^H(k-1)\mathbf{z}(k) \quad (4.27)$$

puis d'introduire le critère approximé suivant :

$$\tilde{V}(\mathbf{\Omega}(t)) = \sum_{k=1}^t \lambda^{t-k} \|\mathbf{z}(k) - \mathbf{\Omega}(t)\mathbf{h}(k)\|^2 \quad (4.28)$$

qui est désormais quadratique en $\mathbf{\Omega}$ et minimisé par⁸ :

$$\mathbf{\Omega}(t) = \mathbf{R}_{\mathbf{zh}}(t)\mathbf{R}_{\mathbf{h}}^{-1}(t). \quad (4.29)$$

Les matrices de covariance sont quant à elles mises à jour comme suit :

$$\mathbf{R}_{\mathbf{zh}}(t) = \lambda\mathbf{R}_{\mathbf{zh}}(t-1) + \mathbf{z}(t)\mathbf{h}^H(t) \quad (4.30)$$

$$\mathbf{R}_{\mathbf{h}}(t) = \lambda\mathbf{R}_{\mathbf{h}}(t-1) + \mathbf{h}(t)\mathbf{h}^H(t). \quad (4.31)$$

En appliquant le lemme d'inversion matricielle (cf. equ. (2.25)) à l'équation (4.29), un algorithme de type moindres carrés récursifs, nommé PAST, est alors facilement construit :

$$\mathbf{h}(t) = \mathbf{\Omega}^H(t-1)\mathbf{z}(t) \quad (4.32a)$$

$$\mathbf{K}(t) = \frac{\mathbf{h}^H(t)\mathbf{L}(t-1)}{\lambda + \mathbf{h}^H(t)\mathbf{L}(t-1)\mathbf{h}(t)} \quad (4.32b)$$

$$\mathbf{L}(t) = \frac{1}{\lambda} (\mathbf{L}(t-1) - \mathbf{L}(t-1)\mathbf{h}(t)\mathbf{K}(t)) \quad (4.32c)$$

$$\mathbf{\Omega}(t) = \mathbf{\Omega}(t-1) + (\mathbf{z}(t) - \mathbf{\Omega}(t-1)\mathbf{h}(t))\mathbf{K}(t) \quad (4.32d)$$

$$\text{avec } \mathbf{L}(t) = \mathbf{R}_{\mathbf{h}}^{-1}(t). \quad (4.32e)$$

Un second algorithme de minimisation fondé sur le même principe que PAST est proposé dans [Yan95b]. Baptisé PASTd, il utilise la technique de déflation [YK88, BA92, GV96]. Son principe consiste à estimer de manière séquentielle les vecteurs propres du sous-espace recherché. Cette procédure débute par la recherche du vecteur propre dominant \mathbf{w} qui minimise le critère (4.28) pour $n_s = 1$. Le vecteur des observations \mathbf{z} est alors projeté sur le sous-espace engendré par ce vecteur propre dominant. Le résultat de cette projection est ensuite soustrait au

⁸En supposant que l'inverse existe.

vecteur \mathbf{z} . A l'itération suivante, le deuxième vecteur propre dominant devient le vecteur principal. Le vecteur des observations est alors de nouveau projeté et le résultat soustrait comme précédemment. Ce processus est répété jusqu'à ce que les n_s vecteurs propres soient calculés. L'algorithme résultant est le suivant :

$$\mathbf{z}_1(t) = \mathbf{z}(t)$$

Pour $k = 1 : n_s$

$$h_k(t) = \boldsymbol{\omega}_k^H(t-1)\mathbf{z}_k(t) \quad (4.33a)$$

$$d_k(t) = \lambda d_k(t-1) + |h_k(t)|^2 \quad (4.33b)$$

$$\boldsymbol{\omega}_k(t) = \boldsymbol{\omega}_k(t-1) + (\mathbf{z}_k(t) - \boldsymbol{\omega}_k(t-1)h_k(t)) \frac{h_k^*(t)}{d_k(t)} \quad (4.33c)$$

$$\mathbf{z}_{k+1}(t) = \mathbf{z}_k(t) - \boldsymbol{\omega}_k(t)h_k(t). \quad (4.33d)$$

fin

Cette technique présente l'avantage d'estimer les vecteurs et valeurs propres de la matrice de covariance des observations. En effet, $\boldsymbol{\omega}_k$ est une estimée du k^e vecteur propre de \mathbf{R}_z tandis que d_k est une estimée pondérée de la k^e valeur propre de cette même matrice. De plus, de par son caractère séquentiel, les valeurs propres estimées sont naturellement classées dans un ordre décroissant. Ces valeurs propres peuvent alors être utilisées pour estimer le nombre de sources émettrices [YG94, Yan95a].

Remarque 4.2. *Ces méthodes possèdent des propriétés asymptotiques intéressantes puisque, sous de faibles conditions [Yan96], les estimées obtenues à l'aide de PAST et PASTd convergent vers les vecteurs et valeurs propres désirés avec une probabilité de 1. Elles présentent également une complexité calculatoire réduite en $\mathcal{O}(n_s n_z)$ [Yan95b].*

Remarque 4.3. *Selon B. Yang [Yan95b] et T. Gustafsson [Gus98], la différence entre $\boldsymbol{\Omega}^H(t)\mathbf{z}(k)$ et $\boldsymbol{\Omega}^H(k-1)\mathbf{z}(k)$ est mince, en particulier lorsque k est proche de t . Cette approximation, reprise en identification récursive des sous-espaces [Gus97a, LGV00], pourrait cependant induire quelques difficultés lors de l'estimation d'une réalisation d'état, spécialement lorsque les méthodes qui en découlent sont utilisées pour la détection de changements [LPV01, ONVV01].*

4.2.3.3 Algorithmes COPAST et COPASTd

Les algorithmes exposés précédemment présentent quelques difficultés à estimer, de manière consistante, le sous-espace signal lorsque les sources émettrices sont proches ou lorsque le rapport signal sur bruit est faible [Yan95b]. J. L. Yu [Yu00] propose d'améliorer les performances de PAST et PASTd en introduisant, au sein du critère de Yang, la matrice de covariance des observations en lieu et place du vecteur \mathbf{z} . Il justifie ce choix en affirmant que l'utilisation d'une telle matrice permet d'atteindre un rapport signal sur bruit quadratiquement supérieur à celui obtenu en utilisant uniquement le vecteur des observations. Le critère considéré s'écrit donc :

$$\bar{V}(\boldsymbol{\Omega}) = \|\mathbf{R}_z - \boldsymbol{\Omega}\boldsymbol{\Omega}^H\mathbf{R}_z\|_F^2 \quad (4.34)$$

Le théorème suivant est alors vérifié [Yu00] :

Théorème 4.2. $\mathbf{\Omega}_{opt}$ est le minimum global du critère $\tilde{V}(\mathbf{\Omega})$ si et seulement si $\mathbf{\Omega}_{opt} = \mathbf{U}_s \mathbf{T}$ où $\mathbf{U}_s \in \mathbb{C}^{n_z \times n_s}$ contient les n_s vecteurs propres dominants de \mathbf{R}_z et $\mathbf{T} \in \mathbb{C}^{n_s \times n_s}$ une matrice unitaire arbitraire.

La minimisation de ce critère est réalisable en introduisant l'approximation suivante :

$$\mathbf{\Omega}^H(t) \mathbf{R}_z(t) = \sum_{k=1}^t \lambda^{t-k} \mathbf{\Omega}^H(t) \mathbf{z}(k) \mathbf{z}^H(k) \approx \sum_{k=1}^t \lambda^{t-k} \mathbf{\Omega}^H(k-1) \mathbf{z}(k) \mathbf{z}^H(k) = \mathbf{R}_{hz}(t) \quad (4.35)$$

en posant comme précédemment (cf. equ. (4.27)) :

$$\mathbf{\Omega}^H(k-1) \mathbf{z}(k) = \mathbf{h}(k). \quad (4.36)$$

Le critère approximé vaut alors :

$$\tilde{V}(\mathbf{\Omega}) = \|\mathbf{R}_z - \mathbf{\Omega} \mathbf{R}_{hz}\|_F^2. \quad (4.37)$$

Ce dernier est minimisé par :

$$\mathbf{\Omega}(t) = \mathbf{R}_z(t) \mathbf{R}_{hz}^\dagger(t) \quad (4.38)$$

où \mathbf{R}_{hz}^\dagger symbolise la pseudo inverse de Moore Penrose de \mathbf{R}_{hz} . Puisque cette matrice n'est pas carrée, l'application directe du lemme d'inversion matricielle est impossible. Il est alors nécessaire de faire appel à un algorithme capable de gérer ce type de matrice. En identification récursive, la méthode de la variable instrumentale étendue [Fri84, SS89] permet de fournir une estimée des paramètres du modèle recherché à partir de matrices non carrées (cf. §2.3.2). Le calcul récursif de $\mathbf{\Omega}$ est donc réalisable en adaptant l'algorithme de la variable instrumentale étendue (cf. equ. (2.71)). L'algorithme final s'écrit :

$$\mathbf{h}(t) = \mathbf{\Omega}^H(t-1) \mathbf{z}(t) \quad (4.39a)$$

$$\mathbf{g}(t) = [\mathbf{R}_z(t-1) \mathbf{z}(t) \quad \mathbf{z}(t)] \quad (4.39b)$$

$$\mathbf{\Lambda}(t) = \begin{bmatrix} -\mathbf{z}^T(t) \mathbf{z}(t) & \lambda \\ \lambda & 0 \end{bmatrix} \quad (4.39c)$$

$$\mathbf{\Psi}(t) = [\mathbf{R}_{hz}(t-1) \mathbf{z}(t) \quad \mathbf{h}(t)] \quad (4.39d)$$

$$\mathbf{K}(t) = (\mathbf{\Lambda}(t) + \mathbf{\Psi}^T(t) \mathbf{L}(t-1) \mathbf{\Psi}(t))^{-1} \mathbf{\Psi}^T(t) \mathbf{L}(t-1) \quad (4.39e)$$

$$\mathbf{\Omega}(t) = \mathbf{\Omega}(t-1) + (\mathbf{g}(t) - \mathbf{\Omega}(t-1) \mathbf{\Psi}(t)) \mathbf{K}(t) \quad (4.39f)$$

$$\mathbf{R}_z(t) = \lambda \mathbf{R}_z(t-1) + \mathbf{z}(t) \mathbf{z}^H(t) \quad (4.39g)$$

$$\mathbf{R}_{hz}(t) = \lambda \mathbf{R}_{hz}(t-1) + \mathbf{h}(t) \mathbf{z}^H(t) \quad (4.39h)$$

$$\mathbf{L}(t) = \frac{1}{\lambda^2} (\mathbf{L}(t-1) - \mathbf{L}(t-1) \mathbf{\Psi}(t) \mathbf{K}(t)) \quad (4.39i)$$

$$\text{avec } \mathbf{L}(t) = (\mathbf{R}_{hz}(t) \mathbf{R}_{hz}^H(t))^{-1}. \quad (4.39j)$$

Cet algorithme, nommé COPAST1, présente une complexité en $\mathcal{O}(n_z^2)$.

J. L. Yu propose de réduire la charge numérique de COPAST1 en développant deux algorithmes séquentiels de minimisation du critère (4.34). Il suggère de minimiser successivement n_z critères \bar{V}_i définis par :

$$\bar{V}_i(\mathbf{\Omega}(t)) = \sum_{k=1}^t \lambda^{t-k} \|\mathbf{r}_z^i(k) - \mathbf{\Omega}(t)\mathbf{\Omega}^H(t)\mathbf{r}_z^i(k)\|^2 \quad (4.40)$$

où \mathbf{r}_z^i désigne la i^{e} colonne de \mathbf{R}_z . Puisque chaque fonction \bar{V}_i est analogue au critère (4.26) en remplaçant simplement \mathbf{z} par \mathbf{r}_z^i , deux solutions, inspirées des algorithmes PAST et PASTd, peuvent être envisagées pour les minimiser. En approximant $\mathbf{\Omega}^H(t)\mathbf{r}_z^i(k)$ par $\mathbf{h}^i(k) = \mathbf{\Omega}^H(k-1)\mathbf{r}_z^i(k)$, le critère approximé suivant peut être considéré :

$$\tilde{V}_i(\mathbf{\Omega}(t)) = \sum_{k=1}^t \lambda^{t-k} \|\mathbf{r}_z^i(k) - \mathbf{\Omega}(t)\mathbf{h}^i(k)\|^2. \quad (4.41)$$

Ce dernier est similaire à la fonction coût (4.28). L'algorithme récursif (de type moindres carrés) suivant, nommé COPAST2, peut donc être appliqué pour le minimiser :

$$i = \text{reste}(t, n_z) + 1, \text{ où } \text{reste} \text{ calcule le reste de la division de } t \text{ par } n_z \quad (4.42a)$$

$$\mathbf{R}_z(t) = \lambda \mathbf{R}_z(t-1) + \mathbf{z}(t)\mathbf{z}^T(t) \quad (4.42b)$$

$$\mathbf{r}_z^i(t) = \mathbf{R}_z(:, i)(t) \quad (4.42c)$$

$$\mathbf{h}^i(t) = \mathbf{\Omega}^H(t-1)\mathbf{r}_z^i(t) \quad (4.42d)$$

$$\mathbf{K}(t) = \frac{\mathbf{h}^{iH}(t)\mathbf{L}(t-1)}{\lambda + \mathbf{h}^{iH}(t)\mathbf{L}(t-1)\mathbf{h}^i(t)} \quad (4.42e)$$

$$\mathbf{L}(t) = \frac{1}{\lambda} (\mathbf{L}(t-1) - \mathbf{L}(t-1)\mathbf{h}^i(t)\mathbf{K}(t)) \quad (4.42f)$$

$$\mathbf{\Omega}(t) = \mathbf{\Omega}(t-1) + (\mathbf{r}_z^i(t) - \mathbf{\Omega}(t-1)\mathbf{h}^i(t)) \mathbf{K}(t). \quad (4.42g)$$

Cet algorithme est ensuite aisément complété par une procédure de type déflation (cf. §4.2.3.2) :

$$i = \text{reste}(t, n_z) + 1$$

$$\mathbf{R}_z(t) = \lambda \mathbf{R}_z(t-1) + \mathbf{z}(t)\mathbf{z}^T(t)$$

$$\mathbf{r}_z^i(t) = \mathbf{R}_z(:, i)(t)$$

$$\mathbf{r}_{z,1}^i(t) = \mathbf{r}_z^i(t)$$

Pour $k = 1 : n_s$

$$h_k^i(t) = \mathbf{\omega}_k^H(t-1)\mathbf{r}_{z,k}^i(t) \quad (4.43a)$$

$$d_k(t) = \lambda d_k(t-1) + |h_k^i(t)|^2 \quad (4.43b)$$

$$\mathbf{\omega}_k(t) = \mathbf{\omega}_k(t-1) + \left(\mathbf{r}_{z,k}^i(t) - \mathbf{\omega}_k(t-1)h_k^i(t) \right) \frac{h_k^{i*}(t)}{d_k(t)} \quad (4.43c)$$

$$\mathbf{r}_{z,k+1}^i(t) = \mathbf{r}_{z,k}^i(t) - \mathbf{\omega}_k(t)h_k^i(t). \quad (4.43d)$$

fin

Il sera est nommé COPASTd.

Puisque les matrices et vecteurs traités ont des dimensions équivalentes à celles utilisées par les algorithmes PAST et PASTd et que le critère \bar{V}_i est minimisé de manière séquentielle, la complexité des procédures (4.42) et (4.43) est en $\mathcal{O}(n_s n_z)$.

4.2.3.4 Algorithme du projecteur

Les algorithmes proposés dans les deux paragraphes précédents présentent l'inconvénient d'être fondés sur une approximation (cf. equ. (4.27)). Pour remédier à cette imprécision tout en conservant un critère quadratique, W. Utschick [Uts02] propose d'utiliser une propriété particulière de la matrice $\mathbf{\Omega}_{opt}$ solution du critère $V(\mathbf{\Omega})$:

$$\mathbf{\Omega}_{opt} \mathbf{\Omega}_{opt}^H = \mathbf{\Pi}_{opt} \text{ est un projecteur.} \quad (4.44)$$

En effet, ce dernier vérifie :

$$\text{rang}(\mathbf{\Pi}_{opt}) = n_s, \quad (4.45)$$

$$\mathbf{\Pi}_{opt} = \mathbf{\Pi}_{opt}^H, \quad (4.46)$$

$$\mathbf{\Pi}_{opt}^2 = \mathbf{\Pi}_{opt}. \quad (4.47)$$

Un nouveau problème d'optimisation peut alors être considéré en introduisant la matrice de projection $\mathbf{\Pi} = \mathbf{\Omega} \mathbf{\Omega}^H$ au sein du critère $V(\mathbf{\Omega})$. La fonction coût à minimiser s'écrit alors :

$$\bar{V}(\mathbf{\Pi}) = \mathbb{E} \|\mathbf{z} - \mathbf{\Pi} \mathbf{z}\|^2 = \text{Tr} \left((\mathbf{R}_z - \mathbf{\Pi} \mathbf{R}_z) (\mathbf{R}_z - \mathbf{R}_z \mathbf{\Pi})^T \right) = \text{Tr}(\mathbf{R}_z) - \text{Tr}(\mathbf{\Pi} \mathbf{R}_z) \quad (4.48)$$

en fixant comme contraintes de vérifier, à chaque itération, que :

$$\text{rang}(\mathbf{\Pi}) = n_s \quad (4.49a)$$

$$\mathbf{\Pi} = \mathbf{\Pi}^H \quad (4.49b)$$

$$\mathbf{\Pi}^2 = \mathbf{\Pi}. \quad (4.49c)$$

L'application directe d'une technique de minimisation de type gradient au critère (4.48) n'est pas concevable puisque celle-ci conduit nécessairement à la violation des contraintes (4.49). La solution proposée par W. Utschick [Uts02] consiste à mettre à jour, à chaque nouvelle acquisition, le projecteur $\mathbf{\Pi}$ à l'aide d'un ensemble de rotations de Givens [Giv58, GV96] comme suit :

$$\mathbf{\Pi}(t) = \mathbf{\Psi}(t) \mathbf{\Pi}(t-1) \mathbf{\Psi}^H(t) \quad (4.50)$$

pour laquelle :

$$\mathbf{\Psi}(t) = \mathbf{\Delta}(t) \mathbf{\Theta}(t) \quad (4.51)$$

avec :

$$\mathbf{\Delta}(t) = \mathbf{\Delta}(\boldsymbol{\delta}(t)) = \begin{bmatrix} e^{i\delta_1(t)} & 0 & \dots & 0 \\ 0 & e^{i\delta_2(t)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & e^{i\delta_{n_z}(t)} \end{bmatrix} \in \mathbb{R}^{n_z \times n_z}, \quad (4.52)$$

et :

$$\Theta(t) = \Theta(\vartheta(t)) = \prod_{k=n_z-1}^1 \prod_{\ell=k+1}^{n_z} \Theta^{k,\ell}(\vartheta_{k,\ell}(t)). \quad (4.53)$$

$\Theta^{k,\ell}(\vartheta_{k,\ell}(t))$ est une matrice de rotations de Givens d'indices k et ℓ définie par [GV96] :

$$\Theta^{k,\ell}(\vartheta_{k,\ell}(t)) = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 & \dots & \dots & \dots & 0 & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots & & & & \vdots & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & & & & \vdots & & & & \vdots \\ 0 & \dots & 0 & 1 & 0 & \dots & \dots & \dots & 0 & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & 0 & c & 0 & \dots & 0 & s & 0 & \dots & \dots & 0 \\ \vdots & & & & 0 & 1 & 0 & \dots & 0 & & & & \vdots \\ \vdots & & & & \vdots & \ddots & \ddots & \ddots & \vdots & & & & \vdots \\ \vdots & & & & 0 & \dots & 0 & 1 & 0 & & & & \vdots \\ 0 & \dots & \dots & 0 & -s & 0 & \dots & 0 & c & 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & 0 & \dots & \dots & \dots & 0 & 1 & 0 & \dots & 0 \\ \vdots & & & & \vdots & & & & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & \vdots & & & & \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & \dots & \dots & 0 & \dots & \dots & \dots & 0 & \dots & \dots & 0 & 1 \end{bmatrix} \quad (4.54)$$

$k \qquad \qquad \ell$

telle que $c = \cos(\vartheta_{k,\ell}(t))$ et $s = \sin(\vartheta_{k,\ell}(t))$. La relation (4.50) repose alors sur le théorème suivant [Uts02] :

Théorème 4.3. *Étant donné \mathcal{P} l'espace des matrices de projection $\mathbf{\Pi} \in \mathbb{C}^{n_z \times n_z}$ de rang n_s , il existe $\varepsilon > 0$, $\mathbf{\Pi}_1 \in \mathcal{P}$, $\mathbf{\Pi}_2 \in \mathcal{U}_\varepsilon(\mathbf{\Pi}_1) \subset \mathcal{P}$, δ et ϑ tels que, pour tout $\partial > 0$:*

$$\|\mathbf{\Pi}_2 - \Psi \mathbf{\Pi}_1 \Psi^H\|_2 < \partial \quad (4.55)$$

est vérifié avec :

$$\mathcal{U}_\varepsilon(\mathbf{\Pi}_1) = \{\tilde{\mathbf{\Pi}} \in \mathcal{P} / \|\tilde{\mathbf{\Pi}} - \mathbf{\Pi}_1\|_2 < \varepsilon\}. \quad (4.56)$$

Ce théorème traduit le fait qu'à partir de tout projecteur $\mathbf{\Pi}_1$ de rang n_s , il est possible d'atteindre un autre projecteur $\mathbf{\Pi}_2$, de même rang, dans son voisinage, à l'aide d'un ensemble de transformations algébriques Ψ . En d'autres termes, la poursuite d'un projecteur est réalisable par l'intermédiaire de matrices de projections contenues dans une boule de rayon ε .

La phase suivante consiste alors à estimer les angles de rotation ϑ et δ afin de construire les matrices Θ et Δ . Ces derniers [Uts02] sont déterminés par une loi de type gradient stochastique :

$$\delta_i(t) = -\mu \frac{\partial \bar{V}(\mathbf{\Pi}(t-1))}{\partial \delta_i} \quad (4.57)$$

$$\vartheta_{k,\ell}(t) = -\mu \frac{\partial \bar{V}(\mathbf{\Pi}(t-1))}{\partial \vartheta_{k,\ell}}. \quad (4.58)$$

Un fois ces angles estimés, la mise à jour de $\mathbf{\Omega}$ est réalisée de la façon suivante :

$$\mathbf{\Omega}(t) = \Delta(t) \Theta(t) \mathbf{\Omega}(t-1). \quad (4.59)$$

Un tel algorithme, nommé PT dans la suite de ce manuscrit, présente une complexité calculatoire en $\mathcal{O}(n_z^3)$.

4.2.4 Méthode du propagateur

Les techniques précédentes nous ont permis d'avoir accès à une base de l'espace signal à l'aide d'une fonction coût spécifique : le critère de Yang. La méthode présentée dans ce paragraphe repose sur l'utilisation d'un opérateur particulier, nommé propagateur [MD91], qui permet de décomposer l'espace d'observation en sous-espace signal et sous-espace bruit sans recherche d'éléments propres. Il est qualifié de linéaire puisque seules des opérations linéaires sont appliquées à la matrice interspectrale des signaux reçus. Son application au problème de détection des directions d'arrivée a été plus particulièrement développée par S. Marcos [MMB95, MBS96].

La méthode du propagateur repose sur l'hypothèse suivante : la matrice directionnelle $\mathbf{\Gamma}(\theta) \in \mathbb{C}^{n_z \times n_s}$ est de rang plein pour $n_s < n_z$. Cette matrice possède alors au moins n_s lignes (non nécessairement contiguës) linéairement indépendantes, les $n_z - n_s$ autres pouvant s'exprimer comme une simple combinaison linéaire de ces dernières. Après réorganisation des sorties de capteurs de telle sorte que les n_s premières lignes de $\mathbf{\Gamma}(\theta)$ soient linéairement indépendantes, il est possible de réaliser la répartition suivante :

$$\mathbf{\Gamma} = \begin{bmatrix} \mathbf{\Gamma}_1 \\ \mathbf{\Gamma}_2 \end{bmatrix} \begin{matrix} \} n_s \\ \} n_z - n_s \end{matrix} \quad (4.60)$$

où $\mathbf{\Gamma}_1 \in \mathbb{C}^{n_s \times n_s}$ contient les n_s lignes linéairement indépendantes et $\mathbf{\Gamma}_2 \in \mathbb{C}^{(n_z - n_s) \times n_s}$ les lignes complémentaires.

Définition 4.2. *Le propagateur est l'unique opérateur linéaire $\mathbf{P} \in \mathbb{C}^{n_s \times (n_z - n_s)}$ tel que :*

$$\mathbf{\Gamma}_2 = \mathbf{P}^H \mathbf{\Gamma}_1 \quad (4.61)$$

Il est alors facile de vérifier que :

$$\mathbf{\Gamma} = \begin{bmatrix} \mathbf{\Gamma}_1 \\ \mathbf{\Gamma}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{\Gamma}_1 \\ \mathbf{P}^H \mathbf{\Gamma}_1 \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{I}_{n_s} \\ \mathbf{P}^H \end{bmatrix}}_{\mathbf{E}_s} \mathbf{\Gamma}_1 \quad (4.62)$$

et

$$\mathbf{\Gamma}_2 - \mathbf{P}^H \mathbf{\Gamma}_1 = 0 \Rightarrow \begin{bmatrix} \mathbf{P}^H & -\mathbf{I}_{n_z - n_s} \end{bmatrix} \begin{bmatrix} \mathbf{\Gamma}_1 \\ \mathbf{\Gamma}_2 \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{P}^H & -\mathbf{I}_{n_z - n_s} \end{bmatrix}}_{\mathbf{E}_b} \mathbf{\Gamma} = 0. \quad (4.63)$$

Ainsi :

$$\text{Im}_{\text{col}}(\mathbf{E}_s) = \text{Im}_{\text{col}}(\mathbf{\Gamma}) \quad (4.64)$$

et :

$$\text{Im}_{\text{col}}(\mathbf{E}_b) = \text{Im}_{\text{col}}(\mathbf{\Gamma})^\perp \quad (4.65)$$

puisque \mathbf{E}_s (resp. \mathbf{E}_b) contient \mathbf{I}_{n_s} (resp. $\mathbf{I}_{n_z - n_s}$). Ces deux relations indiquent qu'il est donc possible d'avoir accès aux sous-espaces bruit et signal en estimant le propagateur \mathbf{P} . L'équation

(4.65) est plus particulièrement utilisée en traitement d'antennes puisque, par exemple, le spectre utilisé par les algorithmes MUSIC s'écrit de la manière suivante (cf. §4.2.2) :

$$Sp(\boldsymbol{\theta}) = \frac{\boldsymbol{\gamma}^H(\boldsymbol{\theta})\boldsymbol{\gamma}(\boldsymbol{\theta})}{\boldsymbol{\gamma}^H(\boldsymbol{\theta})\mathbf{E}_b\mathbf{E}_b^H\boldsymbol{\gamma}(\boldsymbol{\theta})}. \quad (4.66)$$

L'objectif ultime est donc d'estimer le propagateur. Plusieurs techniques d'estimation ont été proposées [MJD88, MD91, MMB95, MBS96]. Celles-ci sont principalement basées sur des algorithmes de type moindres carrés (en version adaptative ou non adaptative) ou moindres carrés totaux [VV91]. Deux situations sont rencontrées dans la littérature :

- Une estimation du propagateur à partir d'une décomposition de la matrice de corrélation des données d'entrée :

$$\mathbf{R}_z = \begin{bmatrix} \mathbf{R}_{z_1} & \mathbf{R}_{z_2} \end{bmatrix} \quad (4.67)$$

avec \mathbf{R}_{z_1} et \mathbf{R}_{z_2} de dimension respective $n_z \times n_s$ et $n_z \times n_z - n_s$. Dans le cas non bruité, ces deux sous-matrices vérifient :

$$\mathbf{R}_{z_2} = \mathbf{R}_{z_1}\mathbf{P} \quad (4.68)$$

Dans le cas bruité, une estimation de \mathbf{P} peut être obtenue en minimisant la fonction coût suivante :

$$J(\mathbf{P}) = \|\mathbf{R}_{z_2} - \mathbf{R}_{z_1}\mathbf{P}\|_F^2, \quad (4.69)$$

dont la solution optimale est :

$$\mathbf{P} = (\mathbf{R}_{z_1}^H\mathbf{R}_{z_1})^{-1}\mathbf{R}_{z_1}^H\mathbf{R}_{z_2}. \quad (4.70)$$

- Une estimation à partir de l'ensemble des données accessibles à la mesure en considérant le critère suivant [MMB95] :

$$J(\mathbf{P}) = \|\mathbf{Z}_{M_2} - \mathbf{P}^H\mathbf{Z}_{M_1}\|_F^2 \quad (4.71)$$

avec :

$$\mathbf{Z}_M = [\mathbf{z}(1) \quad \dots \quad \mathbf{z}(M)] = \begin{bmatrix} \mathbf{Z}_{M_1} \\ \mathbf{Z}_{M_2} \end{bmatrix}, \quad (4.72)$$

la solution optimale étant :

$$\mathbf{P} = (\mathbf{Z}_{M_1}\mathbf{Z}_{M_1}^H)^{-1}\mathbf{Z}_{M_1}\mathbf{Z}_{M_2}^H. \quad (4.73)$$

Les solutions proposées pour minimiser (4.69) et (4.71) sont au nombre de trois :

- appliquer les moindres carrés simples en supposant que le bruit soit temporellement et spatialement blanc ;
- utiliser les moindres carrés totaux associés à une décomposition en valeurs singulières permettant ainsi d'atteindre le même espace bruit que celui fourni par MUSIC [MMB95] ;
- considérer le critère (4.69) et mettre à jour le propagateur de manière adaptative en appliquant un algorithme de descente du gradient [MBS96].

Ces trois techniques, efficaces dans la plupart des problèmes de traitement d'antennes, souffrent de plusieurs inconvénients :

- les méthodes utilisant les moindres carrés simples ou l'algorithme de descente du gradient fournissent des estimées asymptotiquement biaisées lorsque le bruit n'est pas temporellement et / ou spatialement blanc ;
- la technique basée sur les moindres carrés totaux produit des estimées non biaisées mais nécessite l'utilisation d'une DVS, outil mathématique robuste mais non utilisable en ligne car trop coûteux en temps de calcul.

Une amélioration de la méthode du propagateur est donc nécessaire si nous désirons l'adapter au problème d'identification récursive des sous-espaces. Cette problématique sera traitée au sein du chapitre 5.

4.2.5 Conclusion

Les paragraphes précédents nous ont permis de présenter deux classes de méthodes de poursuite des directions d'arrivée :

- les algorithmes reposant sur le critère de Yang et ses fonctions coût dérivées ;
- les techniques basées sur l'opérateur propagateur.

Ces méthodes présentent l'avantage d'un coût numérique relativement faible (au maximum en $\mathcal{O}(n_z^3)$) et la capacité de suivre les évolutions spatiales des sources émettrices de par leur caractère adaptatif. Elles sont également, à notre connaissance, les seules techniques de traitement d'antennes utilisées en identification récursive des sous-espaces.

L'objectif de la section suivante est d'exposer un panorama des méthodes d'identification récursive des sous-espaces fondées sur l'adaptation des algorithmes de Yang et de ses dérivées. Une description détaillée de l'utilisation du propagateur en identification sera présentée dans le chapitre 5.

4.3 Adaptation des algorithmes fondés sur le critère de Yang au problème d'identification récursive des sous-espaces

Considérons de nouveau la représentation d'état suivante :

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\tilde{\mathbf{u}}(t) + \mathbf{w}(t) \quad (4.74a)$$

$$\tilde{\mathbf{y}}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\tilde{\mathbf{u}}(t) \quad (4.74b)$$

$$\mathbf{u}(t) = \tilde{\mathbf{u}}(t) + \mathbf{v}(t) \quad (4.74c)$$

$$\mathbf{y}(t) = \tilde{\mathbf{y}}(t) + \mathbf{v}(t). \quad (4.74d)$$

Les méthodes d'identification récursive des sous-espaces ont pour but d'estimer, à chaque nouvelle acquisition de données d'entrée-sortie \mathbf{u} et \mathbf{y} , les matrices d'état du système connaissant leurs estimées à l'instant précédent. Bien que les algorithmes hors ligne des sous-espaces aient démontré leur efficacité dans de nombreuses situations pratiques [BGRS97, BNSR98, FDV00], ceux-ci sont difficilement applicables en temps réel de par leur coût calculatoire. Le principal obstacle se trouve être la décomposition en valeurs singulières nécessaire à l'estimation de l'espace engendré par les colonnes de la matrice d'observabilité. Il a donc été impératif de développer de nouvelles techniques évitant de recourir à cet outil numériquement lourd. Plusieurs solutions ont été proposées [VD91, CK95, Gus97a, LGV00, OK02, Lov03]. Leur principe réside principalement en l'adaptation, au problème d'identification, de certaines méthodes

de poursuite des sous-espaces utilisées en traitement d'antennes. Les plus efficaces reposent sur l'utilisation du critère de Yang présenté §4.2.3. L'objectif des paragraphes suivants est de présenter les principales méthodes d'identification adaptant les algorithmes de Yang et ses dérivées au problème d'estimation récursive du sous-espace d'observabilité. Nous tenterons de mettre en évidence la manière dont chaque auteur ajuste ces méthodes de traitement d'antennes à la problématique qu'il se pose, *i.e.* traiter les perturbations agissant sur le procédé, diminuer le coût calculatoire des algorithmes, augmenter la fiabilité numérique des techniques d'estimation... Nous proposerons également une adaptation des méthodes COPAST et COPASTd à notre problème d'estimation récursive.

4.3.1 Algorithmes développés par T. Gustafsson [Gus97a]

Bien que des méthodes d'identification récursive des sous-espaces aient été développées avant 1997 [VD91, CK95], une étape capitale fut franchie lorsque T. Gustafsson proposa un algorithme de mise à jour de l'espace d'observabilité à partir de données perturbées par des bruits colorés pouvant être mutuellement corrélés. En effet, les techniques suggérées jusqu'alors supposaient que le bruit agissant sur le procédé était temporellement et spatialement blanc, ce qui, pour des systèmes ne vérifiant pas cette hypothèse, conduisait nécessairement à des estimées asymptotiquement biaisées. L'algorithme de T. Gustafsson repose sur une amélioration de l'algorithme PAST (cf. §4.2.3.2). Il propose d'associer au critère PAST (cf. equ. (4.28)) une variable instrumentale afin de le rendre applicable lorsque le bruit n'est pas blanc. Considérons donc de nouveau le modèle d'antenne linéaire uniforme suivant :

$$\mathbf{z}(t) = \mathbf{\Gamma}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t) \quad (4.75)$$

et introduisons une variable instrumentale $\boldsymbol{\xi} \in \mathbb{C}^{n_\xi \times 1}$, $n_\xi \geq n_s$, telle que [Gus98] :

$$\mathbb{E} \left\{ \mathbf{n}(t) \boldsymbol{\xi}^H(t) \right\} = 0 \quad (4.76)$$

$$\text{rang} \left(\mathbb{E} \left\{ \mathbf{s}(t) \boldsymbol{\xi}^H(t) \right\} \right) = n_s. \quad (4.77)$$

La matrice de corrélation $\mathbf{R}_{z\xi}$ vérifie alors :

$$\mathbf{R}_{z\xi} = \mathbf{\Gamma}(\boldsymbol{\theta})\mathbf{R}_{s\xi}. \quad (4.78)$$

L'estimation de la matrice directionnelle $\mathbf{\Gamma}(\boldsymbol{\theta})$ est réalisable en considérant le critère suivant :

$$V_{IV}(\boldsymbol{\Omega}(t)) = \left\| \mathbf{R}_{z\xi}(t) - \boldsymbol{\Omega}(t)\boldsymbol{\Omega}^H(t)\mathbf{R}_{z\xi}(t) \right\|_F^2. \quad (4.79)$$

En effet [Gus98] :

Théorème 4.4. *Le minimum global de $V_{IV}(\boldsymbol{\Omega}(t))$ est atteint si et seulement si $\boldsymbol{\Omega}(t) = \bar{\mathbf{U}}_s \mathbf{T}$ où $\bar{\mathbf{U}}_s$ contient les n_s vecteurs singuliers dominants de $\mathbf{R}_{z\xi}$ et \mathbf{T} est une matrice unitaire arbitraire.*

En appliquant l'approximation de Yang (cf. equ. (4.27)) :

$$\boldsymbol{\Omega}^H(t)\mathbf{R}_{z\xi}(t) = \sum_{k=1}^t \lambda^{t-k} \boldsymbol{\Omega}^H(t) \mathbf{z}(k) \boldsymbol{\xi}^H(k) \approx \sum_{k=1}^t \lambda^{t-k} \underbrace{\boldsymbol{\Omega}^H(k-1) \mathbf{z}(k)}_{\mathbf{h}(k)} \boldsymbol{\xi}^H(k) = \mathbf{R}_{\mathbf{h}\xi}(t), \quad (4.80)$$

le critère (4.79) s'écrit :

$$\tilde{V}_{IV}(\mathbf{\Omega}(t)) = \left\| \mathbf{R}_{z\xi}(t) - \mathbf{\Omega}(t)\mathbf{R}_{h\xi}(t) \right\|_F^2. \quad (4.81)$$

Son minimum est atteint pour :

$$\mathbf{\Omega}(t) = \mathbf{R}_{z\xi}(t)\mathbf{R}_{h\xi}^\dagger(t). \quad (4.82)$$

Deux algorithmes récursifs peuvent alors être construits selon la valeur de n_ξ .

- Lorsque $n_\xi = n_s$, puisque $\mathbf{R}_{h\xi}$ est une matrice carrée, en supposant la matrice inverse existe, l'application du lemme d'inversion matricielle à l'équation (4.82) fournit un algorithme récursif, de complexité en $\mathcal{O}(n_s n_z)$, nommé IVPAST [Gus98], très proche de la méthode classique de la variable instrumentale récursive [LS83] :

$$\mathbf{h}(t) = \mathbf{\Omega}^H(t-1)\mathbf{z}(t) \quad (4.83a)$$

$$\mathbf{K}(t) = \frac{\boldsymbol{\xi}^H(t)\mathbf{L}(t-1)}{\lambda + \boldsymbol{\xi}^H(t)\mathbf{L}(t-1)\mathbf{h}(t)} \quad (4.83b)$$

$$\mathbf{L}(t) = \frac{1}{\lambda} (\mathbf{L}(t-1) - \mathbf{L}(t-1)\mathbf{h}(t)\mathbf{K}(t)) \quad (4.83c)$$

$$\mathbf{\Omega}(t) = \mathbf{\Omega}(t-1) + (\mathbf{z}(t) - \mathbf{\Omega}(t-1)\mathbf{h}(t))\mathbf{K}(t). \quad (4.83d)$$

- Lorsque $n_\xi > n_s$, l'application directe du lemme d'inversion matricielle n'est plus envisageable. Il est alors nécessaire de faire appel à l'algorithme de la variable instrumentale étendue [Fri84, SS89] (cf. §2.3.2). L'application d'un tel algorithme à l'équation (4.82) conduit à la formulation récursive suivante de complexité $\mathcal{O}(n_z n_\xi)$:

$$\mathbf{h}(t) = \mathbf{\Omega}^H(t-1)\mathbf{z}(t) \quad (4.84a)$$

$$\mathbf{g}(t) = [\mathbf{R}_{z\xi}(t-1)\boldsymbol{\xi}(t) \quad \mathbf{z}(t)] \quad (4.84b)$$

$$\mathbf{\Lambda}(t) = \begin{bmatrix} -\boldsymbol{\xi}^H(t)\boldsymbol{\xi}(t) & \lambda \\ \lambda & 0 \end{bmatrix} \quad (4.84c)$$

$$\mathbf{\Psi}(t) = [\mathbf{R}_{h\xi}(t-1)\boldsymbol{\xi}(t) \quad \mathbf{h}(t)] \quad (4.84d)$$

$$\mathbf{K}(t) = (\mathbf{\Lambda}(t) + \mathbf{\Psi}^H(t)\mathbf{L}(t-1)\mathbf{\Psi}(t))^{-1}\mathbf{\Psi}^H(t)\mathbf{L}(t-1) \quad (4.84e)$$

$$\mathbf{\Omega}(t) = \mathbf{\Omega}(t-1) + (\mathbf{g}(t) - \mathbf{\Omega}(t-1)\mathbf{\Psi}(t))\mathbf{K}(t) \quad (4.84f)$$

$$\mathbf{R}_{z\xi}(t) = \lambda\mathbf{R}_{z\xi}(t-1) + \mathbf{z}(t)\boldsymbol{\xi}^H(t) \quad (4.84g)$$

$$\mathbf{R}_{h\xi}(t) = \lambda\mathbf{R}_{h\xi}(t-1) + \mathbf{h}(t)\boldsymbol{\xi}^H(t) \quad (4.84h)$$

$$\mathbf{L}(t) = \frac{1}{\lambda^2} (\mathbf{L}(t-1) - \mathbf{L}(t-1)\mathbf{\Psi}(t)\mathbf{K}(t)) \quad (4.84i)$$

$$\text{avec } \mathbf{L}(t) = \left(\mathbf{R}_{h\xi}(t)\mathbf{R}_{h\xi}^H(t) \right)^{-1}. \quad (4.84j)$$

Cet algorithme est nommé EIVPAST.

Bien que l'algorithme précédent présente des propriétés intéressantes pour l'identification des systèmes⁹, il n'est pas directement applicable aux données d'entrée-sortie mesurées sur le

⁹Rappelons qu'il cherche à supprimer l'effet des perturbations.

procédé à identifier. En effet, son utilisation pour l'estimation d'une base de l'espace d'observabilité nécessite de transformer le système d'état (4.74) afin que ce dernier ressemble¹⁰ au modèle de traitement d'antennes (4.75). T. Gustafsson réalise cette transformation en introduisant le vecteur de sorties décalées \mathbf{y}_f^+ que nous savons vérifier (cf. equ. (3.27)) :

$$\mathbf{y}_f^+(t) = \mathbf{\Gamma}_f \mathbf{x}(t) + \mathbf{H}_f \mathbf{u}_f^+(t) + \mathbf{n}_f^+(t). \quad (4.85)$$

En considérant le signal de sortie modifié suivant :

$$\mathbf{z}_f^+(t) = \mathbf{y}_f^+ - \mathbf{H}_f \mathbf{u}_f^+(t), \quad (4.86)$$

l'équation (4.85) devient :

$$\mathbf{z}_f^+(t) = \mathbf{\Gamma}_f \mathbf{x}(t) + \mathbf{n}_f^+(t), \quad (4.87)$$

qui n'est autre qu'une expression image du modèle d'antenne (4.75). L'estimation d'une base de l'espace d'observabilité semble donc réalisable à l'aide des algorithmes IVPAST et EIVPAST si nous parvenons à calculer, à chaque instant, le vecteur de sortie modifié \mathbf{z}_f^+ . Par définition, ce vecteur dépend de la matrice de Toeplitz \mathbf{H}_f . Or, à l'instant t , aucune estimée de $\mathbf{H}_f(t)$ n'est accessible. T. Gustafsson propose alors d'approximer $\mathbf{H}_f(t)$ par $\hat{\mathbf{H}}_f(t-1)$ qui est, quant à elle, calculable à partir des estimées des matrices d'état de l'instant $t-1$. Il simplifie sa construction en remarquant que le premier bloc colonne de \mathbf{H}_f vaut :

$$\begin{bmatrix} \mathbf{D} \\ \mathbf{\Gamma}_f(1 : n_y(f-1), :) \mathbf{B} \end{bmatrix}, \quad (4.88)$$

les autres blocs colonnes pouvant s'exprimer comme une troncature de ce dernier. Il est alors aisé de construire $\hat{\mathbf{H}}_f(t-1)$ en remplaçant, dans l'équation précédente, les matrices considérées par leur estimée. Cette technique permet ainsi d'éviter de calculer $\hat{\mathbf{A}}^k$. L'équation suivante est alors vérifiée :

$$\check{\mathbf{z}}_f^+(t) = \mathbf{y}_f^+ - \hat{\mathbf{H}}_f(t-1) \mathbf{u}_f^+(t) \approx \mathbf{\Gamma}_f \mathbf{x}(t) + \mathbf{n}_f^+(t). \quad (4.89)$$

L'algorithme (4.84) peut alors être appliqué au vecteur de sortie modifié $\check{\mathbf{z}}_f^+$.

Remarque 4.4. *Lorsqu'on utilise le critère PAST ou ses dérivées, le sous-espace signal est estimé en minimisant la fonction coût modifiée $\tilde{V}(\mathbf{\Omega})$ au lieu de $V(\mathbf{\Omega})$ (cf. equ. (4.21) et (4.28)). Le sous-espace colonne estimé est alors légèrement différent de celui obtenu avec la fonction originelle. Théoriquement, les colonnes de $\mathbf{\Omega}$ sont orthonormales. Bien que cette propriété ne soit pas nécessaire à l'extraction des matrices d'état, la minimisation de $\tilde{V}(\mathbf{\Omega})$ conduit à une matrice dont les colonnes ne sont pas orthonormales. Cette caractéristique évolue au cours de la minimisation puisque, sous certaines conditions, le minimiseur de $\tilde{V}(\mathbf{\Omega})$ converge vers une matrice orthonormale [Yan96] avec une probabilité de 1. Cette évolution peut être interprétée comme un lent changement de base. Il est alors impossible de garantir que $\mathbf{\Gamma}_f(t)$ et $\mathbf{\Gamma}_f(t-1)$ s'expriment dans le même repère d'état. Cette variation pourrait poser problème lors de l'estimation de la réalisation d'état, en particulier lorsque cette technique d'estimation est utilisée pour détecter des changements [LPV01, ONVV01].*

¹⁰Une description détaillée de l'analogie traitement d'antennes / identification des sous-espaces est exposée dans le chapitre 5.

4.3.2 Algorithmes développés par M. Lovera [Lov98, LGV00]

Le principal avantage de l'algorithme précédent est son faible coût calculatoire. Malheureusement, la fiabilité numérique de l'étape d'estimation du vecteur de sortie modifié (4.89) peut être remise en question puisque celle-ci est fondée sur un ensemble d'approximations. M. Lovera propose d'améliorer l'algorithme de T. Gustafsson en développant des versions récursives des décompositions QR des schémas d'identification `Ordinary` et `PI/PO MOESP`. Il suggère de mettre à jour, à chaque nouvelle acquisition, les matrices $\bar{\mathbf{R}}_{22}$ et $\bar{\mathbf{R}}_{32}$ des factorisations suivantes :

$$\begin{bmatrix} \mathbf{U}_f^+ \\ \mathbf{Y}_f^+ \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_{11} & \mathbf{0} \\ \bar{\mathbf{R}}_{21} & \bar{\mathbf{R}}_{22} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1 \\ \bar{\mathbf{Q}}_2 \end{bmatrix} \quad (4.90)$$

$$\begin{bmatrix} \mathbf{U}_f^+ \\ \bar{\boldsymbol{\Xi}} \\ \mathbf{Y}_f^+ \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_{11} & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{21} & \bar{\mathbf{R}}_{22} & \mathbf{0} \\ \bar{\mathbf{R}}_{31} & \bar{\mathbf{R}}_{32} & \bar{\mathbf{R}}_{33} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1 \\ \bar{\mathbf{Q}}_2 \\ \bar{\mathbf{Q}}_3 \end{bmatrix} \quad (4.91)$$

sachant que ces dernières constituent, en identification hors ligne, le point de départ d'une estimation consistante de la matrice d'observabilité du système (cf. §3.3.1 et 3.3.2).

4.3.2.1 Mise à jour de la factorisation QR du schéma d'identification `Ordinary MOESP`

Considérons la décomposition QR suivante accessible à l'instant $\bar{t} - 1 = t + M - 2$:

$$\begin{bmatrix} \mathbf{U}_f^+(\bar{t}-1) \\ \mathbf{Y}_f^+(\bar{t}-1) \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_{11}(\bar{t}-1) & \mathbf{0} \\ \bar{\mathbf{R}}_{21}(\bar{t}-1) & \bar{\mathbf{R}}_{22}(\bar{t}-1) \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1(\bar{t}-1) \\ \bar{\mathbf{Q}}_2(\bar{t}-1) \end{bmatrix} \quad (4.92)$$

utilisée dans le schéma d'identification `Ordinary MOESP` (cf. §3.3.1). Lorsqu'un nouveau couple de données est mesuré, la décomposition précédente est mise à jour de la façon suivante¹¹ :

$$\begin{bmatrix} \sqrt{\lambda} \begin{bmatrix} \bar{\mathbf{R}}_{11}(\bar{t}-1) & \mathbf{0} \\ \bar{\mathbf{R}}_{21}(\bar{t}-1) & \bar{\mathbf{R}}_{22}(\bar{t}-1) \end{bmatrix} \begin{bmatrix} \mathbf{u}_f^+(\bar{t}) \\ \mathbf{y}_f^+(\bar{t}) \end{bmatrix} \\ \begin{bmatrix} \bar{\mathbf{Q}}_1(\bar{t}-1) & \mathbf{0} \\ \bar{\mathbf{Q}}_2(\bar{t}-1) & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \end{bmatrix}. \quad (4.93)$$

Une séquence de rotations de Givens [Giv58, GV96] peut alors être utilisée pour remodeler la matrice $\bar{\mathbf{R}}$ de telle sorte qu'elle soit à nouveau triangulaire par bloc et ainsi annihiler le vecteur $\mathbf{u}_f^+(\bar{t})$:

$$\begin{bmatrix} \sqrt{\lambda} \begin{bmatrix} \bar{\mathbf{R}}_{11}(\bar{t}-1) & \mathbf{0} \\ \bar{\mathbf{R}}_{21}(\bar{t}-1) & \bar{\mathbf{R}}_{22}(\bar{t}-1) \end{bmatrix} \begin{bmatrix} \mathbf{u}_f^+(\bar{t}) \\ \mathbf{y}_f^+(\bar{t}) \end{bmatrix} \\ \mathbf{Rot}_G(\bar{t}) \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_{11}(\bar{t}) & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{21}(\bar{t}) & \sqrt{\lambda} \bar{\mathbf{R}}_{22}(\bar{t}-1) & \bar{\mathbf{z}}_f^+(\bar{t}) \end{bmatrix}. \quad (4.94)$$

$\mathbf{Rot}_G(\bar{t})$ est la matrice résultante des produits de matrices de rotations de Givens nécessaires à l'annulation du terme $\mathbf{u}_f^+(\bar{t})$. $\bar{\mathbf{z}}_f^+(\bar{t})$ est quant à lui le vecteur obtenu en modifiant $\mathbf{y}_f^+(\bar{t})$ de façon à intégrer l'information contenue dans $\mathbf{u}_f^+(\bar{t})$ et le bloc $[\bar{\mathbf{R}}_{11}^T(\bar{t}-1) \bar{\mathbf{R}}_{21}^T(\bar{t}-1)]^T$. La relation suivante est alors vérifiée [Lov98] :

$$\bar{\mathbf{R}}_{22}(\bar{t}) \bar{\mathbf{R}}_{22}^T(\bar{t}) = \lambda \bar{\mathbf{R}}_{22}(\bar{t}-1) \bar{\mathbf{R}}_{22}^T(\bar{t}-1) + \bar{\mathbf{z}}_f^+(\bar{t}) \bar{\mathbf{z}}_f^{+T}(\bar{t}). \quad (4.95)$$

¹¹Remarquez l'introduction du facteur d'oubli permettant de pondérer les informations passées.

De plus (cf. preuve en annexe A.3), le vecteur de sortie modifié $\bar{\mathbf{z}}_f^+$ est lié à la matrice d'observabilité par la relation suivante [GLV98] :

$$\begin{aligned} \bar{\mathbf{z}}_f^+(\bar{t}) = & \mathbf{\Gamma}_f \left(\sqrt{\lambda} \mathbf{X}^+(\bar{t}-1) \bar{\mathbf{Q}}_1^T(\bar{t}-1) \mathbf{Rot}_G^{12}(\bar{t}) + \mathbf{x}(\bar{t}-1) \mathbf{Rot}_G^{22}(\bar{t}) \right) \\ & + \sqrt{\lambda} \mathbf{N}_f^+(\bar{t}-1) \bar{\mathbf{Q}}_1^T(\bar{t}-1) \mathbf{Rot}_G^{12}(\bar{t}) + \mathbf{n}_f^+(\bar{t}) \mathbf{Rot}_G^{22}(\bar{t}) \end{aligned} \quad (4.96)$$

en posant :

$$\mathbf{Rot}_G(\bar{t}) = \begin{bmatrix} \mathbf{Rot}_G^{11}(\bar{t}) & \mathbf{0} & \mathbf{Rot}_G^{12}(\bar{t}) \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{Rot}_G^{21}(\bar{t}) & \mathbf{0} & \mathbf{Rot}_G^{22}(\bar{t}) \end{bmatrix}. \quad (4.97)$$

Le sous-espace d'observabilité est donc estimable à partir du vecteur de sortie modifié obtenu par mise à jour de la factorisation QR (4.93). Malheureusement, quelle que soit la nature des perturbations agissant sur le procédé, le terme $\mathbf{N}_f^+(\bar{t}-1) \bar{\mathbf{Q}}_1^T(\bar{t}-1) \mathbf{Rot}_G^{12}(\bar{t}) + \mathbf{n}_f^+(\bar{t}) \mathbf{Rot}_G^{22}(\bar{t})$ relatif au bruit n'est pas blanc. M. Lovera propose donc d'introduire ce vecteur de sortie modifié au sein de l'algorithme EIVPAST afin d'estimer, de manière consistante, une base de $\mathbf{\Gamma}_f$. Ce dernier contient en effet une variable instrumentale capable d'atténuer l'effet des perturbations contenues dans $\bar{\mathbf{z}}_f^+$.

4.3.2.2 Mise à jour de la factorisation QR des schémas d'identification PI et PO MOESP

Supposons que la factorisation du bloc matriciel $\left[\mathbf{U}_f^{+T} \quad \mathbf{\Xi}^T \quad \mathbf{Y}_f^{+T} \right]^T$ à l'instant $\bar{t}-1$ soit donnée par :

$$\begin{bmatrix} \mathbf{U}_f^+(\bar{t}-1) \\ \mathbf{\Xi}(\bar{t}-1) \\ \mathbf{Y}_f^+(\bar{t}-1) \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_{11}(\bar{t}-1) & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{21}(\bar{t}-1) & \bar{\mathbf{R}}_{22}(\bar{t}-1) & \mathbf{0} \\ \bar{\mathbf{R}}_{31}(\bar{t}-1) & \bar{\mathbf{R}}_{32}(\bar{t}-1) & \bar{\mathbf{R}}_{33}(\bar{t}-1) \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1(\bar{t}-1) \\ \bar{\mathbf{Q}}_2(\bar{t}-1) \\ \bar{\mathbf{Q}}_3(\bar{t}-1) \end{bmatrix}. \quad (4.98)$$

Lorsqu'un nouveau couple de données est mesuré, la décomposition précédente peut alors être mise à jour comme suit :

$$\begin{bmatrix} \sqrt{\lambda} \begin{bmatrix} \bar{\mathbf{R}}_{11}(\bar{t}-1) & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{21}(\bar{t}-1) & \bar{\mathbf{R}}_{22}(\bar{t}-1) & \mathbf{0} \\ \bar{\mathbf{R}}_{31}(\bar{t}-1) & \bar{\mathbf{R}}_{32}(\bar{t}-1) & \bar{\mathbf{R}}_{33}(\bar{t}-1) \end{bmatrix} \begin{bmatrix} \mathbf{u}_f^+(\bar{t}) \\ \boldsymbol{\xi}(\bar{t}) \\ \mathbf{y}_f^+(\bar{t}) \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1(\bar{t}-1) & \mathbf{0} \\ \bar{\mathbf{Q}}_2(\bar{t}-1) & \mathbf{0} \\ \bar{\mathbf{Q}}_3(\bar{t}-1) & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \end{bmatrix} \quad (4.99)$$

avec $\boldsymbol{\xi}(\bar{t}) = \mathbf{u}_p^-(\bar{t})$ pour la classe PI MOESP et $\boldsymbol{\xi}(\bar{t}) = \left[\mathbf{u}_p^{-T}(\bar{t}) \quad \mathbf{y}_p^{-T}(\bar{t}) \right]^T$ pour PO MOESP. Puisque, dans le cas hors ligne (cf. §3.3.2) :

$$\text{Im}_{\text{col}}(\bar{\mathbf{R}}_{32}) = \text{Im}_{\text{col}}(\mathbf{\Gamma}_f), \quad (4.100)$$

il semble intéressant de mettre à jour $\bar{\mathbf{R}}_{32}$ afin d'accéder au vecteur de sortie modifié de l'instant \bar{t} . Pour cela, appliquons un ensemble de rotations de Givens à (4.99) afin de retriangulariser ce bloc matriciel. Ainsi, considérons, dans un premier temps, un ensemble de rotations de Givens,

nommé \mathbf{Rot}_{G_1} , permettant d'annuler $\mathbf{u}_f^+(\bar{t})$:

$$\begin{aligned} & \begin{bmatrix} \sqrt{\lambda} \begin{bmatrix} \bar{\mathbf{R}}_{11}(\bar{t}-1) & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{21}(\bar{t}-1) & \bar{\mathbf{R}}_{22}(\bar{t}-1) & \mathbf{0} \\ \bar{\mathbf{R}}_{31}(\bar{t}-1) & \bar{\mathbf{R}}_{32}(\bar{t}-1) & \bar{\mathbf{R}}_{33}(\bar{t}-1) \end{bmatrix} \begin{bmatrix} \mathbf{u}_f^+(\bar{t}) \\ \boldsymbol{\xi}(\bar{t}) \\ \mathbf{y}_f^+(\bar{t}) \end{bmatrix} \end{bmatrix} \mathbf{Rot}_{G_1}(\bar{t}) \\ &= \begin{bmatrix} \bar{\mathbf{R}}_{11}(\bar{t}) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{21}(\bar{t}) & \sqrt{\lambda} \begin{bmatrix} \bar{\mathbf{R}}_{22}(\bar{t}-1) & \mathbf{0} \\ \bar{\mathbf{R}}_{32}(\bar{t}-1) & \bar{\mathbf{R}}_{33}(\bar{t}-1) \end{bmatrix} & \bar{\mathbf{z}}_p^-(\bar{t}) \\ \bar{\mathbf{R}}_{31}(\bar{t}) & & \bar{\mathbf{z}}_f^+(\bar{t}) \end{bmatrix}. \end{aligned} \quad (4.101)$$

De même, appliquons un autre ensemble de rotations \mathbf{Rot}_{G_2} permettant d'annihiler $\bar{\mathbf{z}}_p^-(\bar{t})$:

$$\begin{aligned} & \begin{bmatrix} \bar{\mathbf{R}}_{11}(\bar{t}) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{21}(\bar{t}) & \sqrt{\lambda} \begin{bmatrix} \bar{\mathbf{R}}_{22}(\bar{t}-1) & \mathbf{0} \\ \bar{\mathbf{R}}_{32}(\bar{t}-1) & \bar{\mathbf{R}}_{33}(\bar{t}-1) \end{bmatrix} & \bar{\mathbf{z}}_p^-(\bar{t}) \\ \bar{\mathbf{R}}_{31}(\bar{t}) & & \bar{\mathbf{z}}_f^+(\bar{t}) \end{bmatrix} \mathbf{Rot}_{G_2}(\bar{t}) \\ &= \begin{bmatrix} \bar{\mathbf{R}}_{11}(\bar{t}) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{21}(\bar{t}) & \bar{\mathbf{R}}_{22}(\bar{t}) & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{31}(\bar{t}) & \bar{\mathbf{R}}_{32}(\bar{t}) & \sqrt{\lambda} \bar{\mathbf{R}}_{33}(\bar{t}-1) & \bar{\mathbf{z}}_f^+(\bar{t}) \end{bmatrix}. \end{aligned} \quad (4.102)$$

Il est alors possible de montrer que la mise à jour du bloc $\bar{\mathbf{R}}_{32}$ est réalisable comme suit [LGV00] :

$$\bar{\mathbf{R}}_{32}(\bar{t}) \bar{\mathbf{R}}_{32}^T(\bar{t}) = \lambda \bar{\mathbf{R}}_{32}(\bar{t}-1) \bar{\mathbf{R}}_{32}^T(\bar{t}-1) + \bar{\mathbf{z}}_f^+(\bar{t}) \bar{\mathbf{z}}_f^+(\bar{t})^T - \bar{\mathbf{z}}_f^+(\bar{t}) \bar{\mathbf{z}}_f^+(\bar{t})^T. \quad (4.103)$$

M. Lovera propose d'estimer récursivement une base de l'espace d'observabilité en suivant l'évolution des vecteurs propres de la matrice de covariance \mathbf{R} définie par :

$$\mathbf{R} = \mathbb{E} \left\{ \bar{\mathbf{z}}_f^+ \bar{\mathbf{z}}_f^+{}^T - \bar{\mathbf{z}}_f^+ \bar{\mathbf{z}}_f^+{}^T \right\}. \quad (4.104)$$

Il suggère plus précisément de considérer le critère suivant [Lov98] :

$$\bar{V}(\boldsymbol{\Omega}(\bar{t})) = \text{Tr}(\mathbf{R}(\bar{t}) - \boldsymbol{\Omega}(\bar{t}) \boldsymbol{\Omega}^T(\bar{t}) \mathbf{R}(\bar{t})) \quad (4.105)$$

qu'il réécrit comme suit :

$$\begin{aligned} \bar{V}(\boldsymbol{\Omega}(\bar{t})) &= \mathbb{E} \left\| \bar{\mathbf{z}}_f^+ - \boldsymbol{\Omega} \boldsymbol{\Omega}^H \bar{\mathbf{z}}_f^+ \right\|^2 - \mathbb{E} \left\| \bar{\mathbf{z}}_f^+ - \boldsymbol{\Omega} \boldsymbol{\Omega}^H \bar{\mathbf{z}}_f^+ \right\|^2 \\ &= \sum_{k=1}^{\bar{t}} \lambda^{\bar{t}-k} \left\| \bar{\mathbf{z}}_f^+(k) - \boldsymbol{\Omega}(\bar{t}) \boldsymbol{\Omega}^H(\bar{t}) \bar{\mathbf{z}}_f^+(k) \right\|^2 - \sum_{k=1}^{\bar{t}} \lambda^{\bar{t}-k} \left\| \bar{\mathbf{z}}_f^+(k) - \boldsymbol{\Omega}(\bar{t}) \boldsymbol{\Omega}^H(\bar{t}) \bar{\mathbf{z}}_f^+(k) \right\|^2. \end{aligned} \quad (4.106)$$

Il propose ensuite d'introduire les approximations classiques suivantes :

$$\bar{\mathbf{h}}_f^+(k) = \boldsymbol{\Omega}^T(k-1) \bar{\mathbf{z}}_f^+(k) \quad (4.107)$$

$$\bar{\mathbf{h}}_f^+(k) = \boldsymbol{\Omega}^T(k-1) \bar{\mathbf{z}}_f^+(k). \quad (4.108)$$

Le critère \bar{V} s'écrit alors :

$$\bar{V}(\mathbf{\Omega}(\bar{t})) = \sum_{k=1}^{\bar{t}} \lambda^{\bar{t}-k} \left\| \bar{\mathbf{z}}_f^+(k) - \mathbf{\Omega}(\bar{t}) \bar{\mathbf{h}}_f^+(k) \right\|^2 - \sum_{k=1}^{\bar{t}} \lambda^{\bar{t}-k} \left\| \bar{\bar{\mathbf{z}}}_f^+(k) - \mathbf{\Omega}(\bar{t}) \bar{\bar{\mathbf{h}}}_f^+(k) \right\|^2. \quad (4.109)$$

Son optimum vaut¹² [Lov98] :

$$\mathbf{\Omega}(\bar{t}) = \left[\mathbf{R}_{\bar{\mathbf{z}}_f^+ \bar{\mathbf{h}}_f^+}(\bar{t}) - \mathbf{R}_{\bar{\bar{\mathbf{z}}}_f^+ \bar{\bar{\mathbf{h}}}_f^+}(\bar{t}) \right] \left[\mathbf{R}_{\bar{\mathbf{h}}_f^+}(\bar{t}) - \mathbf{R}_{\bar{\bar{\mathbf{h}}}_f^+}(\bar{t}) \right]^{-1}. \quad (4.110)$$

La minimisation de la fonction coût \bar{V} est alors réalisable en appliquant par deux fois le lemme d'inversion matricielle¹³ :

$$\bar{\mathbf{h}}_f^+(\bar{t}) = \mathbf{\Omega}^T(\bar{t}-1) \bar{\mathbf{z}}_f^+(\bar{t}) \quad (4.111a)$$

$$\bar{\bar{\mathbf{h}}}_f^+(\bar{t}) = \mathbf{\Omega}^T(\bar{t}-1) \bar{\bar{\mathbf{z}}}_f^+(\bar{t}) \quad (4.111b)$$

$$\bar{\mathbf{L}}_f(\bar{t}) = \frac{1}{\lambda} \left(\bar{\mathbf{L}}_f(\bar{t}-1) - \frac{\bar{\mathbf{L}}_f(\bar{t}-1) \bar{\mathbf{h}}_f^+(\bar{t}) \bar{\mathbf{h}}_f^{+T}(\bar{t}) \bar{\mathbf{L}}_f(\bar{t}-1)}{\lambda + \bar{\mathbf{h}}_f^{+T}(\bar{t}) \bar{\mathbf{L}}_f(\bar{t}-1) \bar{\mathbf{h}}_f^+(\bar{t})} \right) \quad (4.111c)$$

$$\bar{\bar{\mathbf{L}}}_f(\bar{t}) = \bar{\mathbf{L}}_f(\bar{t}) + \frac{\bar{\mathbf{L}}_f(\bar{t}) \bar{\mathbf{h}}_f^+(\bar{t}) \bar{\mathbf{h}}_f^{+T}(\bar{t}) \bar{\mathbf{L}}_f(\bar{t})}{1 - \bar{\mathbf{h}}_f^{+T}(\bar{t}) \bar{\mathbf{L}}_f(\bar{t}) \bar{\mathbf{h}}_f^+(\bar{t})} \quad (4.111d)$$

$$\begin{aligned} \mathbf{\Omega}(\bar{t}) &= \mathbf{\Omega}(\bar{t}-1) + \left(\bar{\mathbf{z}}_f^+(\bar{t}) - \mathbf{\Omega}^T(\bar{t}-1) \bar{\mathbf{h}}_f^+(\bar{t}) \right) \bar{\mathbf{h}}_f^{+T}(\bar{t}) \bar{\mathbf{L}}_f(\bar{t}) \\ &\quad - \left(\bar{\bar{\mathbf{z}}}_f^+(\bar{t}) - \mathbf{\Omega}^T(\bar{t}-1) \bar{\bar{\mathbf{h}}}_f^+(\bar{t}) \right) \bar{\mathbf{h}}_f^{+T}(\bar{t}) \bar{\mathbf{L}}_f(\bar{t}). \end{aligned} \quad (4.111e)$$

Cet algorithme sera nommé UDPAST dans la suite de ce manuscrit. Sa complexité numérique est en $\mathcal{O}(n_x n_z)$.

Remarque 4.5. *M. Lovera présente également un algorithme de mise à jour du schéma d'identification PO_EIV MOESP [LGV00]. Ce dernier repose sur une approximation équivalente à celle développée par T. Gustafsson [Gus97a] puisqu'il utilise $\hat{\mathbf{H}}_f(t-1)$ construite à partir des estimées des matrices d'état du système. Il suggère également d'appliquer la version récursive PI MOESP pour estimer en ligne la partie linéaire d'un modèle de Wiener.*

4.3.3 Adaptation des algorithmes COPAST et COPASTd

Bien que les méthodes précédentes aient montré leur efficacité sur de nombreux exemples de simulation [Lov98, LGV00], elles présentent l'inconvénient d'utiliser deux algorithmes d'estimation distincts selon l'origine du vecteur de sortie modifié estimé (EIVPAST ou UDPAST). Or, pour faciliter son utilisation, il serait intéressant de pouvoir appliquer un unique algorithme de minimisation capable de traiter, de manière uniforme, les données obtenues par mise à jour des matrices $\bar{\mathbf{R}}_{22}$ et $\bar{\mathbf{R}}_{32}$. C'est ce que nous réalisons au sein de ce paragraphe en adaptant les

¹²En supposant que la matrice inverse existe.

¹³A notre connaissance, cet algorithme n'a jamais été présenté explicitement.

techniques COPAST et COPASTd (cf. §4.2.3.3) à notre problème d'identification. Afin de bien comprendre la démarche suivie dans ce paragraphe, considérons de nouveau les algorithmes développés par M. Lovera (cf. §4.3.2). Une analyse rapide de ces derniers montrent que l'estimation de $\mathbf{\Omega}$ est réalisée, pour chaque schéma d'identification (cf. §4.3.2.1 et 4.3.2.2), à partir de matrices de corrélation spécifiques :

- $\mathbf{R}_{z\xi} = \mathbb{E} \left\{ \mathbf{z}\xi^T \right\}$ pour la classe d'identification Ordinary MOESP ;
- $\mathbf{R} = \mathbb{E} \left\{ \bar{\mathbf{z}}_f^+ \bar{\mathbf{z}}_f^{+T} \right\} - \left\{ \bar{\bar{\mathbf{z}}}_f^+ \bar{\bar{\mathbf{z}}}_f^{+T} \right\}$ pour la classe d'identification PI/PO MOESP.

Notre objectif est de proposer un algorithme capable de traiter indifféremment l'une ou l'autre de ces matrices de covariance et fournir une estimée consistante de la matrice d'observabilité. Dans le paragraphe 4.2.3, nous avons présenté trois algorithmes (COPAST1, COPAST2 et COPASTd) minimisant des critères fondés sur des matrices de covariance. Il semble donc naturel de tenter de les adapter à notre problème d'estimation. Le critère considéré est le suivant :

$$\bar{V}(\mathbf{\Omega}) = \|\check{\mathbf{R}} - \mathbf{\Omega}\mathbf{\Omega}^H\check{\mathbf{R}}\|_F^2 \quad (4.112)$$

où $\check{\mathbf{R}}$ représente respectivement $\mathbf{R}_{z\xi}$ ou \mathbf{R} selon le schéma d'identification étudié. Deux situations sont alors à envisagées¹⁴ :

- L'algorithme COPAST1 (cf. equ. (4.39)) est basé sur l'introduction de la variable \mathbf{h} définie par :

$$\mathbf{h}(k) = \mathbf{\Omega}^H(k-1)\mathbf{z}(k). \quad (4.113)$$

Cette expression nécessite, par définition, de connaître à chaque instant le vecteur \mathbf{z} . Ce dernier n'est malheureusement pas directement estimé par la classe d'identification PI/PO MOESP. Elle ne fournit en effet qu'une estimation de son carré et ne permet pas d'extraire le vecteur \mathbf{z} . L'algorithme COPAST1 n'est donc pas utilisable pour résoudre notre objectif initial : proposer un algorithme de minimisation capable de traiter, de manière uniforme, les données fournies par les classes d'identification Ordinary et PI/PO MOESP.

- La seconde solution consiste quant à elle à minimiser séquentiellement $n_{c_{\check{\mathbf{R}}}}$ fonctions coût définies comme suit :

$$\bar{V}_i(\mathbf{\Omega}(t)) = \sum_{k=1}^t \lambda^{t-k} \|\check{\mathbf{r}}^i(k) - \mathbf{\Omega}(t)\mathbf{\Omega}^H(t)\check{\mathbf{r}}^i(k)\|^2 \quad (4.114)$$

où $n_{c_{\check{\mathbf{R}}}}$ symbolise le nombre de colonnes de $\check{\mathbf{R}}$ et $\check{\mathbf{r}}^i(k)$ désigne la i^e colonne de $\check{\mathbf{R}}$. L'approximation suivante est ensuite considérée :

$$\mathbf{\Omega}^H(t)\check{\mathbf{r}}^i(k) \approx \mathbf{\Omega}^H(k-1)\check{\mathbf{r}}^i(k) = \check{\mathbf{h}}^i(k). \quad (4.115)$$

Le critère précédent s'écrit donc :

$$\tilde{\bar{V}}_i(\mathbf{\Omega}(t)) = \sum_{k=1}^t \lambda^{t-k} \|\check{\mathbf{r}}^i(k) - \mathbf{\Omega}(t)\check{\mathbf{h}}^i(k)\|^2. \quad (4.116)$$

¹⁴Rappelons que l'algorithme COPASTd n'est en fait qu'une réécriture de l'algorithme COPAST2.

Les méthodes de traitement d'antennes COPAST2 et COPASTd peuvent alors être appliquées pour minimiser cette fonction coût approximée. L'adaptation de la technique COPAST2 à la fonction coût (4.116) conduit plus précisément à l'algorithme suivant :

$$i = \text{reste}(t, n_{c_{\check{\mathbf{R}}}}) + 1 \quad (4.117a)$$

$$\check{\mathbf{r}}^i(t) = \check{\mathbf{R}}(:, i)(t) \quad (4.117b)$$

$$\check{\mathbf{h}}^i(t) = \mathbf{\Omega}^T(t-1)\check{\mathbf{r}}^i(t) \quad (4.117c)$$

$$\mathbf{K}(t) = \frac{\check{\mathbf{h}}^{iT}(t)\mathbf{L}(t-1)}{\lambda + \check{\mathbf{h}}^{iT}(t)\mathbf{L}(t-1)\check{\mathbf{h}}^i(t)} \quad (4.117d)$$

$$\mathbf{L}(t) = \frac{1}{\lambda} (\mathbf{L}(t-1) - \mathbf{L}(t-1)\check{\mathbf{h}}^i(t)\mathbf{K}(t)) \quad (4.117e)$$

$$\mathbf{\Omega}(t) = \mathbf{\Omega}(t-1) + (\check{\mathbf{r}}^i(t) - \mathbf{\Omega}(t-1)\check{\mathbf{h}}^i(t)) \mathbf{K}(t). \quad (4.117f)$$

L'application de COPASTd permet quant à lui de construire l'algorithme suivant :

$$i = \text{reste}(t, n_{c_{\check{\mathbf{R}}}}) + 1$$

$$\check{\mathbf{r}}^i(t) = \check{\mathbf{R}}(:, i)(t)$$

$$\check{\mathbf{r}}_1^i(t) = \check{\mathbf{r}}^i(t)$$

Pour $k = 1 : n_x$

$$\check{h}_k^i(t) = \boldsymbol{\omega}_k^T(t-1)\check{\mathbf{r}}_k^i(t) \quad (4.118a)$$

$$d_k(t) = \lambda d_k(t-1) + |\check{h}_k^i(t)|^2 \quad (4.118b)$$

$$\boldsymbol{\omega}_k(t) = \boldsymbol{\omega}_k(t-1) + (\check{\mathbf{r}}_k^i(t) - \boldsymbol{\omega}_k(t-1)\check{h}_k^i(t)) \frac{\check{h}_k^{i*}(t)}{d_k(t)} \quad (4.118c)$$

$$\check{\mathbf{r}}_{k+1}^i(t) = \check{\mathbf{r}}_k^i(t) - \boldsymbol{\omega}_k(t)\check{h}_k^i(t). \quad (4.118d)$$

fin

Ces deux algorithmes d'identification récursive seront respectivement nommés COPAST et COPASTd dans la suite de ce manuscrit. Nous les appellerons COIVPAST et COIVPASTd lorsqu'il sera nécessaire de préciser que la matrice de covariance $\check{\mathbf{R}}$ contient une variable instrumentale.

Remarque 4.6. *L'atout majeur de l'approche présentée au sein de ce paragraphe est qu'elle repose sur une démarche unifiée pour deux schémas d'identification distincts. De plus, l'utilisation d'un algorithme de déflation permet d'avoir accès à des estimées instantanées des valeurs propres du système. Ces dernières pourraient être utilisées pour estimer en ligne l'ordre du procédé.*

4.3.4 Algorithmes développés par H. Oku [OK99, OK02]

Nous avons montré dans le paragraphe 3.3.5 que les matrices de résidus des schémas d'identification Ordinary et PI/PO MOESP vérifiaient respectivement (cf. equ. (3.119) et (3.133)) :

$$\mathbf{R}_\varepsilon^{\text{Ord}} = \mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{Y}_f^{+T} = \mathbf{\Gamma}_f \left(\mathbf{R}_x - \mathbf{R}_{\mathbf{xu}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{\mathbf{xu}_f^+}^T \right) \mathbf{\Gamma}_f^T + \mathbf{R}_{\mathbf{v}_f^+} \quad (4.119)$$

$$\begin{aligned} \mathbf{R}_\varepsilon^{\text{PI/PO}} &= \mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T \left(\mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T \right)^{-1} \mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{Y}_f^+ \\ &= \mathbf{\Gamma}_f \left(\mathbf{R}_{\mathbf{x}\xi} - \mathbf{R}_{\mathbf{xu}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{\mathbf{u}_f^+} \xi \right) \left(\mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T \right)^{-1} \left(\mathbf{R}_{\mathbf{x}\xi} - \mathbf{R}_{\mathbf{xu}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{\mathbf{u}_f^+} \xi \right)^T \mathbf{\Gamma}_f^T. \end{aligned} \quad (4.120)$$

L'estimation de la matrice d'observabilité du système est alors réalisable en appliquant à ces dernières une décomposition en valeurs propres [Liu92, Vib95] puisque le sous-espace engendré par les vecteurs propres gauches de $\mathbf{R}_\varepsilon^{\text{Ord}}$ (resp. $\mathbf{R}_\varepsilon^{\text{PI/PO}}$) génère la matrice d'observabilité $\mathbf{\Gamma}_f$. H. Oku s'inspire de cette constatation pour proposer deux algorithmes récursifs d'estimation d'une base de la matrice d'observabilité du procédé à identifier. Il suggère, dans un premier temps, de mettre à jour les relations (4.119) et (4.120) à l'aide du lemme d'inversion matricielle [OK99, OK02]. Les algorithmes suivants sont alors énoncés :

Proposition 4.1. *Supposons que les matrices $\mathbf{T}_f \triangleq \mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{Y}_f^{+T}$, $\mathbf{M}_f \triangleq \left(\mathbf{U}_f^+ \mathbf{U}_f^{+T} \right)^{-1}$ et $\mathbf{Y}_f^+ \mathbf{U}_f^{+T}$ aient été estimées à l'instant $\bar{t} - 1$. Lorsque de nouvelles données sont accessibles à la mesure, la mise à jour de la matrice \mathbf{T}_f est réalisée récursivement en adoptant la procédure suivante :*

$$\boldsymbol{\alpha}_f(\bar{t}) = \left(1 + \mathbf{u}_f^{+T}(\bar{t}) \mathbf{M}_f(\bar{t} - 1) \mathbf{u}_f^+(\bar{t}) \right)^{-1} \quad (4.121a)$$

$$\mathbf{e}_f^+(\bar{t}) = \mathbf{y}_f^+(\bar{t}) - \mathbf{Y}_f^+(\bar{t} - 1) \mathbf{U}_f^{+T}(\bar{t} - 1) \mathbf{u}_f^+(\bar{t}) \quad (4.121b)$$

$$\mathbf{T}_f(\bar{t}) = \mathbf{T}_f(\bar{t} - 1) + \boldsymbol{\alpha}_f(\bar{t}) \mathbf{e}_f^+(\bar{t}) \mathbf{e}_f^{+T}(\bar{t}) \quad (4.121c)$$

$$\mathbf{M}_f(\bar{t}) = \mathbf{M}_f(\bar{t} - 1) - \boldsymbol{\alpha}_f(\bar{t}) \mathbf{M}_f(\bar{t} - 1) \mathbf{u}_f^+(\bar{t}) \mathbf{u}_f^{+T}(\bar{t}) \mathbf{M}_f(\bar{t} - 1) \quad (4.121d)$$

$$\mathbf{Y}_f^+(\bar{t}) \mathbf{U}_f^{+T}(\bar{t}) = \mathbf{Y}_f^+(\bar{t} - 1) \mathbf{U}_f^{+T}(\bar{t} - 1) + \mathbf{y}_f^+(\bar{t}) \mathbf{u}_f^{+T}(\bar{t}). \quad (4.121e)$$

Proposition 4.2. *Supposons que les matrices $\mathbf{M}_f \triangleq \left(\mathbf{U}_f^+ \mathbf{U}_f^{+T} \right)^{-1}$, $\mathbf{Y}_f^+ \mathbf{U}_f^{+T}$, $\mathbf{\Xi} \mathbf{U}_f^{+T}$, $\Phi_f \triangleq \left(\mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T \right)$ et $\mathbf{T}_f \triangleq \mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T \left(\mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{\Xi}^T \right)^{-1} \mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^+} \mathbf{Y}_f^{+T}$ aient été estimées à l'instant $\bar{t} - 1$. Lorsque de nouvelles données sont accessibles à la mesure, la mise à jour de la matrice \mathbf{T}_f est réalisée récursivement en adoptant la procédure suivante :*

$$\boldsymbol{\alpha}_f(\bar{t}) = \left(1 + \mathbf{u}_f^{+T}(\bar{t}) \mathbf{M}_f(\bar{t} - 1) \mathbf{u}_f^+(\bar{t}) \right)^{-1} \quad (4.122a)$$

$$\mathbf{e}_f^+(\bar{t}) = \mathbf{y}_f^+(\bar{t}) - \mathbf{Y}_f^+(\bar{t} - 1) \mathbf{U}_f^{+T}(\bar{t} - 1) \mathbf{u}_f^+(\bar{t}) \quad (4.122b)$$

$$\mathbf{q}_f^+(\bar{t}) = \Phi_f(\bar{t} - 1) \mathbf{U}_f^{+T}(\bar{t} - 1) \mathbf{M}_f(\bar{t} - 1) \mathbf{u}_f^+(\bar{t}) - \xi(\bar{t}) \quad (4.122c)$$

$$\mathbf{g}_f^+(\bar{t}) = \mathbf{e}_f^+(\bar{t}) + \mathbf{Y}_f^+(\bar{t} - 1) \mathbf{\Pi}_{\mathbf{U}_f^+}(\bar{t} - 1) \mathbf{\Xi}^T(\bar{t} - 1) \Phi_f(\bar{t} - 1) \mathbf{q}_f^+(\bar{t}) \quad (4.122d)$$

$$\boldsymbol{\beta}_f(\bar{t}) = \left(\frac{1}{\boldsymbol{\alpha}_f(\bar{t})} + \mathbf{q}_f^{+T}(\bar{t})\boldsymbol{\Phi}_f(\bar{t}-1)\mathbf{q}_f^+(\bar{t}) \right)^{-1} \quad (4.122e)$$

$$\mathbf{Y}_f^+(\bar{t})\boldsymbol{\Pi}_{\mathbf{U}_f^{+\perp}}(\bar{t})\boldsymbol{\Xi}^T(\bar{t}) = \mathbf{Y}_f^+(\bar{t}-1)\boldsymbol{\Pi}_{\mathbf{U}_f^{+\perp}}(\bar{t}-1)\boldsymbol{\Xi}^T(\bar{t}-1) - \boldsymbol{\alpha}_f(\bar{t})\mathbf{e}_f^+(\bar{t})\mathbf{q}_f^{+T}(\bar{t}) \quad (4.122f)$$

$$\boldsymbol{\Phi}_f(\bar{t}) = \boldsymbol{\Phi}_f(\bar{t}-1) - \boldsymbol{\beta}_f(\bar{t})\boldsymbol{\Phi}_f(\bar{t}-1)\mathbf{q}_f^+(\bar{t})\mathbf{q}_f^{+T}(\bar{t})\boldsymbol{\Phi}_f(\bar{t}-1) \quad (4.122g)$$

$$\mathbf{T}_f(\bar{t}) = \mathbf{T}_f(\bar{t}-1) + \boldsymbol{\alpha}_f(\bar{t})\mathbf{e}_f^+(\bar{t})\mathbf{e}_f^{+T}(\bar{t}) - \boldsymbol{\beta}_f(\bar{t})\mathbf{g}_f^+(\bar{t})\mathbf{g}_f^{+T}(\bar{t}) \quad (4.122h)$$

$$\mathbf{M}_f(\bar{t}) = \mathbf{M}_f(\bar{t}-1) - \boldsymbol{\alpha}_f(\bar{t})\mathbf{M}_f(\bar{t}-1)\mathbf{u}_f^+(\bar{t})\mathbf{u}_f^{+T}(\bar{t})\mathbf{M}_f(\bar{t}-1) \quad (4.122i)$$

$$\mathbf{Y}_f^+(\bar{t})\mathbf{U}_f^{+T}(\bar{t}) = \mathbf{Y}_f^+(\bar{t}-1)\mathbf{U}_f^{+T}(\bar{t}-1) + \mathbf{y}_f^+(\bar{t})\mathbf{u}_f^{+T}(\bar{t}) \quad (4.122j)$$

$$\boldsymbol{\Xi}(\bar{t})\mathbf{U}_f^{+T}(\bar{t}) = \boldsymbol{\Xi}(\bar{t}-1)\mathbf{U}_f^{+T}(\bar{t}-1) + \boldsymbol{\xi}(\bar{t})\mathbf{u}_f^{+T}(\bar{t}). \quad (4.122k)$$

Une fois les matrices de covariance $\mathbf{R}_{\mathbf{e}}^{\text{Ord}}$ et $\mathbf{R}_{\mathbf{e}}^{\text{PI/PO}}$ mises à jour récursivement, l'étape suivante consiste à extraire la matrice d'observabilité du système identifié. Puisque l'application en ligne d'une décomposition en valeurs propres n'est pas concevable de par son coût calculatoire élevé, H. Oku propose d'introduire certaines estimées des algorithmes précédents au sein de l'algorithme de descente du gradient présenté au paragraphe 4.2.3.1 :

$$\begin{aligned} \boldsymbol{\Omega}(\bar{t}) &= \boldsymbol{\Omega}(\bar{t}-1) \\ &- \mu \left(-2\mathbf{R}_z(\bar{t}) + \mathbf{R}_z(\bar{t})\boldsymbol{\Omega}(\bar{t}-1)\boldsymbol{\Omega}^H(\bar{t}-1) + \boldsymbol{\Omega}(\bar{t}-1)\boldsymbol{\Omega}^H(\bar{t}-1)\mathbf{R}_z(\bar{t}) \right) \boldsymbol{\Omega}(\bar{t}-1). \end{aligned} \quad (4.123)$$

Il justifie cette démarche pour la classe d'identification Ordinary MOESP en remarquant que le vecteur \mathbf{e}_f^+ vérifie :

$$\mathbf{e}_f^+(\bar{t}) = \mathbf{y}_f^+(\bar{t}) - \mathbf{Y}_f^+(\bar{t}-1)\mathbf{U}_f^{+T}(\bar{t}-1)\mathbf{u}_f^+(\bar{t}) = \tilde{\mathbf{e}}_f^+(\bar{t}) + \boldsymbol{\delta}_{\mathbf{e}_f^+}(\bar{t}) \quad (4.124)$$

en posant :

$$\tilde{\mathbf{e}}_f^+(\bar{t}) = \boldsymbol{\Gamma}_f \left(\mathbf{x}(\bar{t}) - \mathbf{X}^+(\bar{t}-1)\mathbf{U}_f^{+T}(\bar{t}-1)\mathbf{M}_f(\bar{t}-1)\mathbf{u}_f^+(\bar{t}) \right) \quad (4.125)$$

$$\boldsymbol{\delta}_{\mathbf{e}_f^+}(\bar{t}) = \mathbf{n}_f^+(\bar{t}) - \mathbf{N}_f^+(\bar{t}-1)\mathbf{M}_f(\bar{t}-1)\mathbf{u}_f^+(\bar{t}). \quad (4.126)$$

H. Oku suggère alors de normaliser¹⁵ l'équation (4.124) qui s'écrit alors :

$$\boldsymbol{\kappa}_f^+(\bar{t}) = \frac{\mathbf{e}_f^+(\bar{t})}{\|\mathbf{e}_f^+(\bar{t})\|} = \tilde{\boldsymbol{\kappa}}_f^+(\bar{t}) + \boldsymbol{\delta}_{\boldsymbol{\kappa}_f^+}(\bar{t}). \quad (4.127)$$

Puisque le vecteur $\boldsymbol{\kappa}_f^+$ appartient à l'espace engendré par les colonnes de $\boldsymbol{\Gamma}_f$, il existe alors un vecteur $\boldsymbol{\chi} \in \mathbb{R}^{n_x \times 1}$ tel que :

$$\tilde{\boldsymbol{\kappa}}_f^+ = \boldsymbol{\Gamma}_f \boldsymbol{\chi}. \quad (4.128)$$

H. Oku applique ensuite l'algorithme de descente de gradient (4.123) à la matrice de covariance $\mathbf{R}_{\boldsymbol{\kappa}_f^+}$ calculée comme suit :

$$\mathbf{R}_{\boldsymbol{\kappa}_f^+}(\bar{t}) = \frac{1}{\bar{t}} \sum_{k=1}^{\bar{t}} \boldsymbol{\kappa}_f^+(k)\boldsymbol{\kappa}_f^{+T}(k). \quad (4.129)$$

¹⁵Il ne s'intéresse qu'aux directions de \mathbf{e}_f^+ pour estimer $\boldsymbol{\Gamma}_f$.

Il démontre que, dans le cas d'un bruit dont la matrice de covariance vérifie $\mathbf{R}_n \ll \mathbf{I}$, l'algorithme de descente de gradient converge vers le sous-espace optimal [OK02] avec une probabilité de 1.

Dans de nombreuses situations pratiques, l'hypothèse $\mathbf{R}_n \ll \mathbf{I}$ n'est pas vérifiée. La solution précédente ne peut alors fournir que des estimées asymptotiquement biaisées. H. Oku propose donc d'utiliser l'algorithme (4.122) pour mettre à jour la matrice de covariance $\mathbf{R}_{\kappa_f^+}$ afin de l'exploiter via la méthode de descente de gradient (4.123). La relation suivante est considérée :

$$\mathbf{R}_{\kappa_f^+}(\bar{t}) = \frac{\bar{t}-1}{\bar{t}} \mathbf{R}_{\kappa_f^+}(\bar{t}-1) + \frac{1}{\bar{t}} \frac{\boldsymbol{\alpha}_f(\bar{t}) \mathbf{e}_f^+(\bar{t}) \mathbf{e}_f^{+T}(\bar{t}) - \boldsymbol{\beta}_f(\bar{t}) \mathbf{g}_f^+(\bar{t}) \mathbf{g}_f^{+T}(\bar{t})}{\left\| \boldsymbol{\alpha}_f(\bar{t}) \mathbf{e}_f^+(\bar{t}) \mathbf{e}_f^{+T}(\bar{t}) - \boldsymbol{\beta}_f(\bar{t}) \mathbf{g}_f^+(\bar{t}) \mathbf{g}_f^{+T}(\bar{t}) \right\|}. \quad (4.130)$$

Son introduction au sein de l'algorithme de descente de gradient (4.24) permet d'estimer en ligne une base de l'espace d'observabilité de manière consistante.

4.3.5 Algorithme développé par M. Lovera [Lov03, MLL04]

Les algorithmes d'estimation de l'espace engendré par les colonnes de $\boldsymbol{\Gamma}_f$ reposent jusqu'ici soit sur un ensemble d'approximations (cf. EIVPAST, UDPAST...), soit sur un algorithme d'ordre quatre (cf. §4.3.4). W. Utschick a développé en 2002 [Uts02] un algorithme de poursuite des sous-espaces fondé sur le critère avec contrainte suivant :

$$\bar{V}(\boldsymbol{\Pi}) = \mathbb{E} \|\mathbf{z} - \boldsymbol{\Pi} \mathbf{z}\|^2 \quad (4.131)$$

qui estime récursivement la matrice de projection sur le sous-espace recherché au lieu d'une base de vecteur propre de ce dernier (cf. §4.2.3.4). M. Lovera propose d'adapter cette technique de minimisation au problème d'identification des sous-espaces. Cette méthode présente en effet l'avantage de fournir un critère quadratique sans approximation. Malheureusement, comme la majorité de ces homologues de traitement d'antennes, l'algorithme originel du projecteur [Uts02] suppose que le bruit agissant sur le système est temporellement et spatialement blanc. Il n'est donc pas utilisable lorsque les perturbations ne vérifient pas cette hypothèse. M. Lovera [Lov03] suggère alors d'introduire une variable instrumentale au sein du critère (4.48). Lorsque les matrices et données utilisées sont à composantes réelles, la fonction coût considérée s'écrit alors :

$$\bar{V}_{IV}(\boldsymbol{\Pi}) = \left\| \mathbf{R}_{z_f^+ \xi} - \boldsymbol{\Pi} \mathbf{R}_{z_f^+ \xi} \right\|_F^2 = \text{Tr} \left(\left(\mathbf{R}_{z_f^+ \xi} - \boldsymbol{\Pi} \mathbf{R}_{z_f^+ \xi} \right) \left(\mathbf{R}_{z_f^+ \xi} - \boldsymbol{\Pi} \mathbf{R}_{z_f^+ \xi} \right)^T \right). \quad (4.132)$$

En posant :

$$\mathbf{R} = \mathbf{R}_{z_f^+ \xi} \mathbf{R}_{z_f^+ \xi}^T, \quad (4.133)$$

ce critère devient :

$$\bar{V}_{IV}(\boldsymbol{\Pi}) = \text{Tr}(\mathbf{R}) - \text{Tr}(\mathbf{R}\boldsymbol{\Pi}). \quad (4.134)$$

La structure de ce dernier est similaire à celle du critère (4.48). Une approche de type variable instrumentale peut donc être appliquée au problème de poursuite du projecteur en suivant la même démarche que celle proposée par W. Utschick. Dans le cas particulier de données

réelles¹⁶, il est montré que [Lov03, MLL04] :

$$\delta_i = 0 \forall i \Rightarrow \mathbf{\Delta} = \mathbf{I} \quad (4.135)$$

$$\vartheta_{k,\ell}(t) = -\mu \frac{\partial \bar{V}(\mathbf{\Pi}(t-1))}{\partial \vartheta_{k,\ell}} = -2\mu \left(\sum_{j=1}^{n_z} \mathbf{\Pi}_{\ell,j} \mathbf{R}_{j,k} - \mathbf{\Pi}_{k,j} \mathbf{R}_{j,\ell} \right) \quad (4.136)$$

où $\mathbf{\Pi}_{i,j}$ et $\mathbf{R}_{i,j}$ symbolisent l'élément (i, j) de $\mathbf{\Pi}$ et \mathbf{R} . La mise à jour de la matrice $\mathbf{\Omega}$ est alors réalisée comme suit :

$$\mathbf{\Omega}(t) = \mathbf{\Theta}(\vartheta(t)) \mathbf{\Omega}(t-1). \quad (4.137)$$

Cette méthode sera nommée IVPT dans la suite de ce manuscrit.

Remarque 4.7. *Contrairement aux algorithmes fondés sur le critère PAST, l'algorithme IVPT ne présente aucune approximation. Il est cependant impératif de lui associer une bonne initialisation afin de garantir au projecteur d'être dans un voisinage « proche » de la solution dès les premières itérations. Cette condition est nécessaire à la validation du théorème 4.3.*

4.3.6 Synthèse

Bien que toutes basées sur l'adaptation d'un seul et même critère, les méthodes récursives d'identification exposées jusqu'ici se différencient par les fonctions coût construites pour estimer l'espace d'observabilité du système à modéliser ainsi que par les algorithmes d'optimisation utilisés pour les minimiser. Nous proposons donc de regrouper ces techniques d'estimation au sein du tableau 4.1 en insistant sur les éléments fondamentaux les distinguant : le critère et l'algorithme de minimisation.

| Nom | Critère | Outil ou algorithme de minimisation |
|----------|--|--|
| IVPAST | $\tilde{V}_{IV}(\mathbf{\Omega}(t)) = \left\ \mathbf{R}_z \boldsymbol{\xi}(t) - \mathbf{\Omega}(t) \mathbf{R}_{h\xi}(t) \right\ _F^2$ | Variable instrumentale récursive |
| EIVPAST | $\tilde{V}_{IV}(\mathbf{\Omega}(t)) = \left\ \mathbf{R}_z \boldsymbol{\xi}(t) - \mathbf{\Omega}(t) \mathbf{R}_{h\xi}(t) \right\ _F^2$ | Variable instrumentale récursive étendue |
| UDPAST | $\bar{V}(\mathbf{\Omega}(\bar{t})) = \text{Tr}(\mathbf{R}(\bar{t}) - \mathbf{\Omega}(\bar{t}) \mathbf{\Omega}^T(\bar{t}) \mathbf{R}(\bar{t}))$ | Deux moindres carrés récursifs combinés |
| COPAST | $\tilde{V}_i(\mathbf{\Omega}(t)) = \sum_{k=1}^t \lambda^{t-k} \left\ \check{\mathbf{r}}^i(k) - \mathbf{\Omega}(t) \check{\mathbf{h}}^i(k) \right\ ^2$ pour $i \in [1, n_{c_R}]$ | Moindres carrés récursifs |
| COPASTd | $\tilde{V}_i(\mathbf{\Omega}(t)) = \sum_{k=1}^t \lambda^{t-k} \left\ \check{\mathbf{r}}^i(k) - \mathbf{\Omega}(t) \check{\mathbf{h}}^i(k) \right\ ^2$ pour $i \in [1, n_{c_R}]$ | Moindres carrés récursifs+déflation |
| Yanggrad | $V(\mathbf{\Omega}) = \mathbb{E} \left\ \mathbf{z} - \mathbf{\Omega} \mathbf{\Omega}^H \mathbf{z} \right\ ^2$ | Descente de gradient |
| IVPT | $\tilde{V}_{IV}(\mathbf{\Pi}) = \left\ \mathbf{R}_{z^+} \boldsymbol{\xi} - \mathbf{\Pi} \mathbf{R}_{z^+} \boldsymbol{\xi} \right\ _F^2$ | Rotations de Givens |

TAB. 4.1: Méthodes d'identification récursives des sous-espaces fondées sur le critère de Yang.

¹⁶En identification, les données utilisées $\in \mathbb{R}$.

4.4 Conclusion

Ce chapitre nous a permis de présenter une vue d'ensemble des méthodes d'identification récursive des sous-espaces développées jusqu'en 2003. Ces techniques ont comme point commun d'être fondées sur l'adaptation au problème d'identification d'un critère particulier du traitement d'antennes : le critère de Yang. Cette étude a ainsi permis de mettre en évidence que les méthodes proposées, bien que tirant leur origine d'une même fonction coût, présentent des caractéristiques différentes en termes de charge calculatoire, de nombre d'approximations utilisées, de contraintes posées... Ainsi :

- les algorithmes développés par T. Gustafsson proposent d'utiliser une version approximée du critère de Yang associé à une variable instrumentale pour estimer la matrice d'observabilité du système dans un contexte stochastique ;
- M. Lovera améliore, en deux temps, les performances et la fiabilité des algorithmes de T. Gustafsson en proposant successivement :
 - * de substituer à l'approche de soustraction la mise à jour des décomposition QR des schémas d'identification Ordinary et PI/PO MOESP pour estimer le vecteur de sortie modifié,
 - * de remanier le critère de Yang pour fournir un critère quadratique avec contraintes en introduisant le concept du projecteur ;
- les travaux de H. Oku reposent sur une approche particulière de la phase de mise à jour des matrices de données d'entrée-sortie des schémas d'identification Ordinary et PI/PO MOESP. Il propose d'utiliser le lemme d'inversion matricielle pour actualiser deux matrices symétriques particulières à chaque acquisition. La méthode du gradient de B. Yang est quant à elle utilisée pour estimer la matrice d'observabilité du système.

L'étude de ces algorithmes montrent plus précisément que la complexité originelle du critère de Yang a conduit les auteurs à proposer plusieurs astuces pour estimer, de manière récursive, la matrice d'observabilité du système. En effet :

- T. Gustafsson emploie une version approximée du critère originel de Yang afin d'appliquer des algorithmes de minimisation de type moindres carrés.
- H. Oku doit utiliser des algorithmes itératifs de type gradient pour atteindre le minimum du critère d'ordre 4 considéré,
- M. Lovera construit des fonctions coût d'ordre 2 avec contraintes complexifiant grandement les techniques de minimisation à appliquer et nécessitant une phase d'initialisation optimale.

Or, comme nous l'avons décrit dans le paragraphe 4.2, il existe, en traitement du signal, des méthodes d'estimation édifiées sur des principes plus simples que ceux exposés par B. Yang. C'est notamment le cas de la méthode du propagateur (cf. §4.2.4) puisque cette dernière permet de décomposer l'espace d'observation en deux sous-espaces complémentaires à l'aide d'opérations strictement linéaires. L'objectif du chapitre 5 consiste donc à développer de nouveaux critères et de nouveaux algorithmes fondés sur cet opérateur linéaire afin d'éliminer les approximations et les contraintes liées à l'application du critère de Yang.

Chapitre 5

Adaptation de la méthode du propagateur à l'identification récursive des sous-espaces

Sommaire

| | | |
|------------|---|------------|
| 5.1 | Problématique | 93 |
| 5.2 | Analogie entre l'identification des sous-espaces et le traitement d'antennes | 95 |
| 5.3 | Mise à jour du vecteur d'observation | 97 |
| 5.3.1 | Dans un contexte déterministe | 100 |
| 5.3.2 | Dans un contexte stochastique | 104 |
| 5.3.3 | Complexité numérique et fiabilité des algorithmes de mise à jour du vecteur d'observation | 107 |
| 5.3.4 | Conclusion | 108 |
| 5.4 | Estimation récursive d'une base de l'espace d'observabilité | 109 |
| 5.4.1 | Adaptation de la méthode du propagateur au problème d'identification | 109 |
| 5.4.2 | Estimation à partir d'un vecteur d'observation épuré des effets des perturbations | 113 |
| 5.4.3 | Introduction d'une variable instrumentale | 116 |
| 5.4.4 | Synthèse | 124 |
| 5.4.5 | Propriétés asymptotiques de la méthode du propagateur | 125 |
| 5.4.6 | Complexité numérique des algorithmes d'estimation de la matrice d'observabilité | 132 |
| 5.5 | Complexité numérique des algorithmes d'identification récursive : une étude de cas | 134 |
| 5.6 | Estimation des matrices d'état | 136 |
| 5.7 | Conclusion | 138 |

5.1 Problématique

L'objectif de ce chapitre est de présenter de nouveaux algorithmes récursifs d'identification permettant la mise à jour, à chaque nouvelle acquisition, d'un modèle sous forme d'état à partir

de données d'entrée-sortie perturbées. Les solutions proposées sont, comme dans le chapitre 4, fondées sur une approche de type sous-espace. Cette approche présente en effet l'avantage de fournir un modèle d'état du procédé sans nécessiter de développer de représentations intermédiaires de ce dernier (fonction de transfert, réponse impulsionnelle...). Les algorithmes développés s'inspirent plus particulièrement du principe des méthodes d'identification MOESP (cf. §3.3). Le principal intérêt de ces techniques d'identification hors ligne repose sur le fait qu'elles permettent d'estimer la matrice d'observabilité étendue à partir des données d'entrée-sortie sans calcul explicite du vecteur d'état¹. Plusieurs algorithmes récursifs partageant ces caractéristiques ont été présentés au sein du chapitre précédent. Ces derniers ont en commun d'être fondés sur l'utilisation d'une fonction coût particulière : le critère de Yang. On peut alors s'interroger sur la nécessité de développer de nouvelles techniques d'estimation récursives puisque les méthodes temps réel proposées jusqu'ici ont montré, tout au moins en simulation [LGV00, OK02, Lov03], des capacités de suivi intéressantes. La description réalisée dans le chapitre 4 a cependant mis en évidence un certain nombre d'inconvénients, à savoir :

- que toutes les méthodes développées sont restreintes à un seul et même principe : minimiser le critère de Yang et ses dérivées ;
- qu'elles utilisent respectivement :
 - * un critère *d'ordre 4* uniquement minimisable par des techniques de type gradient stochastique [OK99, OK02],
 - * un critère quadratique *approximé* [Gus98, Lov98, LGV00],
 - * un critère quadratique *avec contrainte* nécessitant d'appliquer des algorithmes de descente incrémentaux associés à une très bonne phase d'initialisation [Lov03].

Toutes ces difficultés nous ont incités à développer un nouveau critère ayant comme principales caractéristiques d'être *quadratique, sans approximation et sans contrainte*. Les algorithmes d'estimation présentés au sein de ce chapitre s'inspirent plus précisément d'une technique particulière de traitement d'antennes : la méthode du propagateur [Fri84] (PM). Celle-ci repose sur une décomposition spécifique de la matrice interspectrale des données des capteurs de l'antenne permettant, à l'aide d'un opérateur nommé propagateur, d'estimer simplement les sous-espaces signal et bruit (cf. §4.2.4). L'adaptation de cet opérateur linéaire au problème d'identification conduit à un critère quadratique, sans contrainte et sans aucune approximation, qu'il est aisé de minimiser par des techniques de type moindres carrés (cf. §5.4). Toutefois, avant de poursuivre dans la description de l'ajustement de la méthode du propagateur au problème d'estimation d'une base de l'espace d'observabilité et des algorithmes récursifs qui en découlent, il est essentiel de prouver qu'il est possible d'utiliser un tel outil de traitement d'antennes en identification des sous-espaces. Une comparaison des approches utilisées au sein de ces deux axes de recherche est donc nécessaire.

Ce chapitre est organisé comme suit :

- Après une étude de l'analogie entre le problème d'estimation de la matrice d'observabilité en identification récursive des sous-espaces et la poursuite du sous-espace signal en traitement d'antennes, les deux étapes nécessaires à l'identification des matrices d'état du système étudié sont mises en évidence au sein de la section 5.2.
- La première de ces étapes est plus particulièrement analysée dans le paragraphe 5.3.

¹Rappelons qu'il est possible d'extraire de la matrice d'observabilité des estimées fiables des matrices d'état du système par des approches classiques de type moindres carrés (cf. §3.3.4).

Quatre algorithmes² d'estimation du vecteur d'observation sont proposés, deux dans un contexte déterministe, deux dans un contexte stochastique. Deux outils mathématiques spécifiques seront utilisés : le lemme d'inversion matricielle et la décomposition QR [Rot01].

- Le problème de mise à jour d'une base de l'espace d'observabilité est quant à lui étudié dans le paragraphe 5.4. Sept algorithmes³, fondés sur l'adaptation de la méthode du propagateur [MD91] au problème d'estimation récursive de la matrice d'observabilité du procédé, sont développés. Une analyse de quelques propriétés asymptotiques de ces derniers conclut cette section.
- Une étude de cas mettant en évidence la complexité numérique des techniques d'identification récursive présentées dans ce manuscrit est proposée dans le paragraphe 5.5.
- Les principales méthodes d'estimation récursive des matrices d'état sont décrites au sein de la section 5.6.

5.2 Analogie entre l'identification des sous-espaces et le traitement d'antennes

En identification récursive des sous-espaces, l'objectif que nous nous imposons est de pouvoir estimer en ligne les paramètres des matrices d'état du système suivant :

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\tilde{\mathbf{u}}(t) \quad (5.1a)$$

$$\tilde{\mathbf{y}}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\tilde{\mathbf{u}}(t) \quad (5.1b)$$

en supposant que ce dernier puisse être perturbé par des bruits d'état et de mesure en entrée et en sortie :

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\tilde{\mathbf{u}}(t) + \mathbf{w}(t) \quad (5.2a)$$

$$\tilde{\mathbf{y}}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\tilde{\mathbf{u}}(t) \quad (5.2b)$$

$$\mathbf{y}(t) = \tilde{\mathbf{y}}(t) + \mathbf{v}(t) \quad (5.2c)$$

$$\mathbf{u}(t) = \tilde{\mathbf{u}}(t) + \mathbf{v}(t). \quad (5.2d)$$

En poursuite de sous-espaces pour le traitement d'antennes, le problème considéré consiste à estimer en ligne les positions des sources émettrices en suivant les directions d'arrivée $\boldsymbol{\theta}$ déduites de la matrice des directions $\boldsymbol{\Gamma}(\boldsymbol{\theta})$:

$$\mathbf{z}(t) = \boldsymbol{\Gamma}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t). \quad (5.3)$$

Lorsque'on étudie en détail les techniques développées dans ces deux thématiques [Vib95, KV96], il est remarquable de constater que l'objectif mathématique final est le même : suivre les valeurs et vecteurs propres ou singuliers de matrices particulières en adaptant, à chaque nouvelle acquisition, un sous-espace spécifique aux dernières observations. Dans chaque discipline, des outils algébriques particuliers ont été développés et appliqués pour fournir des solutions consistantes aux problèmes posés. Notre approche consiste à adapter certains algorithmes de poursuite des sous-espaces à notre problème d'identification.

²ROM_{im}, RPIPO_{im}, ROM_{qr} et RPIPO_{qr}.

³RPM1, RPM2, COPM, COIVPM, IVPM, EIVPM et EIV_{sqr}tPM.

Lorsqu'on compare directement la forme des représentations (5.2) et (5.3), l'analogie n'est pas spécialement évidente. Pourtant, lorsqu'on analyse les différentes variables du système différentiel (5.2) et de l'équation (5.3), on peut constater que certaines d'entre elles partagent des caractéristiques communes. Ceci est particulièrement flagrant lorsqu'on considère le vecteur d'état \mathbf{x} et le vecteur source \mathbf{s} . Ils regroupent en effet un ensemble de variables, non mesurables, considérées, la plupart du temps, comme des variables aléatoires [Jan97] contenant l'information nécessaire et suffisante sur le passé du système étudié pour déterminer son évolution ultérieure. Ces premières constatations montrent donc qu'il existe un véritable lien entre ces deux représentations. Puisque nous désirons adapter, à notre problème d'estimation récursive, certains algorithmes de traitement d'antennes développés pour le suivi des caractéristiques du processus émetteur, il serait intéressant de pouvoir récrire le système différentiel (5.2) sous une forme équivalente à la relation (5.3). Cette dernière équation contient un vecteur d'observation \mathbf{z} composé de réponses spatialement décalées correspondant aux signaux mesurés au niveau des n_z capteurs de l'antenne considérée. En identification, nous n'avons accès qu'à des données temporellement décalées. L'idée qui vient alors naturellement à l'esprit consiste à introduire, en identification, une fenêtre temporelle similaire à la fenêtre spatiale des n_z capteurs de l'antenne. Considérons donc les vecteurs d'entrées et de sorties décalées suivants :

$$\mathbf{y}_f^+(t) = \begin{bmatrix} \mathbf{y}(t) \\ \vdots \\ \mathbf{y}(t+f-1) \end{bmatrix} \quad \text{et} \quad \mathbf{u}_f^+(t) = \begin{bmatrix} \mathbf{u}(t) \\ \vdots \\ \mathbf{u}(t+f-1) \end{bmatrix} \quad (5.4)$$

acquis sur une fenêtre temporelle de taille f . Nous savons que ces derniers vérifient la relation suivante (cf. §3.3.2) :

$$\mathbf{y}_f^+(t) = \mathbf{\Gamma}_f \mathbf{x}(t) + \mathbf{H}_f \mathbf{u}_f^+(t) + \underbrace{\mathbf{G}_f \mathbf{w}_f^+(t) - \mathbf{H}_f \mathbf{v}_f^+(t) + \mathbf{v}_f^+(t)}_{\mathbf{n}_f^+(t)}, \quad (5.5)$$

les vecteurs de bruit décalés étant construits de la même manière que \mathbf{y}_f^+ et \mathbf{u}_f^+ . L'analogie de notation entre les équations (5.3) et (5.5) est alors manifeste en posant⁴ :

$$\mathbf{z}_f^+(t) = \mathbf{y}_f^+(t) - \mathbf{H}_f \mathbf{u}_f^+(t) \quad (5.6)$$

puisque :

$$\mathbf{z}_f^+(t) = \mathbf{\Gamma}_f \mathbf{x}(t) + \mathbf{n}_f^+(t). \quad (5.7)$$

Cette similitude (cf. tab. 5.1) permet ainsi de décomposer le problème d'identification en deux étapes⁵ :

1. la mise à jour du vecteur \mathbf{z}_f^+ , nommé vecteur d'observation dans la suite de ce manuscrit, à partir des nouvelles données d'entrée-sortie u et y accessibles à l'instant t :

$$\mathbf{z}_f^+(t) = \mathbf{y}_f^+(t) - \mathbf{H}_f \mathbf{u}_f^+(t); \quad (5.8)$$

2. l'estimation d'une base de la matrice d'observabilité, à partir de ce vecteur, en adaptant des méthodes particulières de poursuite des sous-espaces.

| Identification | DDA |
|---|--|
| $\mathbf{z}_f^+(t) \in \mathbb{R}^{n_y f \times 1}$ | $\mathbf{z}(t) \in \mathbb{C}^{n_z \times 1}$ |
| $\mathbf{\Gamma}_f \in \mathbb{R}^{n_y f \times n_x}$ | $\mathbf{\Gamma}(\boldsymbol{\theta}) \in \mathbb{C}^{n_z \times n_s}$ |
| $\mathbf{x}(t) \in \mathbb{R}^{n_x \times 1}$ | $\mathbf{s}(t) \in \mathbb{C}^{n_s \times 1}$ |
| $\mathbf{n}_f^+(t) \in \mathbb{R}^{n_y f \times 1}$ | $\mathbf{n}(t) \in \mathbb{C}^{n_z \times 1}$ |

TAB. 5.1: Analogie des notations utilisées en identification des sous-espaces et en traitement d'antennes.

Ces deux phases seront respectivement traitées au sein des sections 5.3 et 5.4.

Remarque 5.1. *Puisque le vecteur d'état est, la plupart du temps, non accessible à la mesure, une estimation directe de $\mathbf{\Gamma}_f$ est impossible. Il est donc tout à fait sensé de se restreindre à estimer une base de l'espace d'observabilité, l'identification étant en effet réalisée dans une base arbitraire. Ce résultat est la conséquence de la pluralité de la représentation d'état [Bro91]. Il sera toutefois nécessaire de vérifier que l'estimation est constamment effectuée dans la même base.*

Il est important de préciser que la plupart des techniques utilisées en traitement d'antennes pour la détection des directions d'arrivée fixent des hypothèses simplificatrices sur le type de bruit agissant sur l'antenne. Ainsi, MUSIC [BK83, Sch86], ESPRIT [RPK86, RK89] ou PAST [Yan95b, Yan96] imposent que la matrice de covariance du bruit soit proportionnelle à l'identité, SWEDE [ESS94] qu'elle soit bloc diagonale. PM [MD91] ne déroge pas à cette règle. Or, dans notre problème d'identification, la matrice de bruit \mathbf{N}_f^+ est, dans le pire des cas, construite comme suit :

$$\mathbf{N}_f^+ = \mathbf{G}_f \mathbf{W}_f^+ - \mathbf{H}_f \mathbf{\Upsilon}_f^+ + \mathbf{V}_f^+ \quad (5.9)$$

où \mathbf{G}_f et \mathbf{H}_f sont des matrices de Toeplitz par blocs. La matrice de covariance des perturbations n'est alors pas proportionnelle à l'identité. Il est donc nécessaire d'inclure une étape de traitement des perturbations afin de fournir des estimations consistantes quel que soit le type de bruit agissant sur le procédé. Ce traitement peut être réalisé soit dans la phase de mise à jour du vecteur d'observation, soit durant l'étape d'estimation de la matrice d'observabilité étendue. Nous verrons plus précisément que, dans chaque situation, l'introduction d'une variable instrumentale [War77, SS83a] permet de supprimer ces perturbations.

5.3 Mise à jour du vecteur d'observation

Comme nous l'avons montré dans le paragraphe précédent, la première étape de notre problème d'identification récursive consiste à déterminer, à chaque nouvelle acquisition, le vecteur d'observation \mathbf{z}_f^+ . Lorsqu'on analyse plus précisément l'équation (5.8), il est facile de

⁴Cette analyse justifie *a posteriori* la représentation proposée par T. Gustafsson [Gus98] (cf. §4.3.1).

⁵Rappelons que la matrice de Toeplitz \mathbf{H}_f est inconnue à l'instant t .

constater que l'unique inconnue de cette relation est la matrice de Toeplitz \mathbf{H}_f . Deux solutions peuvent alors être envisagées pour évaluer \mathbf{z}_f^+ :

- construire la matrice \mathbf{H}_f à partir des informations et estimées accessibles à l'instant t et calculer explicitement le résultat de la soustraction $\mathbf{z}_f^+ = \mathbf{y}_f^+ - \mathbf{H}_f \mathbf{u}_f^+$;
- estimer directement le vecteur d'observation \mathbf{z}_f^+ , sans avoir recours à la matrice \mathbf{H}_f , à l'aide des dernières acquisitions et des connaissances sur le procédé acquises depuis le début de la procédure d'identification.

La première solution est de loin la plus facile à mettre en œuvre. Elle repose simplement sur la définition de la matrice \mathbf{H}_f :

$$\mathbf{H}_f = \begin{bmatrix} \mathbf{D} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{CB} & \mathbf{D} & \cdots & \mathbf{0} \\ \mathbf{CAB} & \mathbf{CB} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{CA}^{f-2}\mathbf{B} & \mathbf{CA}^{f-3}\mathbf{B} & \cdots & \mathbf{D} \end{bmatrix} \quad (5.10)$$

qui indique que cette dernière est directement calculable à partir des matrices d'état du système. Or, au début de l'itération t de la procédure récursive d'identification, nous n'avons accès qu'aux estimées de l'instant $t - 1$. Il est donc inévitable d'approximer $\mathbf{H}_f(t)$ par une matrice de Toeplitz construite à partir des matrices d'état estimées à l'instant $t - 1$. Sa construction s'appuie sur le constat suivant [Gus97a] : le premier bloc colonne de \mathbf{H}_f vaut :

$$\begin{bmatrix} \mathbf{D} \\ \mathbf{\Gamma}_f(1 : n_y(f-1), :)\mathbf{B} \end{bmatrix}, \quad (5.11)$$

les autres blocs colonnes pouvant s'exprimer comme une troncature de ce dernier. Le vecteur d'observation utilisé s'écrit alors :

$$\mathbf{z}_f^+(t) \approx \mathbf{y}_f^+(t) - \hat{\mathbf{H}}_f(t-1)\mathbf{u}_f^+(t) \quad (5.12)$$

où $\hat{\mathbf{H}}_f(t-1)$ est calculé en remplaçant, dans l'équation (5.11), les matrices considérées par leur estimée.

La seconde solution, moins approximative, consiste à estimer, à chaque nouvelle acquisition, le vecteur d'observation sans recourir explicitement à la matrice \mathbf{H}_f . La méthode proposée est fondée sur l'adaptation, au cas récursif, d'un problème similaire rencontré en identification hors ligne des sous espaces : éliminer de l'espace $\text{Im}_{\text{col}}(\mathbf{Y}_f^+)$ le terme relatif à $\text{Im}_{\text{col}}(\mathbf{H}_f)$. Afin de bien comprendre la démarche suivie dans la suite de cette section, considérons la matrice de Hankel des observations \mathbf{Z}_f^+ définie par :

$$\mathbf{Z}_f^+(\bar{t}) = [\mathbf{z}_f^+(t) \quad \mathbf{z}_f^+(t+1) \quad \cdots \quad \mathbf{z}_f^+(\bar{t})]. \quad (5.13)$$

Il est alors facile de vérifier que :

$$\mathbf{Z}_f^+(\bar{t}) = \mathbf{Y}_f^+(\bar{t}) - \mathbf{H}_f \mathbf{U}_f^+(\bar{t}). \quad (5.14)$$

En effet :

$$\begin{aligned}\mathbf{Z}_f^+(\bar{t}) &= [\mathbf{z}_f^+(t) \quad \mathbf{z}_f^+(t+1) \quad \cdots \quad \mathbf{z}_f^+(\bar{t})] = [\mathbf{y}_f^+(t) - \mathbf{H}_f \mathbf{u}_f^+(t) \quad \cdots \quad \mathbf{y}_f^+(\bar{t}) - \mathbf{H}_f \mathbf{u}_f^+(\bar{t})] \\ &= [\mathbf{y}_f^+(t) \quad \cdots \quad \mathbf{y}_f^+(\bar{t})] - \mathbf{H}_f [\mathbf{u}_f^+(t) \quad \cdots \quad \mathbf{u}_f^+(\bar{t})] = \mathbf{Y}_f^+(\bar{t}) - \mathbf{H}_f \mathbf{U}_f^+(\bar{t}).\end{aligned}\quad (5.15)$$

Ainsi :

$$\begin{aligned}\mathbf{Z}_f^+(\bar{t}) &= \mathbf{Y}_f^+(\bar{t}) - \mathbf{H}_f \mathbf{U}_f^+(\bar{t}) \\ &= [\mathbf{Y}_f^+(\bar{t}-1) \quad \mathbf{y}_f^+(\bar{t})] - \mathbf{H}_f [\mathbf{U}_f^+(\bar{t}-1) \quad \mathbf{u}_f^+(\bar{t})] \\ &= [\mathbf{Y}_f^+(\bar{t}-1) - \mathbf{H}_f \mathbf{U}_f^+(\bar{t}-1) \quad \mathbf{y}_f^+(\bar{t}) - \mathbf{H}_f \mathbf{u}_f^+(\bar{t})] \\ &= [\mathbf{Z}_f^+(\bar{t}-1) \quad \mathbf{z}_f^+(\bar{t})].\end{aligned}\quad (5.16)$$

La mise à jour de la matrice de Hankel \mathbf{Z}_f^+ permet donc d'avoir accès au vecteur d'observation. Or, en identification hors ligne, la matrice d'observation \mathbf{Z}_f^+ peut être calculée de différentes manières selon les hypothèses fixées sur les signaux agissant sur le procédé. En effet :

– Dans le cas déterministe (cf. §3.3.1) :

$$\mathbf{Z}_f^+ = \mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \quad (5.17)$$

avec :

$$\mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} = \mathbf{I}_{n_{uf}} - \mathbf{U}_f^{+T} (\mathbf{U}_f^+ \mathbf{U}_f^{+T})^{-1} \mathbf{U}_f^+ \quad (5.18)$$

ou :

$$\mathbf{Z}_f^+ = \bar{\mathbf{R}}_{22} \bar{\mathbf{Q}}_2 \quad (5.19)$$

avec :

$$\begin{bmatrix} \mathbf{U}_f^+ \\ \mathbf{Y}_f^+ \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_{11} & \mathbf{0} \\ \bar{\mathbf{R}}_{21} & \bar{\mathbf{R}}_{22} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1 \\ \bar{\mathbf{Q}}_2 \end{bmatrix}. \quad (5.20)$$

– Dans le cas stochastique (cf. §3.3.2) :

$$\mathbf{Z}_f^+ = \mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \mathbf{\Xi}^T \left(\mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \mathbf{\Xi}^T \right)^{-1} \mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \quad (5.21)$$

ou

$$\mathbf{Z}_f^+ = \bar{\mathbf{R}}_{32} \bar{\mathbf{Q}}_2 \quad (5.22)$$

avec :

$$\begin{bmatrix} \mathbf{U}_f^+ \\ \mathbf{\Xi} \\ \mathbf{Y}_f^+ \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_{11} & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{21} & \bar{\mathbf{R}}_{22} & \mathbf{0} \\ \bar{\mathbf{R}}_{31} & \bar{\mathbf{R}}_{32} & \bar{\mathbf{R}}_{33} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1 \\ \bar{\mathbf{Q}}_2 \\ \bar{\mathbf{Q}}_3 \end{bmatrix}. \quad (5.23)$$

En associant ces relations à la constatation (5.16), plusieurs algorithmes récursifs peuvent être envisagés pour estimer le vecteur d'observation à chaque nouvelle acquisition. Quatre d'entre eux vont être présentés dans la suite de cette section. Deux cas seront plus particulièrement étudiés :

– une mise à jour du vecteur d'observation dans un contexte déterministe ;

– une mise à jour du vecteur d'observation dans un contexte stochastique.

Chacune de ces situations conduira à présenter deux algorithmes récursifs d'estimation respectivement fondés sur les relations (5.17) et (5.19) dans le cas déterministe et (5.21) et (5.22) dans le cas stochastique. Deux d'entre eux (cf. equ. (5.17) et (5.21)) sont fondés sur le lemme d'inversion matricielle [Rot01] et s'inspirent par là même des méthodes développées par H. Oku [OK99]. Ils diffèrent cependant par un choix simplifié des matrices utilisées pour la récursion puisque nous proposons d'utiliser directement les matrices (5.17) et (5.21) et non leur carré. Les deux suivants reposent sur la mise à jour des décompositions QR (5.19) et (5.22) à l'aide de rotations de Givens. Ces techniques ont déjà été exposées au sein du chapitre précédent (cf. §4.3.2). Notre objectif n'est donc pas de les redévelopper mais de prouver qu'elles permettent d'avoir accès au vecteur d'observation.

5.3.1 Dans un contexte déterministe

5.3.1.1 Mise à jour de $\mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}}$

L'algorithme développé dans ce paragraphe est fondé sur la définition de la projection orthogonale présente dans l'équation (5.17) :

$$\mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} = \mathbf{I} - \mathbf{U}_f^{+T} \left(\mathbf{U}_f^+ \mathbf{U}_f^{+T} \right)^{-1} \mathbf{U}_f^+ \quad (5.24)$$

et l'utilisation d'un outil mathématique courant en identification récursive : le lemme d'inversion matricielle. En effet, la relation (5.24) contient une inversion matricielle que nous voulons mettre à jour lorsque de nouvelles données sont mesurées. Or, en appliquant astucieusement le lemme d'inversion matricielle, il est possible de mettre explicitement en évidence l'apport des informations fournies par les dernières acquisitions. Considérons donc qu'un nouveau couple de données d'entrée-sortie soit accessible à l'instant $\bar{t} = t + M - 1$. Les matrices de Hankel d'entrée et de sortie sont alors modifiées de la façon suivante⁶ :

$$\mathbf{Y}_f^+(\bar{t}) = \begin{bmatrix} \sqrt{\lambda} \mathbf{Y}_f^+(\bar{t}-1) & \mathbf{y}_f^+(\bar{t}) \end{bmatrix} \quad (5.25)$$

$$\mathbf{U}_f^+(\bar{t}) = \begin{bmatrix} \sqrt{\lambda} \mathbf{U}_f^+(\bar{t}-1) & \mathbf{u}_f^+(\bar{t}) \end{bmatrix}. \quad (5.26)$$

Nous désirons mettre à jour directement la relation suivante :

$$\mathbf{Y}_f^+(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^+(\bar{t})^\perp} = \mathbf{Y}_f^+(\bar{t}) \left\{ \mathbf{I} - \mathbf{U}_f^+(\bar{t})^T \left(\mathbf{U}_f^+(\bar{t}) \mathbf{U}_f^{+T}(\bar{t}) \right)^{-1} \mathbf{U}_f^+(\bar{t}) \right\} \quad (5.27)$$

afin d'estimer le vecteur d'observation \mathbf{z}_f^+ . Pour cela, calculons cette relation étape par étape. Dans un premier temps, à partir de l'équation (5.26) :

$$\begin{aligned} \mathbf{U}_f^+(\bar{t}) \mathbf{U}_f^{+T}(\bar{t}) &= \begin{bmatrix} \sqrt{\lambda} \mathbf{U}_f^+(\bar{t}-1) & \mathbf{u}_f^+(\bar{t}) \end{bmatrix} \begin{bmatrix} \sqrt{\lambda} \mathbf{U}_f^+(\bar{t}-1) & \mathbf{u}_f^+(\bar{t}) \end{bmatrix}^T \\ &= \lambda \mathbf{U}_f^+(\bar{t}-1) \mathbf{U}_f^{+T}(\bar{t}-1) + \mathbf{u}_f^+(\bar{t}) \mathbf{u}_f^{+T}(\bar{t}). \end{aligned} \quad (5.28)$$

⁶ λ est un facteur d'oubli facultatif permettant de pondérer l'effet des données passées.

En appliquant le lemme d'inversion matricielle⁷ à (5.28), nous obtenons :

$$\left(\mathbf{U}_f^+(\bar{t})\mathbf{U}_f^{+T}(\bar{t})\right)^{-1} = \left[\left(\mathbf{U}_f^+(\bar{t}-1)\mathbf{U}_f^{+T}(\bar{t}-1)\right)^{-1} - \frac{\boldsymbol{\gamma}_f^+(\bar{t})\boldsymbol{\gamma}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\delta}_f(\bar{t})}\right] / \lambda \quad (5.29)$$

en posant :

$$\boldsymbol{\gamma}_f^+(\bar{t}) = \left(\mathbf{U}_f^+(\bar{t}-1)\mathbf{U}_f^{+T}(\bar{t}-1)\right)^{-1} \mathbf{u}_f^+(\bar{t}) \quad (5.30)$$

$$\boldsymbol{\delta}_f(\bar{t}) = \mathbf{u}_f^{+T}(\bar{t})\boldsymbol{\gamma}_f^+(\bar{t}). \quad (5.31)$$

En postmultipliant cette équation par $\mathbf{U}_f^+(\bar{t})$:

$$\begin{aligned} &\left(\mathbf{U}_f^+(\bar{t})\mathbf{U}_f^{+T}(\bar{t})\right)^{-1} \mathbf{U}_f^+(\bar{t}) = \\ &\left[\frac{1}{\lambda} \left(\mathbf{U}_f^+(\bar{t}-1)\mathbf{U}_f^{+T}(\bar{t}-1)\right)^{-1} - \frac{1}{\lambda} \frac{\boldsymbol{\gamma}_f^+(\bar{t})\boldsymbol{\gamma}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\delta}_f(\bar{t})}\right] \left[\sqrt{\lambda} \mathbf{U}_f^+(\bar{t}-1) \quad \mathbf{u}_f^+(\bar{t})\right] = \\ &\left[\frac{1}{\sqrt{\lambda}} \left(\mathbf{U}_f^+(\bar{t}-1)\mathbf{U}_f^{+T}(\bar{t}-1)\right)^{-1} \mathbf{U}_f^+(\bar{t}-1) - \frac{1}{\sqrt{\lambda}} \frac{\boldsymbol{\gamma}_f^+(\bar{t})\boldsymbol{\gamma}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\delta}_f(\bar{t})} \mathbf{U}_f^+(\bar{t}-1) \quad \frac{\boldsymbol{\gamma}_f^+(\bar{t})}{\lambda + \boldsymbol{\delta}_f(\bar{t})}\right]. \end{aligned} \quad (5.32)$$

En prémultipliant cette équation par $\mathbf{U}_f^{+T}(\bar{t})$:

$$\begin{aligned} &\mathbf{U}_f^{+T}(\bar{t}) \left(\mathbf{U}_f^+(\bar{t})\mathbf{U}_f^{+T}(\bar{t})\right)^{-1} \mathbf{U}_f^+(\bar{t}) = \\ &\left[\begin{array}{c} \sqrt{\lambda} \mathbf{U}_f^{+T}(\bar{t}-1) \\ \mathbf{u}_f^{+T}(\bar{t}) \end{array}\right] \left[\frac{1}{\sqrt{\lambda}} \left(\mathbf{U}_f^+(\bar{t}-1)\mathbf{U}_f^{+T}(\bar{t}-1)\right)^{-1} \mathbf{U}_f^+(\bar{t}-1) - \frac{1}{\sqrt{\lambda}} \frac{\boldsymbol{\gamma}_f^+(\bar{t})\boldsymbol{\gamma}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\delta}_f(\bar{t})} \mathbf{U}_f^+(\bar{t}-1) \quad \frac{\boldsymbol{\gamma}_f^+(\bar{t})}{\lambda + \boldsymbol{\delta}_f(\bar{t})}\right] = \\ &\left[\begin{array}{cc} \mathbf{U}_f^{+T}(\bar{t}-1) \left(\mathbf{U}_f^+(\bar{t}-1)\mathbf{U}_f^{+T}(\bar{t}-1)\right)^{-1} \mathbf{U}_f^+(\bar{t}-1) - \mathbf{U}_f^{+T}(\bar{t}-1) \frac{\boldsymbol{\gamma}_f^+(\bar{t})\boldsymbol{\gamma}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\delta}_f(\bar{t})} \mathbf{U}_f^+(\bar{t}-1) & \sqrt{\lambda} \mathbf{U}_f^{+T}(\bar{t}-1) \frac{\boldsymbol{\gamma}_f^+(\bar{t})}{\lambda + \boldsymbol{\delta}_f(\bar{t})} \\ \sqrt{\lambda} \frac{\boldsymbol{\gamma}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\delta}_f(\bar{t})} \mathbf{U}_f^+(\bar{t}-1) & \frac{\boldsymbol{\delta}_f(\bar{t})}{\lambda + \boldsymbol{\delta}_f(\bar{t})} \end{array}\right]. \end{aligned} \quad (5.33)$$

La projection orthogonale sur le noyau de $\mathbf{U}_f^+(\bar{t})$ vaut alors :

$$\begin{aligned} \Pi_{\mathbf{U}_f^{+}(\bar{t})} &= \mathbf{I}_M - \mathbf{U}_f^{+T}(\bar{t}) \left(\mathbf{U}_f^+(\bar{t})\mathbf{U}_f^{+T}(\bar{t})\right)^{-1} \mathbf{U}_f^+(\bar{t}) = \\ &\left[\begin{array}{cc} \mathbf{I}_{M-1} - \mathbf{U}_f^{+T}(\bar{t}-1) \left(\mathbf{U}_f^+(\bar{t}-1)\mathbf{U}_f^{+T}(\bar{t}-1)\right)^{-1} \mathbf{U}_f^+(\bar{t}-1) + \mathbf{U}_f^{+T}(\bar{t}-1) \frac{\boldsymbol{\gamma}_f^+(\bar{t})\boldsymbol{\gamma}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\delta}_f(\bar{t})} \mathbf{U}_f^+(\bar{t}-1) & -\sqrt{\lambda} \mathbf{U}_f^{+T}(\bar{t}-1) \frac{\boldsymbol{\gamma}_f^+(\bar{t})}{\lambda + \boldsymbol{\delta}_f(\bar{t})} \\ -\sqrt{\lambda} \frac{\boldsymbol{\gamma}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\delta}_f(\bar{t})} \mathbf{U}_f^+(\bar{t}-1) & \frac{\lambda}{\lambda + \boldsymbol{\delta}_f(\bar{t})} \end{array}\right] = \\ &\left[\begin{array}{cc} \Pi_{\mathbf{U}_f^{+}(\bar{t}-1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{array}\right] + \frac{1}{\lambda + \boldsymbol{\delta}_f(\bar{t})} \left[\begin{array}{cc} \mathbf{U}_f^{+T}(\bar{t}-1) \boldsymbol{\gamma}_f^+(\bar{t}) \boldsymbol{\gamma}_f^{+T}(\bar{t}) \mathbf{U}_f^+(\bar{t}-1) & -\sqrt{\lambda} \mathbf{U}_f^{+T}(\bar{t}-1) \boldsymbol{\gamma}_f^+(\bar{t}) \\ -\sqrt{\lambda} \boldsymbol{\gamma}_f^{+T}(\bar{t}) \mathbf{U}_f^+(\bar{t}-1) & \lambda \end{array}\right]. \end{aligned} \quad (5.34)$$

⁷ $(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{D} \mathbf{A}^{-1} \mathbf{B} + \mathbf{C}^{-1})^{-1} \mathbf{D} \mathbf{A}^{-1}$ [Rot01].

Finalement, nous obtenons :

$$\begin{aligned}
\mathbf{Y}_f^+(\bar{t})\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})} &= \left[\sqrt{\lambda}\mathbf{Y}_f^+(\bar{t}-1) \quad \mathbf{y}_f^+(\bar{t}) \right] \mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})} = \left[\sqrt{\lambda}\mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} \quad \mathbf{0} \right] \\
&+ \frac{1}{\lambda + \delta_f(\bar{t})} \left[\sqrt{\lambda}\mathbf{Y}_f^+(\bar{t}-1)\mathbf{U}_f^{\perp T}(\bar{t}-1)\boldsymbol{\gamma}_f^+(\bar{t})\boldsymbol{\gamma}_f^{+T}(\bar{t})\mathbf{U}_f^{\perp}(\bar{t}-1) - \sqrt{\lambda}\mathbf{y}_f^+(\bar{t})\boldsymbol{\gamma}_f^{+T}(\bar{t})\mathbf{U}_f^{\perp}(\bar{t}-1) \right. \\
&\quad \left. - \lambda\mathbf{Y}_f^+(\bar{t}-1)\mathbf{U}_f^{\perp T}(\bar{t}-1)\boldsymbol{\gamma}_f^+(\bar{t}) + \lambda\mathbf{y}_f^+(\bar{t}) \right] \\
&= \left[\sqrt{\lambda}\mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} - \frac{1}{\sqrt{\lambda + \delta_f(\bar{t})}}\check{\mathbf{z}}_f^+(\bar{t})\boldsymbol{\gamma}_f^{+T}(\bar{t})\mathbf{U}_f^{\perp}(\bar{t}-1) \quad \sqrt{\frac{\lambda}{\lambda + \delta_f(\bar{t})}}\check{\mathbf{z}}_f^+(\bar{t}) \right]
\end{aligned} \tag{5.35}$$

avec :

$$\check{\mathbf{z}}_f^+(\bar{t}) = \sqrt{\frac{\lambda}{\lambda + \delta_f(\bar{t})}} \left(\mathbf{y}_f^+(\bar{t}) - \mathbf{Y}_f^+(\bar{t}-1)\mathbf{U}_f^{\perp T}(\bar{t}-1)\boldsymbol{\gamma}_f^+(\bar{t}) \right). \tag{5.36}$$

L'étape fondamentale de cette démonstration consiste à prouver que cet algorithme récursif donne accès au vecteur d'observation \mathbf{z}_f^+ . La relation entre $\check{\mathbf{z}}_f^+(\bar{t})$ et $\mathbf{z}_f^+(\bar{t})$ est démontrée en considérant le produit $\mathbf{Z}_f^+(\bar{t})\mathbf{Z}_f^{+T}(\bar{t})$. En effet, puisque :

$$\mathbf{Z}_f^+(\bar{t}) = \mathbf{Y}_f^+(\bar{t})\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})}, \tag{5.37}$$

nous avons⁸ :

$$\mathbf{Z}_f^+(\bar{t})\mathbf{Z}_f^{+T}(\bar{t}) = \mathbf{Y}_f^+(\bar{t})\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})} \left(\mathbf{Y}_f^+(\bar{t})\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})} \right)^T = \mathbf{Y}_f^+(\bar{t})\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})} \mathbf{Y}_f^{+T}(\bar{t}). \tag{5.38}$$

Or, en associant les équations (5.25) et (5.35), il est facile de montrer que :

$$\mathbf{Y}_f^+(\bar{t})\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})} \mathbf{Y}_f^{+T}(\bar{t}) = \lambda\mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} \mathbf{Y}_f^{+T}(\bar{t}-1) + \check{\mathbf{z}}_f^+(\bar{t})\check{\mathbf{z}}_f^{+T}(\bar{t}). \tag{5.39}$$

Ainsi :

$$\mathbf{Z}_f^+(\bar{t})\mathbf{Z}_f^{+T}(\bar{t}) = \lambda\mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} \mathbf{Y}_f^{+T}(\bar{t}-1) + \check{\mathbf{z}}_f^+(\bar{t})\check{\mathbf{z}}_f^{+T}(\bar{t}). \tag{5.40}$$

De plus, nous savons que :

$$\begin{aligned}
\mathbf{Z}_f^+(\bar{t})\mathbf{Z}_f^{+T}(\bar{t}) &= \\
&\left[\sqrt{\lambda}\mathbf{Z}_f^+(\bar{t}-1) \quad \mathbf{z}_f^+(\bar{t}) \right] \begin{bmatrix} \sqrt{\lambda}\mathbf{Z}_f^{+T}(\bar{t}-1) \\ \mathbf{z}_f^{+T}(\bar{t}) \end{bmatrix} = \lambda\mathbf{Z}_f^+(\bar{t}-1)\mathbf{Z}_f^{+T}(\bar{t}-1) + \mathbf{z}_f^+(\bar{t})\mathbf{z}_f^{+T}(\bar{t}).
\end{aligned} \tag{5.41}$$

Puisque $\mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} \mathbf{Y}_f^{+T}(\bar{t}-1) = \mathbf{Z}_f^+(\bar{t}-1)\mathbf{Z}_f^{+T}(\bar{t}-1)$, la relation suivante devient évidente :

$$\check{\mathbf{z}}_f^+(\bar{t}) = \pm\mathbf{z}_f^+(\bar{t}). \tag{5.42}$$

⁸ $\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})} \left(\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})} \right)^T = \mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})}$.

La mise à jour du vecteur d'observation est ainsi réalisable en appliquant l'algorithme ROM_{im} récursif suivant :

$$\boldsymbol{\gamma}_f^+(\bar{t}) = \left(\mathbf{U}_f^+(\bar{t}-1) \mathbf{U}_f^{+T}(\bar{t}-1) \right)^{-1} \mathbf{u}_f^+(\bar{t}) \quad (5.43a)$$

$$\boldsymbol{\delta}_f(\bar{t}) = \mathbf{u}_f^{+T}(\bar{t}) \boldsymbol{\gamma}_f^+(\bar{t}) \quad (5.43b)$$

$$\mathbf{z}_f^+(\bar{t}) = \sqrt{\frac{\lambda}{\lambda + \boldsymbol{\delta}_f(\bar{t})}} \left(\mathbf{y}_f^+(\bar{t}) - \mathbf{Y}_f^+(\bar{t}-1) \mathbf{U}_f^{+T}(\bar{t}-1) \boldsymbol{\gamma}_f^+(\bar{t}) \right) \quad (5.43c)$$

$$\mathbf{Y}_f^+(\bar{t}) \mathbf{U}_f^{+T}(\bar{t}) = \lambda \mathbf{Y}_f^+(\bar{t}-1) \mathbf{U}_f^{+T}(\bar{t}-1) + \mathbf{y}_f^+(\bar{t}) \mathbf{u}_f^{+T}(\bar{t}) \quad (5.43d)$$

$$\left(\mathbf{U}_f^+(\bar{t}) \mathbf{U}_f^{+T}(\bar{t}) \right)^{-1} = \left(\left(\mathbf{U}_f^+(\bar{t}-1) \mathbf{U}_f^{+T}(\bar{t}-1) \right)^{-1} - \frac{\boldsymbol{\gamma}_f^+(\bar{t}) \boldsymbol{\gamma}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\delta}_f(\bar{t})} \right) / \lambda. \quad (5.43e)$$

Remarque 5.2. *Il est fondamental de constater que les matrices et vecteurs composant l'algorithme (5.43) ne voient en aucun cas leur dimension évoluer au cours de la récursion. Cette constatation confond les affirmations de H. Oku [OK99] selon lequel l'utilisation de matrices non symétriques n'est pas appropriée à la récursion à cause d'une taille croissante en les acquisitions. Elle légitime ainsi a posteriori notre choix de mettre à jour $\mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+}}$.*

5.3.1.2 Mise à jour de $\bar{\mathbf{R}}_{22}$

Au sein de la section 4.3.2, nous avons vu qu'il était possible de mettre à jour, à chaque nouvelle acquisition, la décomposition QR du schéma d'identification Ordinary MOESP de la façon suivante :

$$\left[\sqrt{\lambda} \begin{bmatrix} \bar{\mathbf{R}}_{11}(\bar{t}-1) & \mathbf{0} \\ \bar{\mathbf{R}}_{21}(\bar{t}-1) & \bar{\mathbf{R}}_{22}(\bar{t}-1) \end{bmatrix} \begin{bmatrix} \mathbf{u}_f^+(\bar{t}) \\ \mathbf{y}_f^+(\bar{t}) \end{bmatrix} \right] \mathbf{Rot}_G(\bar{t}) = \left[\begin{bmatrix} \bar{\mathbf{R}}_{11}(\bar{t}) & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{21}(\bar{t}) & \lambda \bar{\mathbf{R}}_{22}(\bar{t}-1) & \bar{\mathbf{z}}_f^+(\bar{t}) \end{bmatrix} \right] \quad (5.44)$$

à l'aide d'un ensemble de rotations de Givens \mathbf{Rot}_G . Cet algorithme, nommé ROM_{qr} dans la suite de ce manuscrit, permet de calculer un vecteur $\bar{\mathbf{z}}_f^+$ lié à la matrice d'observabilité du système par la relation suivante (cf. annexe A.3) :

$$\begin{aligned} \bar{\mathbf{z}}_f^+(\bar{t}) = & \mathbf{\Gamma}_f \left(\sqrt{\lambda} \mathbf{X}^+(\bar{t}-1) \bar{\mathbf{Q}}_1^T(\bar{t}-1) \mathbf{Rot}_G^{12}(\bar{t}) + \mathbf{x}(\bar{t}-1) \mathbf{Rot}_G^{22}(\bar{t}) \right) \\ & + \sqrt{\lambda} \mathbf{N}_f^+(\bar{t}-1) \bar{\mathbf{Q}}_1^T(\bar{t}-1) \mathbf{Rot}_G^{12}(\bar{t}) + \mathbf{n}_f^+(\bar{t}) \mathbf{Rot}_G^{22}(\bar{t}) \end{aligned} \quad (5.45)$$

avec [GLV98] :

$$\mathbf{Rot}_G(\bar{t}) = \begin{bmatrix} \mathbf{Rot}_G^{11}(\bar{t}) & \mathbf{0} & \mathbf{Rot}_G^{12}(\bar{t}) \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{Rot}_G^{21}(\bar{t}) & \mathbf{0} & \mathbf{Rot}_G^{22}(\bar{t}) \end{bmatrix}. \quad (5.46)$$

Cette équation, intéressante d'un point de vue théorique, ne montre malheureusement pas la relation liant $\bar{\mathbf{z}}_f^+$ au vecteur d'observation \mathbf{z}_f^+ . Or, l'objectif de cette section est de mettre à jour récursivement ce vecteur. Il semble donc nécessaire d'explicitier la relation liant $\bar{\mathbf{z}}_f^+$ et \mathbf{z}_f^+ . C'est ce que nous faisons dans la suite de ce paragraphe. Nous savons en effet que :

$$\left[\sqrt{\lambda} \bar{\mathbf{R}}_{22}(\bar{t}-1) \quad \bar{\mathbf{z}}_f^+(\bar{t}) \right] = \bar{\mathbf{R}}_{22}(\bar{t}) \quad (5.47)$$

donc :

$$\bar{\mathbf{R}}_{22}(\bar{t})\bar{\mathbf{R}}_{22}^T(\bar{t}) = \lambda \bar{\mathbf{R}}_{22}(\bar{t}-1)\bar{\mathbf{R}}_{22}^T(\bar{t}-1) + \bar{\mathbf{z}}_f^+(\bar{t}) \bar{\mathbf{z}}_f^{+T}(\bar{t}). \quad (5.48)$$

De plus, il est possible de montrer que [Vib95] :

$$\bar{\mathbf{R}}_{22}\bar{\mathbf{Q}}_2 = \mathbf{Y}_f^+ - \mathbf{H}_f^{MC}\mathbf{U}_f^+ \quad (5.49)$$

où \mathbf{H}_f^{MC} représente l'argument minimisant le critère $\left\| \mathbf{Y}_f^+ - \mathbf{H}_f \mathbf{U}_f^+ \right\|_F^2$. Ainsi, puisque $\bar{\mathbf{Q}}_2$ est une matrice orthogonale :

$$\bar{\mathbf{R}}_{22}(\bar{t})\bar{\mathbf{R}}_{22}^T(\bar{t}) = \left[\mathbf{Y}_f^+(\bar{t}) - \mathbf{H}_f^{MC}\mathbf{U}_f^+(\bar{t}) \right] \left[\mathbf{Y}_f^+(\bar{t}) - \mathbf{H}_f^{MC}\mathbf{U}_f^+(\bar{t}) \right]^T. \quad (5.50)$$

Or :

$$\mathbf{U}_f^+(\bar{t}) = \left[\sqrt{\lambda} \mathbf{U}_f^+(\bar{t}-1) \quad \mathbf{u}_f^+(\bar{t}) \right] \quad (5.51)$$

$$\mathbf{Y}_f^+(\bar{t}) = \left[\sqrt{\lambda} \mathbf{Y}_f^+(\bar{t}-1) \quad \mathbf{y}_f^+(\bar{t}) \right]. \quad (5.52)$$

En introduisant ces deux relations au sein de l'équation (5.50), il est facile de vérifier que :

$$\begin{aligned} \bar{\mathbf{R}}_{22}(\bar{t})\bar{\mathbf{R}}_{22}^T(\bar{t}) &= \lambda \left[\mathbf{Y}_f^+(\bar{t}-1) - \mathbf{H}_f^{MC}\mathbf{U}_f^+(\bar{t}-1) \right] \left[\mathbf{Y}_f^+(\bar{t}-1) - \mathbf{H}_f^{MC}\mathbf{U}_f^+(\bar{t}-1) \right]^T \\ &\quad + \left[\mathbf{y}_f^+(\bar{t}) - \mathbf{H}_f^{MC}\mathbf{u}_f^+(\bar{t}) \right] \left[\mathbf{y}_f^+(\bar{t}) - \mathbf{H}_f^{MC}\mathbf{u}_f^+(\bar{t}) \right]^T. \end{aligned} \quad (5.53)$$

Puisque $\mathbf{z}_f^+ = \mathbf{y}_f^+ - \mathbf{H}_f^{MC}\mathbf{u}_f^+$:

$$\bar{\mathbf{R}}_{22}(\bar{t})\bar{\mathbf{R}}_{22}^T(\bar{t}) = \lambda \bar{\mathbf{R}}_{22}(\bar{t}-1)\bar{\mathbf{R}}_{22}^T(\bar{t}-1) + \mathbf{z}_f^+(\bar{t}) \mathbf{z}_f^{+T}(\bar{t}). \quad (5.54)$$

En exploitant les équations (5.48) et (5.54), nous vérifions alors que :

$$\mathbf{z}_f^+(\bar{t}) = \pm \bar{\mathbf{z}}_f^+(\bar{t}). \quad (5.55)$$

Le vecteur $\bar{\mathbf{z}}_f^+$ est, au signe près, le vecteur d'observation recherché.

5.3.2 Dans un contexte stochastique

5.3.2.1 Mise à jour de $\mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^{+\perp}} \Xi^T \left(\Xi \Pi_{\mathbf{U}_f^{+\perp}} \Xi^T \right)^{-1} \Xi \Pi_{\mathbf{U}_f^{+\perp}}$

Lorsque les perturbations agissant sur le procédé sont inconnues et colorées, l'application d'une simple projection orthogonale n'est pas suffisante pour obtenir une estimée consistante de la matrice d'observabilité du système (cf. chapitre 3). Il est alors nécessaire de faire appel à une variable instrumentale dont l'action conjointe avec la projection orthogonale $\Pi_{\mathbf{U}_f^{+\perp}}$ permet de supprimer le régime forcé et l'effet des perturbations. Dans le cas particulier de la classe d'identification MOESP, cette opération cumulée a pour expression (cf. §3.4) :

$$\mathbf{Z}_f^+ = \mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^{+\perp}} \Xi^T \left(\Xi \Pi_{\mathbf{U}_f^{+\perp}} \Xi^T \right)^{-1} \Xi \Pi_{\mathbf{U}_f^{+\perp}}. \quad (5.56)$$

Notre objectif est de mettre à jour cette relation à chaque nouvelle acquisition afin de calculer un vecteur d'observation asymptotiquement épuré des perturbations. La démonstration que nous proposons fait de nouveau appel au lemme d'inversion matricielle. La démarche suivie est identique à celle proposée dans le paragraphe 5.3.1.1. Elle demande cependant un développement mathématique fastidieux dont la présentation au sein même de ce paragraphe n'apporterait au lecteur aucune aide pour la compréhension de l'algorithme. La preuve est donc présentée en annexe A.4. Elle permet d'énoncer la proposition suivante :

Proposition 5.1. *Supposons que les matrices $\mathbf{Y}_f^+ \mathbf{U}_f^{+T}$, $\mathbf{U}_f^+ \mathbf{U}_f^{+T}$, $\Xi \mathbf{U}_f^{+T}$, Φ_f et $\mathbf{Y}_f^+ \Pi_{\mathbf{U}_f^{+\perp}} \Xi^T$ aient été estimées à l'instant $\bar{t} - 1$. Lorsque de nouvelles données d'entrée-sortie sont accessibles à la mesure, la mise à jour du vecteur d'observation est réalisée en appliquant l'algorithme récursif $RPIPO_{im}$ suivant :*

$$\boldsymbol{\gamma}_f^+(\bar{t}) = \left(\mathbf{U}_f^+(\bar{t}-1) \mathbf{U}_f^{+T}(\bar{t}-1) \right)^{-1} \mathbf{u}_f^+(\bar{t}) \quad (5.57a)$$

$$\boldsymbol{\delta}_f(\bar{t}) = \mathbf{u}_f^{+T}(\bar{t}) \boldsymbol{\gamma}_f^+(\bar{t}) \quad (5.57b)$$

$$\check{\mathbf{z}}_f^+(\bar{t}) = \sqrt{\frac{\lambda}{\lambda + \boldsymbol{\delta}_f(\bar{t})}} \left(\mathbf{y}_f^+(\bar{t}) - \mathbf{Y}_f^+(\bar{t}-1) \mathbf{U}_f^{+T}(\bar{t}-1) \boldsymbol{\gamma}_f^+(\bar{t}) \right) \quad (5.57c)$$

$$\mathbf{Y}_f^+(\bar{t}) \mathbf{U}_f^{+T}(\bar{t}) = \lambda \mathbf{Y}_f^+(\bar{t}-1) \mathbf{U}_f^{+T}(\bar{t}-1) + \mathbf{y}_f^+(\bar{t}) \mathbf{u}_f^{+T}(\bar{t}) \quad (5.57d)$$

$$\left(\mathbf{U}_f^+(\bar{t}) \mathbf{U}_f^{+T}(\bar{t}) \right)^{-1} = \left(\left(\mathbf{U}_f^+(\bar{t}-1) \mathbf{U}_f^{+T}(\bar{t}-1) \right)^{-1} - \frac{\boldsymbol{\gamma}_f^+(\bar{t}) \boldsymbol{\gamma}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\delta}_f(\bar{t})} \right) / \lambda \quad (5.57e)$$

$$\check{\boldsymbol{\xi}}(\bar{t}) = \sqrt{\frac{\lambda}{\lambda + \boldsymbol{\delta}_f(\bar{t})}} \left(\boldsymbol{\xi}(\bar{t}) - \Xi(\bar{t}-1) \mathbf{U}_f^{+T}(\bar{t}-1) \boldsymbol{\gamma}_f^+(\bar{t}) \right) \quad (5.57f)$$

$$\boldsymbol{\chi}_f^+(\bar{t}) = \Phi_f(\bar{t}-1) \check{\boldsymbol{\xi}}(\bar{t}) \quad (5.57g)$$

$$\boldsymbol{\zeta}_f(\bar{t}) = \boldsymbol{\chi}_f^{+T}(\bar{t}) \check{\boldsymbol{\xi}}(\bar{t}) \quad (5.57h)$$

$$\check{\check{\mathbf{z}}}_f^+(\bar{t}) = \sqrt{\frac{\lambda}{\lambda + \boldsymbol{\zeta}_f(\bar{t})}} \left(\check{\mathbf{z}}_f^+(\bar{t}) - \mathbf{Y}_f^+(\bar{t}-1) \Pi_{\mathbf{U}_f^{+\perp}(\bar{t}-1)} \Xi^T(\bar{t}-1) \boldsymbol{\chi}_f^+(\bar{t}) \right) \quad (5.57i)$$

$$\check{\check{\check{\mathbf{z}}}}_f^+(\bar{t}) = \check{\mathbf{z}}_f^+(\bar{t}) - \sqrt{\frac{\lambda}{\lambda + \boldsymbol{\zeta}_f(\bar{t})}} \check{\check{\mathbf{z}}}_f^+(\bar{t}) \quad (5.57j)$$

$$\Xi(\bar{t}) \mathbf{U}_f^{+T}(\bar{t}) = \lambda \Xi(\bar{t}-1) \mathbf{U}_f^{+T}(\bar{t}-1) + \boldsymbol{\xi}(\bar{t}) \mathbf{u}_f^{+T}(\bar{t}) \quad (5.57k)$$

$$\Phi_f(\bar{t}) = \left(\Phi_f(\bar{t}-1) - \frac{\boldsymbol{\chi}_f^+(\bar{t}) \boldsymbol{\chi}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\zeta}_f(\bar{t})} \right) / \lambda \quad (5.57l)$$

$$\mathbf{Y}_f^+(\bar{t}) \Pi_{\mathbf{U}_f^{+\perp}(\bar{t})} \Xi^T(\bar{t}) = \lambda \mathbf{Y}_f^+(\bar{t}-1) \Pi_{\mathbf{U}_f^{+\perp}(\bar{t}-1)} \Xi^T(\bar{t}-1) + \check{\check{\mathbf{z}}}_f^+(\bar{t}) \check{\boldsymbol{\xi}}^T(\bar{t}). \quad (5.57m)$$

puisque :

$$\mathbf{z}_f^+(\bar{t}) \mathbf{z}_f^{+T}(\bar{t}) = \check{\mathbf{z}}_f^+(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t}) - \check{\check{\mathbf{z}}}_f^+(\bar{t}) \check{\check{\mathbf{z}}}_f^{+T}(\bar{t}). \quad (5.57n)$$

Remarque 5.3. *Il est important de remarquer que nous n'avons pas directement accès au vecteur d'observation mais à la dernière composante de sa matrice de covariance. Cette constatation est essentielle pour la phase d'estimation de la matrice d'observabilité du système. Il sera*

en effet nécessaire de développer des algorithmes utilisant la matrice $\mathbf{z}_f^+ \mathbf{z}_f^{+T}$ au lieu du vecteur \mathbf{z}_f^+ .

Remarque 5.4. Comme dans le cas déterministe, les vecteurs et matrices composant cet algorithme ne subissent aucune évolution de leur dimension au cours de la récursion.

5.3.2.2 Mise à jour de $\bar{\mathbf{R}}_{32}$

Comme dans le cas déterministe, il est possible, d'estimer, dans un cadre stochastique, la matrice des observations \mathbf{Z}_f^+ à l'aide d'une factorisation QR :

$$\mathbf{Z}_f^+ = \bar{\mathbf{R}}_{32} \bar{\mathbf{Q}}_2 \quad (5.58)$$

avec :

$$\begin{bmatrix} \mathbf{U}_f^+ \\ \bar{\boldsymbol{\Xi}} \\ \mathbf{Y}_f^+ \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_{11} & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{21} & \bar{\mathbf{R}}_{22} & \mathbf{0} \\ \bar{\mathbf{R}}_{31} & \bar{\mathbf{R}}_{32} & \bar{\mathbf{R}}_{33} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1 \\ \bar{\mathbf{Q}}_2 \\ \bar{\mathbf{Q}}_3 \end{bmatrix}. \quad (5.59)$$

La mise à jour de cette relation à l'aide de rotations de Givens a déjà été exposée dans le chapitre précédent (cf. §4.3.2.2) :

$$\begin{aligned} & \left[\sqrt{\lambda} \begin{bmatrix} \bar{\mathbf{R}}_{11}(\bar{t}-1) & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{21}(\bar{t}-1) & \bar{\mathbf{R}}_{22}(\bar{t}-1) & \mathbf{0} \\ \bar{\mathbf{R}}_{31}(\bar{t}-1) & \bar{\mathbf{R}}_{32}(\bar{t}-1) & \bar{\mathbf{R}}_{33}(\bar{t}-1) \end{bmatrix} \begin{bmatrix} \mathbf{u}_f^+(\bar{t}) \\ \bar{\boldsymbol{\xi}}(\bar{t}) \\ \mathbf{y}_f^+(\bar{t}) \end{bmatrix} \mathbf{Rot}_{G_1}(\bar{t}) \mathbf{Rot}_{G_2}(\bar{t}) = \\ & \begin{bmatrix} \bar{\mathbf{R}}_{11}(\bar{t}) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{21}(\bar{t}) & \sqrt{\lambda} \begin{bmatrix} \bar{\mathbf{R}}_{22}(\bar{t}-1) & \mathbf{0} \\ \bar{\mathbf{R}}_{32}(\bar{t}-1) & \bar{\mathbf{R}}_{33}(\bar{t}-1) \end{bmatrix} \begin{bmatrix} \bar{\mathbf{z}}_p^-(\bar{t}) \\ \bar{\mathbf{z}}_f^+(\bar{t}) \end{bmatrix} \mathbf{Rot}_{G_2}(\bar{t}) = \\ & \begin{bmatrix} \bar{\mathbf{R}}_{11}(\bar{t}) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{21}(\bar{t}) & \bar{\mathbf{R}}_{22}(\bar{t}) & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{R}}_{31}(\bar{t}) & \bar{\mathbf{R}}_{32}(\bar{t}) & \sqrt{\lambda} \bar{\mathbf{R}}_{33}(\bar{t}-1) & \bar{\mathbf{z}}_f^+(\bar{t}) \end{bmatrix}. \end{aligned} \quad (5.60)$$

Notre objectif n'est donc pas de redévelopper cet algorithme (nommé RPIPO_{qr} dans la suite de ce manuscrit) mais de prouver qu'il existe une relation liant $\bar{\mathbf{z}}_f^+$ et $\bar{\mathbf{z}}_f^+$ au vecteur d'observation, relation qui légitimerait leur utilisation pour l'estimation d'une base de l'espace d'observabilité. Pour cela, considérons dans un premier temps l'équation (5.60). En multipliant les termes du milieu et de droite de la relation (5.60) par leur transposée respective⁹ et en comparant terme à terme les membres de l'égalité ainsi obtenue, il est aisé de montrer que :

$$\bar{\mathbf{R}}_{32}(\bar{t}) \bar{\mathbf{R}}_{32}^T(\bar{t}) = \lambda \bar{\mathbf{R}}_{32}(\bar{t}-1) \bar{\mathbf{R}}_{32}^T(\bar{t}-1) + \bar{\mathbf{z}}_f^+(\bar{t}) \bar{\mathbf{z}}_f^{+T}(\bar{t}) - \bar{\mathbf{z}}_f^+(\bar{t}) \bar{\mathbf{z}}_f^{+T}(\bar{t}). \quad (5.61)$$

Comme dans le cas déterministe, le résultat de la factorisation QR $\bar{\mathbf{R}}_{32} \bar{\mathbf{Q}}_2$ vérifie asymptotiquement la relation suivante :

$$\bar{\mathbf{R}}_{32} \bar{\mathbf{Q}}_2 = \tilde{\mathbf{Y}}_f^+ - \mathbf{H}_f^{MC} \tilde{\mathbf{U}}_f^+ \quad (5.62)$$

où $\tilde{\mathbf{U}}_f^+$ et $\tilde{\mathbf{Y}}_f^+$ sont les matrices de Hankel des données d'entrée-sortie non bruitées et \mathbf{H}_f^{MC} l'optimum du critère $\left\| \mathbf{Y}_f^+ - \mathbf{H}_f \mathbf{U}_f^+ - \mathbf{N}_f^+ \right\|_F^2$. Cette écriture se justifie par le fait que cette projection

⁹Rappelons que pour toute rotation de Givens, $\mathbf{Rot}_G \mathbf{Rot}_G^T = \mathbf{I}$ [GV96].

se compose successivement, à une matrice de pondération près, d'une projection orthogonale sur le noyau de \mathbf{U}_f^+ et d'une variable instrumentale. Puisque $\bar{\mathbf{Q}}_2$ est une matrice orthogonale :

$$\bar{\mathbf{R}}_{32}(\bar{t})\bar{\mathbf{R}}_{32}^T(\bar{t}) = \left[\tilde{\mathbf{Y}}_f^+(\bar{t}) - \mathbf{H}_f^{MC}\tilde{\mathbf{U}}_f^+(\bar{t}) \right] \left[\tilde{\mathbf{Y}}_f^+(\bar{t}) - \mathbf{H}_f^{MC}\tilde{\mathbf{U}}_f^+(\bar{t}) \right]^T. \quad (5.63)$$

Or :

$$\tilde{\mathbf{U}}_f^+(\bar{t}) = \left[\sqrt{\lambda}\tilde{\mathbf{U}}_f^+(\bar{t}-1) \quad \tilde{\mathbf{u}}_f^+(\bar{t}) \right] \quad (5.64)$$

$$\tilde{\mathbf{Y}}_f^+(\bar{t}) = \left[\sqrt{\lambda}\tilde{\mathbf{Y}}_f^+(\bar{t}-1) \quad \tilde{\mathbf{y}}_f^+(\bar{t}) \right]. \quad (5.65)$$

En introduisant ces deux relations au sein de l'équation (5.63), il est facile de vérifier que :

$$\begin{aligned} \bar{\mathbf{R}}_{32}(\bar{t})\bar{\mathbf{R}}_{32}^T(\bar{t}) &= \lambda \left[\tilde{\mathbf{Y}}_f^+(\bar{t}-1) - \mathbf{H}_f^{MC}\tilde{\mathbf{U}}_f^+(\bar{t}-1) \right] \left[\tilde{\mathbf{Y}}_f^+(\bar{t}-1) - \mathbf{H}_f^{MC}\tilde{\mathbf{U}}_f^+(\bar{t}-1) \right]^T \\ &\quad + \left[\tilde{\mathbf{y}}_f^+(\bar{t}) - \mathbf{H}_f^{MC}\tilde{\mathbf{u}}_f^+(\bar{t}) \right] \left[\tilde{\mathbf{y}}_f^+(\bar{t}) - \mathbf{H}_f^{MC}\tilde{\mathbf{u}}_f^+(\bar{t}) \right]^T. \end{aligned} \quad (5.66)$$

Puisque le vecteur d'observation épuré du bruit vérifie :

$$\mathbf{z}_f^+ = \tilde{\mathbf{y}}_f^+ - \mathbf{H}_f^{MC}\tilde{\mathbf{u}}_f^+, \quad (5.67)$$

nous avons :

$$\bar{\mathbf{R}}_{32}(\bar{t})\bar{\mathbf{R}}_{32}^T(\bar{t}) = \lambda\bar{\mathbf{R}}_{32}(\bar{t}-1)\bar{\mathbf{R}}_{32}^T(\bar{t}-1) + \mathbf{z}_f^+(\bar{t})\mathbf{z}_f^{+T}(\bar{t}). \quad (5.68)$$

En exploitant les équations (5.61) et (5.68), on vérifie alors que :

$$\mathbf{z}_f^+(\bar{t})\mathbf{z}_f^{+T}(\bar{t}) = \bar{\bar{\mathbf{z}}}_f^+(\bar{t})\bar{\bar{\mathbf{z}}}_f^{+T}(\bar{t}) - \bar{\bar{\mathbf{z}}}_f^+(\bar{t})\bar{\bar{\mathbf{z}}}_f^{+T}(\bar{t}). \quad (5.69)$$

L'algorithme RPIPO_{qr} permet donc d'estimer la dernière composante de la matrice de covariance des observations à chaque nouvelle acquisition.

5.3.3 Complexité numérique et fiabilité des algorithmes de mise à jour du vecteur d'observation

Une caractéristique essentielle de toute méthode récursive est sa complexité numérique. En effet, il est nécessaire de proposer des techniques ayant un faible coût calculatoire afin qu'elles soient utilisables dans des procédures temps réel¹⁰.

Comme le montre le tableau 5.2, les algorithmes ROM_{im} et RPIPO_{im} (cf. §5.3.1.1 et 5.3.2.1) possèdent un coût en nombre de multiplications tout à fait raisonnable¹¹. Malheureusement, sous certaines conditions d'excitation, ces derniers peuvent présenter une instabilité numérique attestée en raison de l'application du lemme d'inversion matricielle. La complexité numérique des méthodes ROM_{qr} et RPIPO_{qr}, est quant à elle supérieure à celle des techniques ROM_{im} et RPIPO_{im} (cf. tab. 5.3). Elles présentent néanmoins l'avantage d'être fiables numériquement puisqu'elles utilisent exclusivement des outils robustes d'algèbre linéaire. Un compromis fiabilité numérique / coût calculatoire sera donc inévitable. Le choix se fera en fonction de la richesse des données d'entrée-sortie et de la puissance des outils informatiques à disposition.

¹⁰A titre de comparaison, la décomposition QR utilisée par l'algorithme Ordinary MOESP demande environ $2(n_u + n_y)^2 f^2 (j - (n_u + n_y)f/3)$ multiplications et celle de l'algorithme PI/PO MOESP $2(n_\xi^2 + (n_u + n_y)^2 f^2) (j - (n_\xi + (n_u + n_y)f)/3)$ avec $j \gg f$.

¹¹ $n_\xi = n_u p$ pour la classe d'identification PI MOESP, $(n_u + n_y)p$ pour PO MOESP.

| Algorithme | Complexité |
|---------------------|--|
| ROM _{im} | $2(n_u f)^2 + 2n_u n_y f^2 + n_u f$ |
| RPIPO _{im} | $2(n_u f)^2 + 2n_u n_y f^2 + n_u f + 2n_\xi n_u f + 2n_\xi n_y f + 2n_\xi^2 + n_\xi$ |

TAB. 5.2: Complexité des algorithmes d'estimation du vecteur d'observation basés sur le lemme d'inversion matricielle.

| Algorithme | Complexité |
|---------------------|--|
| ROM _{qr} | $4 \left(\frac{n_u}{2} + n_y \right) n_u f^2 + 6n_u f$ |
| RPIPO _{qr} | $4 \left(\frac{n_u f + n_\xi}{2} + n_y f \right) (n_u f + n_\xi) + 6 (n_u f + n_\xi)$ |

TAB. 5.3: Complexité des algorithmes d'estimation du vecteur d'observation basés sur les rotations de Givens.

5.3.4 Conclusion

Quatre algorithmes de mise à jour du vecteur d'observation viennent d'être présentés. Deux contextes (déterministe et stochastique) et deux outils mathématiques d'algèbre linéaire (lemme d'inversion matricielle et rotations de Givens) ont été plus particulièrement considérés. Il a été démontré que, dans le contexte déterministe, les deux méthodes proposées estiment, au signe près, le vecteur d'observation recherché. Dans le contexte stochastique, il n'est possible d'avoir accès qu'à la nouvelle composante de la matrice de covariance des observations. Cette différence sera à prendre en compte dans la seconde phase de l'estimation i.e. la mise à jour de la matrice d'observabilité. La complexité numérique de ces algorithmes a également été analysée. Cette étude a montré que les techniques utilisant le lemme d'inversion matricielle possèdent un coût calculatoire inférieur à leurs homologues utilisant des rotations de Givens. Ces derniers présentent néanmoins l'avantage d'une robustesse numérique établie. Le choix de l'outil d'algèbre linéaire utilisé dépendra donc de la qualité des mesures accessibles ainsi que de la durée séparant deux acquisitions.

La seconde étape de notre problème d'identification récursive des sous-espaces peut donc être désormais considérée. Comme précédemment, différents cas devront être étudiés selon la technique utilisée pour estimer le vecteur d'observation. Plusieurs algorithmes vont donc être développés. Ces derniers ont en commun de reposer sur l'adaptation de la méthode du propagateur (cf. §4.2.4).

5.4 Estimation récursive d'une base de l'espace d'observabilité

La section précédente nous a permis de proposer quatre algorithmes récursifs de mise à jour du vecteur d'observation. Nous nous focalisons donc désormais sur la seconde étape de notre problème d'identification : estimer, en ligne, une base de l'espace d'observabilité. L'objectif de cette section consiste plus exactement à développer de nouvelles techniques d'estimation en ligne de la matrice d'observabilité alternatives aux méthodes proposées jusqu'ici (cf. chapitre 4). Il s'agit d'une phase incontournable en identification des sous-espaces puisque c'est à partir de celle-ci que sont extraites des estimées consistantes des matrices d'état du système via, par exemple, la propriété d'A-invariance de $\mathbf{\Gamma}_f$ et l'application d'algorithmes de type moindres carrés (cf. §5.6). Les techniques d'estimation récursive présentées dans la suite de ce manuscrit sont toutes fondées sur l'adaptation d'une méthode particulière utilisée en traitement d'antennes pour la détection des directions d'arrivée : la méthode du propagateur (cf. §4.2.4). Cette technique présente en effet l'avantage de fournir une décomposition de l'espace d'observation en deux sous-espaces complémentaires via l'application d'un simple opérateur linéaire. Nous montrerons plus précisément que, sous certaines hypothèses peu contraignantes, il est possible de construire un critère quadratique sans approximation et sans contrainte à partir du propagateur. Plusieurs fonctions coût et plusieurs algorithmes récursifs de minimisation seront développés compte tenu des hypothèses fixées sur les signaux agissant sur le procédé. Nous verrons plus particulièrement qu'il est possible d'estimer de manière consistante une base de l'espace d'observabilité de procédé soumis à des bruits colorés inconnus.

5.4.1 Adaptation de la méthode du propagateur au problème d'identification

Nous avons vu dans la section 4.2.4 qu'il était possible d'estimer une base du sous-espace signal (resp. sous-espace bruit) en évaluant un opérateur linéaire. Or, de par l'analogie entre le problème d'identification et celui de détection d'arrivée mise en évidence au sein de la section 5.1, il semblerait intéressant d'adapter cet opérateur à notre problème d'identification. Pour cela, il est nécessaire de vérifier un ensemble d'hypothèses.

5.4.1.1 Hypothèses à vérifier

Supposons que la paire $\{\mathbf{A}, \mathbf{C}\}$ soit observable. Puisque $\mathbf{\Gamma}_f \in \mathbb{R}^{n_y f \times n_x}$ avec $n_y f > n_x$, la matrice d'observabilité vérifie :

$$\text{rang}(\mathbf{\Gamma}_f) = n_x. \quad (5.70)$$

$\mathbf{\Gamma}_f$ possède donc au moins un ensemble de n_x lignes, non nécessairement contiguës, linéairement indépendantes. Soit $\mathbf{\Gamma}_{f_1} \in \mathbb{R}^{n_x \times n_x}$ la sous-matrice composée d'un de ces ensembles et $\mathbf{\Gamma}_{f_2} \in \mathbb{R}^{n_y f - n_x \times n_x}$ celle contenant les lignes de $\mathbf{\Gamma}_f$ non présentes dans $\mathbf{\Gamma}_{f_1}$. La décomposition suivante peut alors être considérée :

$$\mathbf{\Gamma}_f = \begin{bmatrix} \mathbf{\Gamma}_{f_1} \\ \mathbf{\Gamma}_{f_2} \end{bmatrix} \begin{matrix} \} \mathbb{R}^{n_x \times n_x} \\ \} \mathbb{R}^{n_y f - n_x \times n_x} \end{matrix}. \quad (5.71)$$

Puisque $\mathbf{\Gamma}_{f_2}$ est la sous-matrice complémentaire de $\mathbf{\Gamma}_{f_1}$, $\mathbf{\Gamma}_{f_2}$ peut s'exprimer comme une combinaison linéaire de $\mathbf{\Gamma}_{f_1}$. Il existe donc un unique opérateur $\mathbf{P}_f \in \mathbb{R}^{n_x \times n_y f - n_x}$, nommé propagateur [MD91], tel que :

$$\mathbf{\Gamma}_{f_2} = \mathbf{P}_f^T \mathbf{\Gamma}_{f_1}. \quad (5.72)$$

La matrice d'observabilité peut alors être réécrite de la façon suivante :

$$\mathbf{\Gamma}_f = \begin{bmatrix} \mathbf{\Gamma}_{f_1} \\ \mathbf{\Gamma}_{f_2} \end{bmatrix} = \begin{bmatrix} \mathbf{\Gamma}_{f_1} \\ \mathbf{P}_f^T \mathbf{\Gamma}_{f_1} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{I}_{n_x} \\ \mathbf{P}_f^T \end{bmatrix}}_{\mathbf{E}_o} \mathbf{\Gamma}_{f_1}. \quad (5.73)$$

Cette équation prouve que les colonnes de $\mathbf{\Gamma}_f$ peuvent s'exprimer comme une combinaison linéaire des colonnes de \mathbf{E}_o . Or, comme $\text{rang}(\mathbf{\Gamma}_{f_1}) = n_x$:

$$\text{Im}_{\text{col}}(\mathbf{\Gamma}_f) = \text{Im}_{\text{col}}(\mathbf{E}_o). \quad (5.74)$$

Cette dernière équation met ainsi en évidence qu'il est possible de déterminer une base de l'espace d'observabilité par l'intermédiaire de \mathbf{E}_o . Or, l'obtention d'une telle matrice nécessite uniquement d'estimer le propagateur \mathbf{P}_f . Par conséquent, en supposant que *l'ordre du système soit a priori connu et que le système étudié soit observable*, une estimée du sous-espace engendré par $\mathbf{\Gamma}_f$ est accessible via l'estimation de \mathbf{P}_f . C'est à cette tâche que nous nous attelons dans la suite de cette section.

5.4.1.2 Critère minimisé dans le cas de données non bruitées

L'étape d'estimation précédente (cf. §5.3) nous a permis d'avoir accès à une estimée consistante du vecteur d'observation. Or, nous savons que ce dernier vérifie :

$$\mathbf{z}_f^+(t) = \mathbf{\Gamma}_f \mathbf{x}(t) + \mathbf{n}_f^+(t). \quad (5.75)$$

Après réorganisation des données de telle sorte que les n_x premières lignes de la matrice $\mathbf{\Gamma}_f$ soient linéairement indépendantes, la partition suivante peut être considérée :

$$\mathbf{z}_f^+(t) = \begin{bmatrix} \mathbf{I}_{n_x} \\ \mathbf{P}_f^T \end{bmatrix} \mathbf{\Gamma}_{f_1} \mathbf{x}(t) + \mathbf{n}_f^+(t) = \begin{cases} \mathbf{z}_{f_1}^+(t) \\ \mathbf{z}_{f_2}^+(t) \end{cases} \begin{cases} \in \mathbb{R}^{n_x \times 1} \\ \in \mathbb{R}^{n_y f - n_x \times 1} \end{cases}. \quad (5.76)$$

$\mathbf{z}_{f_1}^+$ et $\mathbf{z}_{f_2}^+$ regroupent les composantes de \mathbf{z}_f^+ correspondant¹² respectivement aux n_x lignes de $\mathbf{\Gamma}_{f_1}$ et aux $n_y f - n_x$ lignes de $\mathbf{\Gamma}_{f_2}$.

Dans le cas d'un vecteur d'observation non bruité, il est facile de montrer que :

$$\mathbf{z}_{f_2}^+ = \mathbf{P}_f^T \mathbf{z}_{f_1}^+ \quad (5.77)$$

puisque $\mathbf{n}_f^+(t) \equiv \mathbf{0}$, que :

$$\begin{cases} \mathbf{z}_{f_1}^+(t) &= \mathbf{\Gamma}_{f_1} \mathbf{x}(t) \\ \mathbf{z}_{f_2}^+(t) &= \mathbf{P}_f^T \mathbf{\Gamma}_{f_1} \mathbf{x}(t) \end{cases} \quad (5.78)$$

¹²Des symboles identiques sont utilisées avant et après réorganisation afin de ne pas surcharger les notations.

et que $\mathbf{\Gamma}_{f_1}$ est inversible. La mise à jour du propagateur est alors réalisable en considérant le critère suivant :

$$J(\mathbf{P}_f) = \mathbb{E} \left\| \mathbf{z}_{f_2}^+ - \mathbf{P}_f^T \mathbf{z}_{f_1}^+ \right\|^2, \quad (5.79)$$

critère qu'il est aisé de récrire comme suit :

$$J(\mathbf{P}_f(t)) = \sum_{k=1}^t \lambda^{t-k} \left\| \mathbf{z}_{f_2}^+(k) - \mathbf{P}_f^T(t) \mathbf{z}_{f_1}^+(k) \right\|^2 \quad (5.80)$$

en remplaçant l'espérance mathématique par une somme pondérée exponentiellement¹³. L'unicité de $\hat{\mathbf{P}}_f$ est alors assurée par la convexité de cette fonction coût, convexité garantie quant à elle par la condition d'excitation persistante. La minimisation d'un tel critère est alors réalisable récursivement en lui appliquant les moindres carrés récursifs [Lju99] :

$$\mathbf{K}_f(t) = \frac{\mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t-1)}{\lambda + \mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t-1) \mathbf{z}_{f_1}^+(t)} \quad (5.81a)$$

$$\mathbf{L}_f(t) = \frac{1}{\lambda} \left(\mathbf{L}_f(t-1) - \mathbf{L}_f(t-1) \mathbf{z}_{f_1}^+(t) \mathbf{K}_f(t) \right) \quad (5.81b)$$

$$\mathbf{P}_f^T(t) = \mathbf{P}_f^T(t-1) + \left(\mathbf{z}_{f_2}^+(t) - \mathbf{P}_f^T(t-1) \mathbf{z}_{f_1}^+(t) \right) \mathbf{K}_f(t). \quad (5.81c)$$

Cet algorithme sera nommé RPM1 dans la suite de ce manuscrit. Sa démonstration est développée en annexe A.5.

5.4.1.3 Effets des perturbations

Supposons maintenant que le bruit \mathbf{n}_f^+ ne soit plus identiquement nul. En appliquant à \mathbf{n}_f^+ une répartition similaire à l'équation (5.76) :

$$\mathbf{n}_f^+ = \begin{bmatrix} \mathbf{n}_{f_1}^+ \\ \mathbf{n}_{f_2}^+ \end{bmatrix} \left. \begin{array}{l} \} \in \mathbb{R}^{n_x \times 1} \\ \} \in \mathbb{R}^{n_{yf} - n_x \times 1} \end{array} \right\} , \quad (5.82)$$

la matrice de covariance du bruit s'écrit :

$$\mathbb{E} \left\{ \mathbf{n}_f^+ \mathbf{n}_f^{+T} \right\} = \mathbf{R}_{\mathbf{n}_f^+} = \begin{bmatrix} \mathbf{R}_{\mathbf{n}_{f_1}^+} & \mathbf{R}_{\mathbf{n}_{f_1}^+ \mathbf{n}_{f_2}^+} \\ \mathbf{R}_{\mathbf{n}_{f_2}^+ \mathbf{n}_{f_1}^+} & \mathbf{R}_{\mathbf{n}_{f_2}^+} \end{bmatrix}. \quad (5.83)$$

Nous savons de plus que l'estimée asymptotique au sens des moindres carrés de \mathbf{P}_f^T est, dans ces conditions, donnée par [Lju99] :

$$\hat{\mathbf{P}}_f^T = \mathbf{R}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+} \mathbf{R}_{\mathbf{z}_{f_1}^+}^{-1} \quad (5.84)$$

avec :

$$\mathbf{R}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+} = \mathbb{E} \left\{ \mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^{+T} \right\} \quad (5.85)$$

$$\mathbf{R}_{\mathbf{z}_{f_1}^+} = \mathbb{E} \left\{ \mathbf{z}_{f_1}^+ \mathbf{z}_{f_1}^{+T} \right\}. \quad (5.86)$$

¹³Rappelons que nous n'avons accès qu'à un nombre fini de données. Il est donc nécessaire d'approximer l'espérance mathématique par une somme finie. Le facteur d'oubli λ est quant à lui utilisé pour pondérer les données passées.

Preuve : Considérons le critère suivant :

$$\mathbf{J}(\mathbf{P}_f) = \mathbb{E} \left\| \mathbf{z}_{f_2}^+ - \mathbf{P}_f^T \mathbf{z}_{f_1}^+ \right\|^2. \quad (5.87)$$

Puisque pour tout vecteur \mathbf{x} :

$$\|\mathbf{x}\|^2 = \text{Tr}(\mathbf{x}\mathbf{x}^T), \quad (5.88)$$

ce critère se réécrit comme suit :

$$\begin{aligned} J(\mathbf{P}_f) &= \mathbb{E} \left\{ \text{Tr} \left(\left(\mathbf{z}_{f_2}^+ - \mathbf{P}_f^T \mathbf{z}_{f_1}^+ \right) \left(\mathbf{z}_{f_2}^{+T} - \mathbf{z}_{f_1}^{+T} \mathbf{P}_f \right) \right) \right\} \\ &= \mathbb{E} \left\{ \text{Tr} \left(\mathbf{z}_{f_2}^+ \mathbf{z}_{f_2}^{+T} - \mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^{+T} \mathbf{P}_f - \mathbf{P}_f^T \mathbf{z}_{f_1}^+ \mathbf{z}_{f_2}^{+T} + \mathbf{P}_f^T \mathbf{z}_{f_1}^+ \mathbf{z}_{f_1}^{+T} \mathbf{P}_f \right) \right\} \\ &= \text{Tr} \left(\mathbf{R}_{\mathbf{z}_{f_2}^+} - \mathbf{R}_{\mathbf{z}_{f_1}^+ \mathbf{z}_{f_2}^+}^T \mathbf{P}_f - \mathbf{P}_f^T \mathbf{R}_{\mathbf{z}_{f_1}^+ \mathbf{z}_{f_2}^+} + \mathbf{P}_f^T \mathbf{R}_{\mathbf{z}_{f_1}^+} \mathbf{P}_f \right). \end{aligned} \quad (5.89)$$

Le minimum de ce dernier est obtenu en le dérivant par rapport à la matrice recherchée \mathbf{P}_f . Ainsi :

$$\frac{\partial J(\mathbf{P}_f)}{\partial \mathbf{P}_f} = -2\mathbf{R}_{\mathbf{z}_{f_1}^+ \mathbf{z}_{f_2}^+} + 2\mathbf{R}_{\mathbf{z}_{f_1}^+} \mathbf{P}_f. \quad (5.90)$$

Le minimum est donc atteint pour $\hat{\mathbf{P}}_f^T = \mathbf{R}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+} \mathbf{R}_{\mathbf{z}_{f_1}^+}^{-1}$. ◀

Il est alors facile de mettre en évidence que la solution au sens des moindres carrés du problème d'optimisation défini par (5.79) conduit à une estimée asymptotiquement biaisée, même lorsque le vecteur de résidus \mathbf{n}_f^+ est spatialement et temporellement blanc. En effet, puisque :

$$\mathbf{R}_z = \begin{bmatrix} \mathbf{R}_{\mathbf{z}_{f_1}^+} & \mathbf{R}_{\mathbf{z}_{f_1}^+ \mathbf{z}_{f_2}^+} \\ \mathbf{R}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+} & \mathbf{R}_{\mathbf{z}_{f_2}^+} \end{bmatrix} = [\Gamma_{f_1} \quad \Gamma_{f_2}] \mathbf{R}_x \begin{bmatrix} \Gamma_{f_1}^T \\ \Gamma_{f_2}^T \end{bmatrix} + \begin{bmatrix} \mathbf{R}_{\mathbf{n}_{f_1}^+} & \mathbf{R}_{\mathbf{n}_{f_1}^+ \mathbf{n}_{f_2}^+} \\ \mathbf{R}_{\mathbf{n}_{f_2}^+ \mathbf{n}_{f_1}^+} & \mathbf{R}_{\mathbf{n}_{f_2}^+} \end{bmatrix}, \quad (5.91)$$

les relations suivantes sont vérifiées :

$$\mathbf{R}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+} = \mathbf{P}_f^T \Gamma_{f_1} \mathbf{R}_x \Gamma_{f_1}^T + \mathbf{R}_{\mathbf{n}_{f_2}^+ \mathbf{n}_{f_1}^+} = \mathbf{P}_f^T \Gamma_{f_1} \mathbf{R}_x \Gamma_{f_1}^T \quad (5.92)$$

$$\mathbf{R}_{\mathbf{z}_{f_1}^+} = \Gamma_{f_1} \mathbf{R}_x \Gamma_{f_1}^T + \mathbf{R}_{\mathbf{n}_{f_1}^+} = \Gamma_{f_1} \mathbf{R}_x \Gamma_{f_1}^T + \sigma^2 \mathbf{I}_{n_x} \quad (5.93)$$

lorsque $\mathbf{R}_{\mathbf{n}_f^+} = \sigma^2 \mathbf{I}_{n_{y,f}}$. Le projecteur s'écrit alors :

$$\hat{\mathbf{P}}_f^T = \mathbf{P}_f^T \Gamma_{f_1} \mathbf{R}_x \Gamma_{f_1}^T (\Gamma_{f_1} \mathbf{R}_x \Gamma_{f_1}^T + \sigma^2 \mathbf{I}_{n_x})^{-1} \neq \mathbf{P}_f^T, \quad (5.94)$$

relation qui montre explicitement que le projecteur est asymptotiquement biaisé. Cette difficulté pourrait être partiellement contournée en considérant une approche de type moindres carrés totaux [VV91]. Malheureusement l'application d'une telle méthode compliquerait fortement l'implémentation récursive de l'algorithme. Deux solutions sont proposées au sein de ce manuscrit :

- La première suppose que le vecteur d'observation a été estimé à l'aide des versions récursives des schémas d'identification PI/PO MOESP. Ce dernier est alors asymptotiquement épuré des effets des perturbations.

- La seconde considère que vecteur d'observation a été calculé en utilisant les versions récursives des schémas d'identification Ordinary MOESP. Aucun traitement des perturbations n'a donc été réalisé *a priori*. La solution consiste alors à introduire une variable instrumentale au sein du critère (5.79) afin qu'il soit exploitable quel que soit le bruit agissant sur le système.

Remarque 5.5. *Il est intéressant de compléter cette étude en mesurant l'ordre d'influence des perturbations sur l'estimation du propagateur. La théorie des perturbations [Vac94] nous permet d'écrire que, pour toutes matrice \mathbf{M} et petite perturbation $\Delta\mathbf{M}$:*

$$(\mathbf{M} + \Delta\mathbf{M})^{-1} \cong \mathbf{M}^{-1} - \mathbf{M}^{-1}\Delta\mathbf{M}\mathbf{M}^{-1}. \quad (5.95)$$

En posant :

$$\mathbf{M} = \Gamma_{f_1} \mathbf{R}_x \Gamma_{f_1}^T \quad (5.96)$$

$$\Delta\mathbf{M} = \sigma^2 \mathbf{I}_{n_x}, \quad (5.97)$$

il est facile de vérifier que :

$$(\Gamma_{f_1} \mathbf{R}_x \Gamma_{f_1}^T + \sigma^2 \mathbf{I}_{n_x})^{-1} = (\Gamma_{f_1} \mathbf{R}_x \Gamma_{f_1}^T)^{-1} - \sigma^2 (\Gamma_{f_1} \mathbf{R}_x \Gamma_{f_1}^T)^{-1} (\Gamma_{f_1} \mathbf{R}_x \Gamma_{f_1}^T)^{-1}. \quad (5.98)$$

La relation (5.94) s'écrit alors :

$$\hat{\mathbf{P}}_f^T \cong \mathbf{P}_f^T - \sigma^2 \mathbf{P}_f^T (\Gamma_{f_1} \mathbf{R}_x \Gamma_{f_1}^T)^{-1}. \quad (5.99)$$

Cette équation indique clairement que la méthode du propagateur est asymptotiquement biaisée au premier ordre.

5.4.2 Estimation à partir d'un vecteur d'observation épuré des effets des perturbations

Supposons que le vecteur d'observation ait été estimé à l'aide des versions récursives des schémas d'identification PI/PO MOESP (cf. §5.3.2.1 et 5.3.2.2). Les algorithmes d'estimation proposés précédemment ne permettent alors d'avoir accès au vecteur d'observation que via les relations suivantes :

$$\mathbf{z}_f^+ \mathbf{z}_f^{+T} = \check{\mathbf{z}}_f^+ \check{\mathbf{z}}_f^{+T} - \check{\check{\mathbf{z}}}_f^+ \check{\check{\mathbf{z}}}_f^{+T} \quad (\text{cf. algo. RPIPO}_{\text{im}}) \quad (5.100)$$

$$\mathbf{z}_f^+ \mathbf{z}_f^{+T} = \bar{\bar{\mathbf{z}}}_f^+ \bar{\bar{\mathbf{z}}}_f^{+T} - \bar{\bar{\bar{\mathbf{z}}}}_f^+ \bar{\bar{\bar{\mathbf{z}}}}_f^{+T} \quad (\text{cf. algo. RPIPO}_{\text{qr}}) \quad (5.101)$$

L'estimation du propagateur n'est alors plus réalisable à partir du critère (5.79) puisque ce dernier ne fait appel qu'au vecteur d'observation et aucunement à son carré. Il est donc nécessaire de proposer d'autres fonctions coût prenant en compte la matrice $\mathbf{z}_f^+ \mathbf{z}_f^{+T}$ pour estimer le propagateur. Deux critères vont être présentés dans la suite de ce paragraphe. Ces derniers émanent du constat suivant :

$$\mathbf{R}_{\mathbf{z}_f^+} = \mathbb{E} \left\{ \mathbf{z}_f^+ \mathbf{z}_f^{+T} \right\} = \Gamma_f \mathbf{R}_x \Gamma_f^T \quad (5.102)$$

puisque $\mathbf{z}_f^+ \mathbf{z}_f^{+T}$ est asymptotiquement épuré des effets des perturbations. Or :

$$\mathbf{\Gamma}_f = \begin{bmatrix} \mathbf{I}_{n_x} \\ \mathbf{P}_f^T \end{bmatrix} \mathbf{\Gamma}_{f_1} \quad (5.103)$$

où $\mathbf{\Gamma}_{f_1}$ peut être considérée comme une matrice de changement de base. Il existe donc un vecteur $\bar{\mathbf{x}} = \mathbf{\Gamma}_{f_1} \mathbf{x}$ tel que :

$$\mathbf{R}_{\mathbf{z}_f^+} = \begin{bmatrix} \mathbf{I}_{n_x} \\ \mathbf{P}_f^T \end{bmatrix} \mathbf{R}_{\bar{\mathbf{x}}} \begin{bmatrix} \mathbf{I}_{n_x} & \mathbf{P}_f \end{bmatrix}. \quad (5.104)$$

Ainsi :

$$\mathbf{R}_{\mathbf{z}_f^+} = \begin{bmatrix} \mathbf{R}_{\mathbf{z}_{f_1}^+} & \mathbf{R}_{\mathbf{z}_{f_1}^+ \mathbf{z}_{f_2}^+} \\ \mathbf{R}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+} & \mathbf{R}_{\mathbf{z}_{f_2}^+} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{\bar{\mathbf{x}}} & \mathbf{R}_{\bar{\mathbf{x}} \mathbf{P}_f} \\ \mathbf{P}_f^T \mathbf{R}_{\bar{\mathbf{x}}} & \mathbf{P}_f^T \mathbf{R}_{\bar{\mathbf{x}} \mathbf{P}_f} \end{bmatrix}. \quad (5.105)$$

Cette dernière relation permet alors de montrer que :

$$\mathbf{P}_f^T = \mathbf{R}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+} \mathbf{R}_{\mathbf{z}_{f_1}^+}^{-1}. \quad (5.106)$$

De cette équation découle alors le critère suivant :

$$\bar{J}(\mathbf{P}_f) = \left\| \mathbf{R}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+} - \mathbf{P}_f^T \mathbf{R}_{\mathbf{z}_{f_1}^+} \right\|_F^2. \quad (5.107)$$

En s'inspirant des travaux réalisés par J. L. Yu [Yu00] sur les algorithmes COPAST (cf. §4.2.3.3), nous proposons de minimiser cette fonction coût en minimisant séquentiellement n_x critères $\bar{J}_i(\mathbf{P}_f)$ définis par :

$$\bar{J}_i(\mathbf{P}_f(t)) = \sum_{k=1}^t \lambda^{t-k} \left\| \mathbf{r}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+}^i(k) - \mathbf{P}_f^T(t) \mathbf{r}_{\mathbf{z}_{f_1}^+}^i(k) \right\|^2 \quad (5.108)$$

où $\mathbf{r}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+}^i$ et $\mathbf{r}_{\mathbf{z}_{f_1}^+}^i$ sont respectivement les i^{e} colonnes de $\mathbf{R}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+}$ et de $\mathbf{R}_{\mathbf{z}_{f_1}^+}$. Le critère (5.108) ainsi obtenu est alors équivalent à la fonction coût (5.80). L'algorithme RPM1 est donc ajustable à notre problème de minimisation. L'algorithme complet, nommé COPM dans la suite de ce manuscrit, s'écrit alors :

$$\mathbf{K}_f(t) = \frac{\mathbf{r}_{\mathbf{z}_{f_1}^+}^i{}^T(t) \mathbf{L}_f(t-1)}{\lambda + \mathbf{r}_{\mathbf{z}_{f_1}^+}^i{}^T(t) \mathbf{L}_f(t-1) \mathbf{r}_{\mathbf{z}_{f_1}^+}^i(t)} \quad (5.109a)$$

$$\mathbf{L}_f(t) = \frac{1}{\lambda} \left(\mathbf{L}_f(t-1) - \mathbf{L}_f(t-1) \mathbf{r}_{\mathbf{z}_{f_1}^+}^i(t) \mathbf{K}_f(t) \right) \quad (5.109b)$$

$$\mathbf{P}_f^T(t) = \mathbf{P}_f^T(t-1) + \left(\mathbf{r}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+}^i(t) - \mathbf{P}_f^T(t-1) \mathbf{r}_{\mathbf{z}_{f_1}^+}^i(t) \right) \mathbf{K}_f(t). \quad (5.109c)$$

La matrice $\mathbf{L}_f(0)$ peut être choisie comme un multiple de la matrice identité. Le premier projecteur peut quant à lui être construit à partir des valeurs des matrices \mathbf{A} et \mathbf{C} initiales.

Cet algorithme nécessite d'extraire les vecteurs $\mathbf{r}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+}^i$ et $\mathbf{r}_{\mathbf{z}_{f_1}^+}^i$. Ces derniers sont obtenus à chaque mise à jour de la matrice de covariance $\mathbf{R}_{\mathbf{z}_f^+}$ en remarquant, dans un premier temps, que

les matrices $\mathbf{R}_{\mathbf{z}_{f_1}^+}$ et $\mathbf{R}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+}$ correspondent respectivement aux n_x premières lignes et $n_{yf} - n_x$ dernières lignes des n_x premières colonnes de $\mathbf{R}_{\mathbf{z}_f^+}$ (cf. equ. (5.105)). Ces matrices étant extraites à chaque nouvelle acquisition, les vecteurs $\mathbf{r}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+}^i$ et $\mathbf{r}_{\mathbf{z}_{f_1}^+}^i$ désirés sont isolés comme suit :

$$i = \text{reste}(t, n_x) + 1 \quad (5.110a)$$

$$\mathbf{r}_{\mathbf{z}_{f_1}^+}^i(t) = \mathbf{R}_{\mathbf{z}_f^+}(1 : n_x, i)(t) \quad (5.110b)$$

$$\mathbf{r}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+}^i(t) = \mathbf{R}_{\mathbf{z}_f^+}(n_x + 1 : n_{yf}, i)(t). \quad (5.110c)$$

Un second critère peut être proposé en complétant la relation (5.106). Afin de faire apparaître simplement cette fonction coût, généralisons, dans un premier temps, les équations (5.100) et (5.101) à l'aide de l'expression suivante :

$$\mathbf{z}_f^+ \mathbf{z}_f^{+T} = \underline{\mathbf{z}}_f^+ \underline{\mathbf{z}}_f^{+T} - \underline{\mathbf{z}}_f^+ \underline{\mathbf{z}}_f^{+T}. \quad (5.111)$$

La matrice de covariance des observations s'écrit alors :

$$\mathbf{R}_{\mathbf{z}_f^+} = \mathbf{R}_{\underline{\mathbf{z}}_f^+} - \mathbf{R}_{\underline{\mathbf{z}}_f^+} = \begin{bmatrix} \mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+} - \mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+} & \mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+ \underline{\mathbf{z}}_{f_2}^+} - \mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+ \underline{\mathbf{z}}_{f_2}^+} \\ \mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_1}^+} - \mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_1}^+} & \mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+} - \mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+} \end{bmatrix}. \quad (5.112)$$

Ainsi :

$$\mathbf{P}_f^T = \left(\mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_1}^+} - \mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_1}^+} \right) \left(\mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+} - \mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+} \right)^{-1}. \quad (5.113)$$

Le second critère que nous proposons est le suivant :

$$\bar{J}(\mathbf{P}_f) = \mathbb{E} \left\| \underline{\mathbf{z}}_{f_2}^+ - \mathbf{P}_f^T \underline{\mathbf{z}}_{f_1}^+ \right\|^2 - \mathbb{E} \left\| \underline{\mathbf{z}}_{f_2}^+ - \mathbf{P}_f^T \underline{\mathbf{z}}_{f_1}^+ \right\|^2 \quad (5.114)$$

dont la solution est donnée par l'équation (5.113).

Preuve : Considérons le critère suivant :

$$\bar{J}(\mathbf{P}_f) = \mathbb{E} \left\| \underline{\mathbf{z}}_{f_2}^+ - \mathbf{P}_f^T \underline{\mathbf{z}}_{f_1}^+ \right\|^2 - \mathbb{E} \left\| \underline{\mathbf{z}}_{f_2}^+ - \mathbf{P}_f^T \underline{\mathbf{z}}_{f_1}^+ \right\|^2. \quad (5.115)$$

Comme pour tout vecteur \mathbf{x} :

$$\|\mathbf{x}\|^2 = \text{Tr}(\mathbf{x}\mathbf{x}^T), \quad (5.116)$$

il est aisé de vérifier que :

$$\begin{aligned} \bar{J}(\mathbf{P}_f) &= \mathbb{E} \left\{ \text{Tr} \left(\left(\underline{\mathbf{z}}_{f_2}^+ - \mathbf{P}_f^T \underline{\mathbf{z}}_{f_1}^+ \right) \left(\underline{\mathbf{z}}_{f_2}^{+T} - \underline{\mathbf{z}}_{f_1}^{+T} \mathbf{P}_f \right) \right) - \text{Tr} \left(\left(\underline{\mathbf{z}}_{f_2}^+ - \mathbf{P}_f^T \underline{\mathbf{z}}_{f_1}^+ \right) \left(\underline{\mathbf{z}}_{f_2}^{+T} - \underline{\mathbf{z}}_{f_1}^{+T} \mathbf{P}_f \right) \right) \right\} \\ &= \mathbb{E} \left\{ \text{Tr} \left(\underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_2}^{+T} - \underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_2}^{+T} - \mathbf{P}_f^T \left(\underline{\mathbf{z}}_{f_1}^+ \underline{\mathbf{z}}_{f_2}^{+T} - \underline{\mathbf{z}}_{f_1}^+ \underline{\mathbf{z}}_{f_2}^{+T} \right) \right. \right. \\ &\quad \left. \left. - \left(\underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_1}^{+T} - \underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_1}^{+T} \right) \mathbf{P}_f + \mathbf{P}_f^T \left(\underline{\mathbf{z}}_{f_1}^+ \underline{\mathbf{z}}_{f_1}^{+T} - \underline{\mathbf{z}}_{f_1}^+ \underline{\mathbf{z}}_{f_1}^{+T} \right) \mathbf{P}_f \right) \right\} \\ &= \text{Tr} \left(\mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+} - \mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+} - 2\mathbf{P}_f^T \left(\mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+ \underline{\mathbf{z}}_{f_2}^+} - \mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+ \underline{\mathbf{z}}_{f_2}^+} \right) + \mathbf{P}_f^T \left(\mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+} - \mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+} \right) \mathbf{P}_f \right). \end{aligned} \quad (5.117)$$

Or, pour toutes matrices \mathbf{A} et \mathbf{X} [Bro91] :

$$\frac{\partial \text{Tr}(\mathbf{X}^T \mathbf{A})}{\partial \mathbf{X}} = \mathbf{A} \quad (5.118)$$

$$\frac{\partial \text{Tr}(\mathbf{X}^T \mathbf{A} \mathbf{X})}{\partial \mathbf{X}} = (\mathbf{A} + \mathbf{A}^T) \mathbf{X}. \quad (5.119)$$

Ainsi :

$$\frac{\partial \bar{J}(\mathbf{P}_f)}{\partial \mathbf{P}_f} = -2 \left(\mathbf{R}_{\underline{z}_{f_1}^+ \underline{z}_{f_2}^+} - \mathbf{R}_{\underline{z}_{f_1}^+ \underline{z}_{f_1}^+} \right) + 2 \left(\mathbf{R}_{\underline{z}_{f_1}^+} - \mathbf{R}_{\underline{z}_{f_1}^+} \right) \mathbf{P}_f \quad (5.120)$$

dont la solution est bien évidemment :

$$\mathbf{P}_f^T = \left(\mathbf{R}_{\underline{z}_{f_2}^+ \underline{z}_{f_1}^+} - \mathbf{R}_{\underline{z}_{f_2}^+ \underline{z}_{f_1}^+} \right) \left(\mathbf{R}_{\underline{z}_{f_1}^+} - \mathbf{R}_{\underline{z}_{f_1}^+} \right)^{-1}. \quad (5.121)$$

◀

La minimisation d'une telle fonction coût est réalisable à l'aide de l'algorithme RPM2 suivant :

$$\underline{\mathbf{L}}_f(t) = \frac{1}{\lambda} \left(\underline{\mathbf{L}}_f(t-1) - \frac{\underline{\mathbf{L}}_f(t-1) \underline{\mathbf{z}}_{f_1}^+(t) \underline{\mathbf{z}}_{f_1}^{+T}(t) \underline{\mathbf{L}}_f(t-1)}{\lambda + \underline{\mathbf{z}}_{f_1}^{+T}(t) \underline{\mathbf{L}}_f(t-1) \underline{\mathbf{z}}_{f_1}^+(t)} \right) \quad (5.122a)$$

$$\underline{\mathbf{L}}_f(t) = \underline{\mathbf{L}}_f(t) + \frac{\underline{\mathbf{L}}_f(t) \underline{\mathbf{z}}_{f_1}^+(t) \underline{\mathbf{z}}_{f_1}^{+T}(t) \underline{\mathbf{L}}_f(t)}{1 - \underline{\mathbf{z}}_{f_1}^{+T}(t) \underline{\mathbf{L}}_f(t) \underline{\mathbf{z}}_{f_1}^+(t)} \quad (5.122b)$$

$$\begin{aligned} \mathbf{P}_f^T(t) &= \mathbf{P}_f^T(t-1) + \left(\underline{\mathbf{z}}_{f_2}^+(t) - \mathbf{P}_f^T(t-1) \underline{\mathbf{z}}_{f_1}^+(t) \right) \underline{\mathbf{z}}_{f_1}^{+T}(t) \underline{\mathbf{L}}_f(t) \\ &\quad - \left(\underline{\mathbf{z}}_{f_2}^+(t) - \mathbf{P}_f^T(t-1) \underline{\mathbf{z}}_{f_1}^+(t) \right) \underline{\mathbf{z}}_{f_1}^{+T}(t) \underline{\mathbf{L}}_f(t) \end{aligned} \quad (5.122c)$$

en appliquant par deux fois le lemme d'inversion matricielle. Sa démonstration est développée en annexe A.6. Comme précédemment, une façon simple d'initialiser cet algorithme consiste à choisir des matrices $\underline{\mathbf{L}}_f(0)$ et $\underline{\mathbf{L}}_f(0)$ proportionnelles à l'identité.

5.4.3 Introduction d'une variable instrumentale

Les techniques proposées précédemment présentent l'inconvénient d'utiliser la matrice de covariance $\mathbf{R}_{\underline{z}_f^+}$ dont l'estimation (cf. tab. 5.2 et 5.3) implique un coût calculatoire plus important qu'en appliquant les algorithmes ROM_{im} et ROM_{qr}. Il serait donc intéressant de développer de nouvelles méthodes récursives d'estimation du propagateur utilisant les solutions de ces derniers. Les techniques ROM_{im} et ROM_{qr} fournissent malheureusement des estimées bruitées lorsque le bruit agissant sur le procédé n'est pas blanc. Il est donc primordial de proposer des algorithmes d'estimation du propagateur supprimant l'effet de ces perturbations.

De nombreuses méthodes d'identification utilisant le principe de la variable instrumentale [SS89, Lju99] ont été développées ces dernières années pour estimer des modèles de systèmes en environnement bruité. Leur principe consiste à introduire, au sein du critère à minimiser, un vecteur construit à partir de données d'entrée-sortie retardées permettant d'éliminer l'influence du bruit contenu dans le régresseur à l'instant t . Afin d'appliquer une telle technique à notre

problème d'identification, considérons¹⁴ donc une variable instrumentale $\boldsymbol{\xi} \in \mathbb{R}^{n_\xi \times 1}$, $n_\xi \geq n_x$, supposée non corrélée avec le bruit $\mathbf{n}_f^+(t)$ mais suffisamment corrélée avec l'état $\mathbf{x}(t)$. En introduisant cette variable au sein de la fonction (5.79), nous obtenons :

$$J_{IV}(\mathbf{P}_f(t)) = \left\| \mathbf{R}_{\mathbf{z}_{f_2}^+} \boldsymbol{\xi}(t) - \mathbf{P}_f^T(t) \mathbf{R}_{\mathbf{z}_{f_1}^+} \boldsymbol{\xi}(t) \right\|_F^2 \quad (5.123)$$

avec $\mathbf{R}_{\mathbf{z}_{f_1}^+} \boldsymbol{\xi} = \mathbb{E} \left\{ \mathbf{z}_{f_1}^+(k) \boldsymbol{\xi}^T(k) \right\} \in \mathbb{R}^{n_x \times n_\xi}$ et $\mathbf{R}_{\mathbf{z}_{f_2}^+} \boldsymbol{\xi} = \mathbb{E} \left\{ \mathbf{z}_{f_2}^+(k) \boldsymbol{\xi}^T(k) \right\} \in \mathbb{R}^{n_{yf} - n_x \times n_\xi}$. En particulier, en remplaçant les espérances mathématiques par des sommes finies et pondérées, ce dernier critère s'écrit :

$$J_{IV}(\mathbf{P}_f(t)) = \sum_{k=1}^t \lambda^{t-k} \left\| \mathbf{z}_{f_2}^+(k) \boldsymbol{\xi}^T(k) - \mathbf{P}_f^T(t) \mathbf{z}_{f_1}^+(k) \boldsymbol{\xi}^T(k) \right\|_F^2. \quad (5.124)$$

La minimisation d'un tel critère est alors réalisable de trois manières différentes.

5.4.3.1 Algorithmes COIVPM, IVPM et EIVPM

Algorithme COIVPM La première solution, nommée COIVPM et inspirée des travaux de J. L. Yu [Yu00], consiste à minimiser séquentiellement n_ξ fonctions coût :

$$J_{IV_i}(\mathbf{P}_f(t)) = \sum_{k=1}^t \lambda^{t-k} \left\| \mathbf{r}_{\mathbf{z}_{f_2}^+}^i \boldsymbol{\xi}(k) - \mathbf{P}_f^T(t) \mathbf{r}_{\mathbf{z}_{f_1}^+}^i \boldsymbol{\xi}(k) \right\|^2 \text{ pour } i \text{ variant de } 1 \text{ à } n_\xi \quad (5.125)$$

où $\mathbf{r}_{\mathbf{z}_{f_1}^+}^i \boldsymbol{\xi}$ et $\mathbf{r}_{\mathbf{z}_{f_2}^+}^i \boldsymbol{\xi}$ sont respectivement les i^e colonnes des matrices de covariance $\mathbf{R}_{\mathbf{z}_{f_1}^+} \boldsymbol{\xi}$ et $\mathbf{R}_{\mathbf{z}_{f_2}^+} \boldsymbol{\xi}$. Cette technique permet ainsi de ne plus considérer directement les matrices de covariance $\mathbf{R}_{\mathbf{z}_{f_1}^+} \boldsymbol{\xi}$ et $\mathbf{R}_{\mathbf{z}_{f_2}^+} \boldsymbol{\xi}$ mais de simples vecteurs $\mathbf{r}_{\mathbf{z}_{f_2}^+}^i \boldsymbol{\xi}$ et $\mathbf{r}_{\mathbf{z}_{f_1}^+}^i \boldsymbol{\xi}$ qui rendent la minimisation des fonctions J_{IV_i} plus aisée. Chacun de ces critères est en effet minimisable à l'aide d'une technique de type moindres carrés [Lju99] équivalente à celles développées pour les algorithmes RPM1 et COPM :

$$\mathbf{K}_f(t) = \frac{\mathbf{r}_{\mathbf{z}_{f_1}^+}^i \boldsymbol{\xi}^T(t) \mathbf{L}_f(t-1)}{\lambda + \mathbf{r}_{\mathbf{z}_{f_1}^+}^i \boldsymbol{\xi}^T(t) \mathbf{L}_f(t-1) \mathbf{r}_{\mathbf{z}_{f_1}^+}^i \boldsymbol{\xi}(t)} \quad (5.126a)$$

$$\mathbf{L}_f(t) = \frac{1}{\lambda} \left(\mathbf{L}_f(t-1) - \mathbf{L}_f(t-1) \mathbf{r}_{\mathbf{z}_{f_1}^+}^i \boldsymbol{\xi}(t) \mathbf{K}_f(t) \right) \quad (5.126b)$$

$$\mathbf{P}_f^T(t) = \mathbf{P}_f^T(t-1) + \left(\mathbf{r}_{\mathbf{z}_{f_2}^+}^i \boldsymbol{\xi}(t) - \mathbf{P}_f^T(t-1) \mathbf{r}_{\mathbf{z}_{f_1}^+}^i \boldsymbol{\xi}(t) \right) \mathbf{K}_f(t). \quad (5.126c)$$

L'extraction des vecteurs $\mathbf{r}_{\mathbf{z}_{f_1}^+}^i \boldsymbol{\xi}$ et $\mathbf{r}_{\mathbf{z}_{f_2}^+}^i \boldsymbol{\xi}$ est quant à elle réalisée comme suit :

$$i = \text{reste}(t, n_\xi) + 1 \quad (5.127a)$$

$$\mathbf{r}_{\mathbf{z}_{f_1}^+}^i \boldsymbol{\xi}(t) = \mathbf{R}_z \boldsymbol{\xi}(1 : n_x, i)(t) \quad (5.127b)$$

$$\mathbf{r}_{\mathbf{z}_{f_2}^+}^i \boldsymbol{\xi}(t) = \mathbf{R}_z \boldsymbol{\xi}(n_x + 1 : n_{yf}, i)(t). \quad (5.127c)$$

¹⁴Un choix de variable instrumentale vérifiant ces conditions est présenté §5.4.3.3.

Remarque 5.6. L'algorithme COIVPM est analogue à l'algorithme COPM du paragraphe 5.4.2 si nous remplaçons, toute chose égale par ailleurs, $\mathbf{r}_{\mathbf{z}_{f_1}^+}^i \boldsymbol{\xi}$ par $\mathbf{r}_{\mathbf{z}_{f_1}^+}^i$ et $\mathbf{r}_{\mathbf{z}_{f_2}^+}^i \boldsymbol{\xi}$ par $\mathbf{r}_{\mathbf{z}_{f_1}^+ \mathbf{z}_{f_2}^+}^i$.

Algorithmes IVPM et EIVPM La minimisation du critère (5.123) est également réalisable récursivement en considérant plus simplement l'expression de son minimum. Ce dernier s'écrit de deux manières différentes selon la dimension de la variable instrumentale utilisée. Afin de mettre en évidence ces deux possibilités, considérons de nouveau la fonction coût (5.123). Cette dernière s'écrit alors :

$$J_{IV}(\mathbf{P}_f(t)) = \text{Tr} \left(\left(\mathbf{R}_{\mathbf{z}_{f_2}^+} \boldsymbol{\xi}(t) - \mathbf{P}_f^T(t) \mathbf{R}_{\mathbf{z}_{f_1}^+} \boldsymbol{\xi}(t) \right) \left(\mathbf{R}_{\mathbf{z}_{f_2}^+}^T \boldsymbol{\xi}(t) - \mathbf{R}_{\mathbf{z}_{f_1}^+}^T \boldsymbol{\xi}(t) \mathbf{P}_f(t) \right) \right) \quad (5.128)$$

puisque, pour toute matrice \mathbf{X} [Bro91] :

$$\|\mathbf{X}\|_F^2 = \text{Tr}(\mathbf{X}\mathbf{X}^T). \quad (5.129)$$

Ainsi :

$$J_{IV}(\mathbf{P}_f(t)) = \text{Tr} \left(\mathbf{R}_{\mathbf{z}_{f_2}^+} \boldsymbol{\xi}(t) \mathbf{R}_{\mathbf{z}_{f_2}^+}^T \boldsymbol{\xi}(t) - \mathbf{P}_f^T(t) \mathbf{R}_{\mathbf{z}_{f_1}^+} \boldsymbol{\xi}(t) \mathbf{R}_{\mathbf{z}_{f_2}^+}^T \boldsymbol{\xi}(t) \right. \\ \left. - \mathbf{R}_{\mathbf{z}_{f_2}^+} \boldsymbol{\xi}(t) \mathbf{R}_{\mathbf{z}_{f_1}^+}^T \boldsymbol{\xi}(t) \mathbf{P}_f(t) + \mathbf{P}_f^T(t) \mathbf{R}_{\mathbf{z}_{f_1}^+} \boldsymbol{\xi}(t) \mathbf{R}_{\mathbf{z}_{f_1}^+}^T \boldsymbol{\xi}(t) \mathbf{P}_f(t) \right) \quad (5.130)$$

Or, pour toutes matrices \mathbf{A} et \mathbf{X} réelles [Bro91] :

$$\frac{\partial \text{Tr}(\mathbf{X}^T \mathbf{A})}{\partial \mathbf{X}} = \frac{\partial \text{Tr}(\mathbf{A}^T \mathbf{X})}{\partial \mathbf{X}} = \mathbf{A} \quad (5.131)$$

$$\frac{\partial \text{Tr}(\mathbf{X}^T \mathbf{A} \mathbf{X})}{\partial \mathbf{X}} = (\mathbf{A} + \mathbf{A}^T) \mathbf{X}. \quad (5.132)$$

Ainsi :

$$\frac{\partial J_{IV}(\mathbf{P}_f)}{\partial \mathbf{P}_f} = -2\mathbf{R}_{\mathbf{z}_{f_1}^+} \boldsymbol{\xi}(t) \mathbf{R}_{\mathbf{z}_{f_2}^+}^T \boldsymbol{\xi}(t) + 2\mathbf{R}_{\mathbf{z}_{f_1}^+} \boldsymbol{\xi}(t) \mathbf{R}_{\mathbf{z}_{f_1}^+}^T \boldsymbol{\xi}(t) \mathbf{P}_f(t). \quad (5.133)$$

Le minimum du critère (5.123) est donc atteint pour \mathbf{P}_f vérifiant :

$$\mathbf{P}_f^T(t) \mathbf{R}_{\mathbf{z}_{f_1}^+} \boldsymbol{\xi}(t) = \mathbf{R}_{\mathbf{z}_{f_2}^+} \boldsymbol{\xi}(t). \quad (5.134)$$

Deux cas doivent alors être envisagés selon le nombre d'instruments composant la variable instrumentale $\boldsymbol{\xi}$:

- Si $n_{\boldsymbol{\xi}} = n_x$, en supposant que la variable instrumentale soit suffisamment corrélée au vecteur $\mathbf{z}_{f_1}^+$ de telle sorte que $\mathbf{R}_{\mathbf{z}_{f_1}^+} \boldsymbol{\xi}$ soit inversible :

$$\mathbf{P}_f^T(t) = \mathbf{R}_{\mathbf{z}_{f_2}^+} \boldsymbol{\xi}(t) \mathbf{R}_{\mathbf{z}_{f_1}^+}^{-1} \boldsymbol{\xi}(t). \quad (5.135)$$

- Si $n_{\boldsymbol{\xi}} > n_x$, la matrice $\mathbf{R}_{\mathbf{z}_{f_1}^+} \boldsymbol{\xi}$ n'est plus carrée. Son inversion nécessite alors d'utiliser la pseudo inverse de Moore Penrose [GV96] :

$$\mathbf{P}_f^T(t) = \mathbf{R}_{\mathbf{z}_{f_2}^+} \boldsymbol{\xi}(t) \mathbf{R}_{\mathbf{z}_{f_1}^+}^\dagger \boldsymbol{\xi}(t). \quad (5.136)$$

Deux algorithmes sont alors développés dans la suite de ce paragraphe pour estimer le propagateur \mathbf{P}_f .

Algorithme IVPM Si $n_\xi = n_x$, l'estimation de \mathbf{P}_f est réalisable en appliquant le lemme d'inversion matricielle à la solution (5.135) :

$$\mathbf{K}_f(t) = \frac{\boldsymbol{\xi}^T(t)\mathbf{L}_f(t-1)}{\lambda + \boldsymbol{\xi}^T(t)\mathbf{L}_f(t-1)\mathbf{z}_{f_1}^+(t)} \quad (5.137a)$$

$$\mathbf{L}_f(t) = \frac{1}{\lambda} \left(\mathbf{L}_f(t-1) - \mathbf{L}_f(t-1)\mathbf{z}_{f_1}^+(t)\mathbf{K}_f(t) \right) \quad (5.137b)$$

$$\mathbf{P}_f^T(t) = \mathbf{P}_f^T(t-1) + \left(\mathbf{z}_{f_2}^+(t) - \hat{\mathbf{P}}_f^T(t-1)\mathbf{z}_{f_1}^+(t) \right) \mathbf{K}_f(t). \quad (5.137c)$$

avec $\mathbf{L}_f(t) = \left(\mathbb{E} \left\{ \mathbf{z}_{f_1}^+(t)\boldsymbol{\xi}^T(t) \right\} \right)^{-1} = \mathbf{R}_{\mathbf{z}_{f_1}^+\boldsymbol{\xi}}^{-1}(t)$. Cet algorithme sera nommé IVPM dans la suite de ce manuscrit. Sa démonstration est équivalente à celle développée pour l'algorithme RPM1 (cf. annexe A.5).

Algorithme EIVPM Lorsque $n_\xi > n_x$, l'utilisation du lemme d'inversion matricielle n'est plus satisfaisante pour estimer rékursivement le propagateur puisque la matrice $\mathbf{R}_{\mathbf{z}_{f_1}^+\boldsymbol{\xi}}$ est, par construction, rectangulaire. Or, cette situation n'est pas rare en pratique puisque, la plupart du temps, la variable instrumentale est réalisée à partir de versions retardées des données d'entrée-sortie ce qui rend difficile de fixer *a priori* $n_\xi = n_x$. De plus, dans l'ouvrage [SS89], il est précisé que la fidélité de l'estimée obtenue à l'aide d'une méthode de type variable instrumentale augmente avec le nombre d'instruments utilisés. Il est donc particulièrement intéressant de développer un algorithme capable de considérer une matrice $\mathbf{R}_{\mathbf{z}_{f_1}^+\boldsymbol{\xi}}$ non carrée. La solution développée dans ce paragraphe fait appel à une technique particulière de minimisation nommée méthode de la variable instrumentale étendue [Fri84] (cf. §2.3.2). L'ajustement d'une telle méthode à la solution (5.136) permet d'obtenir un nouvel algorithme d'estimation du propagateur, nommé EIVPM, dont les formules de mise à jour sont les suivantes :

$$\mathbf{g}_f(t) = \begin{bmatrix} \mathbf{R}_{\mathbf{z}_{f_2}^+\boldsymbol{\xi}}(t)\boldsymbol{\xi}(t) & \mathbf{z}_{f_2}^+(t) \end{bmatrix} \quad (5.138a)$$

$$\boldsymbol{\Lambda}(t) = \begin{bmatrix} -\boldsymbol{\xi}^T(t)\boldsymbol{\xi}(t) & \lambda \\ \lambda & 0 \end{bmatrix} \quad (5.138b)$$

$$\boldsymbol{\Psi}_f(t) = \begin{bmatrix} \mathbf{R}_{\mathbf{z}_{f_1}^+\boldsymbol{\xi}}(t-1)\boldsymbol{\xi}(t) & \mathbf{z}_{f_1}^+(t) \end{bmatrix} \quad (5.138c)$$

$$\mathbf{K}_f(t) = \left(\boldsymbol{\Lambda}(t) + \boldsymbol{\Psi}_f^T(t)\mathbf{L}_f(t-1)\boldsymbol{\Psi}_f(t) \right)^{-1} \boldsymbol{\Psi}_f^T(t)\mathbf{L}_f(t-1) \quad (5.138d)$$

$$\mathbf{P}_f^T(t) = \mathbf{P}_f^T(t-1) + \left(\mathbf{g}_f(t) - \mathbf{P}_f^T(t-1)\boldsymbol{\Psi}_f(t) \right) \mathbf{K}_f(t) \quad (5.138e)$$

$$\mathbf{R}_{\mathbf{z}_{f_1}^+\boldsymbol{\xi}}(t) = \lambda \mathbf{R}_{\mathbf{z}_{f_1}^+\boldsymbol{\xi}}(t-1) + \mathbf{z}_{f_1}^+(t)\boldsymbol{\xi}^T(t) \quad (5.138f)$$

$$\mathbf{R}_{\mathbf{z}_{f_2}^+\boldsymbol{\xi}}(t) = \lambda \mathbf{R}_{\mathbf{z}_{f_2}^+\boldsymbol{\xi}}(t-1) + \mathbf{z}_{f_2}^+(t)\boldsymbol{\xi}^T(t) \quad (5.138g)$$

$$\mathbf{L}_f(t) = \frac{1}{\lambda^2} \left(\mathbf{L}_f(t-1) - \mathbf{L}_f(t-1)\boldsymbol{\Psi}(t)\mathbf{K}_f(t) \right) \quad (5.138h)$$

$$\text{avec } \mathbf{L}_f(t) = \left(\mathbf{R}_{\mathbf{z}_{f_1}^+\boldsymbol{\xi}}(t)\mathbf{R}_{\mathbf{z}_{f_1}^+\boldsymbol{\xi}}^T(t) \right)^{-1}. \quad (5.138i)$$

La démonstration de cet algorithme est présentée en annexe A.7. Les matrices utilisées pour l'initialiser peuvent être simplement choisies égales à la matrice identité ou comme des tronca- tures de cette dernière.

5.4.3.2 Amélioration de la méthode de la variable étendue : l'algorithme EIV_{sqr t}PM

Comme tout algorithme type moindres carrés récurrents, une condition nécessaire de stabi- lité numérique est d'imposer, à chaque itération, que la matrice \mathbf{L}_f soit semi-définie positive [Rot01]. En effet, lorsque cette matrice est mal conditionnée, la non vérification de cette pro- priété peut conduire à la divergence de l'algorithme. Or, de par sa formule de mise à jour (cf. equ. (5.138h)), cette condition ne peut être assurée. En effet, même si la matrice \mathbf{L}_f est semi- définie positive à l'instant $t - 1$, l'expression :

$$\mathbf{L}_f(t) = \frac{1}{\lambda^2} \left(\mathbf{L}_f(t-1) - \mathbf{L}_f(t-1) \boldsymbol{\Psi}_f(t) (\boldsymbol{\Lambda}(t) + \boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1) \boldsymbol{\Psi}_f(t))^{-1} \boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1) \right) \quad (5.139)$$

ne respecte pas cette propriété à l'instant t puisque :

$$\boldsymbol{\Lambda}(t) + \boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1) \boldsymbol{\Psi}_f(t) = \begin{bmatrix} -\boldsymbol{\xi}^T(t) \left(\mathbf{I} - \mathbf{R}_{z_{f_1}^+}^T \boldsymbol{\xi}(t-1) \mathbf{L}_f(t-1) \mathbf{R}_{z_{f_1}^+} \boldsymbol{\xi}(t-1) \right) \boldsymbol{\xi}(t) & \lambda + \boldsymbol{\xi}^T(t) \mathbf{R}_{z_{f_1}^+}^T \boldsymbol{\xi}(t-1) \mathbf{L}_f(t-1) \mathbf{z}_{f_1}^+(t) \\ \lambda + \mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t-1) \mathbf{R}_{z_{f_1}^+} \boldsymbol{\xi}(t-1) \boldsymbol{\xi}(t) & \mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t-1) \mathbf{z}_{f_1}^+(t) \end{bmatrix} \quad (5.140)$$

possède deux valeurs propres de signe opposé, $\mathbf{I} - \mathbf{R}_{z_{f_1}^+}^T \boldsymbol{\xi}(t-1) \mathbf{L}_f(t-1) \mathbf{R}_{z_{f_1}^+} \boldsymbol{\xi}(t-1)$ étant une matrice semi-définie positive. Il est alors impossible de garantir une bonne convergence de EIVPM (ou tout autre algorithme utilisant la technique EIV classique [SS89]). La solution proposée jusqu'ici reposait sur l'évaluation de \mathbf{L}_f à partir de l'équation (5.138h), de l'extraction de la partie triangulaire supérieure puis de l'utilisation de sa propriété de symétrie pour compléter sa mise à jour [Gus98]. Cette manipulation ne permet malheureusement pas d'assu- rer la définie-positivité de \mathbf{L}_f . La technique développée au sein de cette section assure quant à elle cette propriété. Elle consiste plus précisément en l'adaptation, au cas multivariable, d'un algorithme de mise à jour de la méthode de la variable instrumentale étendue proposée par B. Porat et B. Friedlander [PF89]. L'idée sous-jacente est de calculer, à chaque itération, la racine carrée de \mathbf{L}_f , condition suffisante à la vérification de la définie positivité de \mathbf{L}_f .

Supposons donc désormais que la matrice \mathbf{L}_f soit semi-définie positive à l'instant t . Si $\boldsymbol{\Lambda}$ était également semi-définie positive, il serait possible d'écrire $\mathbf{L}_f(t)$ de la façon suivante :

$$\mathbf{L}_f(t) = \frac{1}{\lambda^2} \left(\mathbf{L}_f^{1/2}(t-1) \mathbf{L}_f^{T/2}(t-1) - \mathbf{L}_f^{1/2}(t-1) \left(\boldsymbol{\Psi}_f^T(t) \mathbf{L}_f^{1/2}(t-1) \right)^T \left(\boldsymbol{\Lambda}^{1/2}(t) \boldsymbol{\Lambda}^{T/2}(t) + \boldsymbol{\Psi}_f^T(t) \mathbf{L}_f^{1/2}(t-1) \left(\boldsymbol{\Psi}_f^T(t) \mathbf{L}_f^{1/2}(t-1) \right)^T \right)^{-1} \boldsymbol{\Psi}_f^T(t) \mathbf{L}_f^{1/2}(t-1) \mathbf{L}_f^{T/2}(t-1) \right). \quad (5.141)$$

La mise à jour des différents éléments composant cette expression serait alors accessible en considérant la matrice suivante :

$$\mathbf{T}_f(t) = \begin{bmatrix} \mathbf{\Lambda}^{1/2}(t) & \mathbf{\Psi}_f^T(t)\mathbf{L}_f^{1/2}(t-1) \\ \mathbf{0} & \mathbf{L}_f^{1/2}(t-1) \end{bmatrix} \quad (5.142)$$

puisque :

$$\mathbf{T}_f(t)\mathbf{T}_f^T(t) = \begin{bmatrix} \mathbf{\Lambda}(t) + \mathbf{\Psi}_f^T(t)\mathbf{L}_f(t-1)\mathbf{\Psi}_f(t) & \mathbf{\Psi}_f^T(t)\mathbf{L}_f(t-1) \\ \mathbf{L}_f(t-1)\mathbf{\Psi}_f(t) & \mathbf{L}_f(t-1) \end{bmatrix} \quad (5.143)$$

contient les composantes de $\mathbf{L}_f(t)$. Malheureusement, $\mathbf{\Lambda}$ (cf. equ. (5.138b)) n'est pas semi-définie positive. Elle peut cependant être réécrite de la façon suivante :

$$\begin{aligned} \mathbf{\Lambda}(t) &= \begin{bmatrix} \sqrt{\xi^T(t)\xi(t)} & 0 \\ -\frac{\lambda}{\sqrt{\xi^T(t)\xi(t)}} & \frac{\lambda}{\sqrt{\xi^T(t)\xi(t)}} \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{\xi^T(t)\xi(t)} & -\frac{\lambda}{\sqrt{\xi^T(t)\xi(t)}} \\ 0 & \frac{\lambda}{\sqrt{\xi^T(t)\xi(t)}} \end{bmatrix} \\ &= \mathbf{\Delta}(t)\mathbf{J}\mathbf{\Delta}^T(t) \end{aligned} \quad (5.144)$$

avec \mathbf{J} matrice de signature de $\mathbf{\Lambda}$ [GV96]. Ainsi, en posant :

$$\bar{\mathbf{T}}_f(t) = \begin{bmatrix} \mathbf{\Delta}(t) & \mathbf{\Psi}_f^T(t)\mathbf{L}_f^{1/2}(t-1) \\ \mathbf{0} & \mathbf{L}_f^{1/2}(t-1) \end{bmatrix}, \quad (5.145)$$

l'expression (5.143) est modifiable comme suit :

$$\bar{\mathbf{T}}_f(t) \begin{bmatrix} \mathbf{J} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n_x} \end{bmatrix} \bar{\mathbf{T}}_f^T(t) = \begin{bmatrix} \mathbf{\Lambda}(t) + \mathbf{\Psi}_f^T(t)\mathbf{L}_f(t-1)\mathbf{\Psi}_f(t) & \mathbf{\Psi}_f^T(t)\mathbf{L}_f(t-1) \\ \mathbf{L}_f(t-1)\mathbf{\Psi}_f(t) & \mathbf{L}_f(t-1) \end{bmatrix}. \quad (5.146)$$

L'étape suivante consiste donc à mettre à jour la matrice $\bar{\mathbf{T}}_f(t)$. En se référant à [GV96], nous savons qu'il existe une matrice $\mathbf{Rot}_H \in \mathbb{R}^{2 \times 2}$ composée de rotations hyperboliques telle que :

$$(\mathbf{Rot}_H)\mathbf{J}(\mathbf{Rot}_H)^T = \mathbf{J} \text{ et } \mathbf{\Delta}\mathbf{Rot}_H \text{ triangulaire inférieure.} \quad (5.147)$$

De même, pour toute rotation de Givens \mathbf{Rot}_G [GV96] :

$$(\mathbf{Rot}_G)\mathbf{I}(\mathbf{Rot}_G)^T = \mathbf{I}. \quad (5.148)$$

La méthode de mise à jour développée dans ce paragraphe et adaptée de [PF89] consiste à associer ces deux types de rotations au sein d'une unique matrice, nommée \mathbf{Rot} , afin de les appliquer à la matrice $\bar{\mathbf{T}}_f(t)$ de telle sorte que leurs actions cumulées fournissent une matrice triangulaire :

$$\bar{\mathbf{T}}_f(t) \underbrace{\begin{bmatrix} \mathbf{Rot}_H & \mathbf{0} \\ \mathbf{0} & \mathbf{Rot}_G \end{bmatrix}}_{\mathbf{Rot}} = \begin{bmatrix} \mathbf{L}_{11} & \mathbf{0} \\ \mathbf{L}_{21} & \mathbf{L}_{22} \end{bmatrix} \quad (5.149)$$

avec \mathbf{L}_{11} et \mathbf{L}_{22} deux matrices triangulaires inférieures de dimension respective 2×2 et $n_x \times n_x$. Les relations liant ces matrices à celles composant $\bar{\mathbf{T}}_f$ sont obtenues en calculant :

$$\begin{aligned} \bar{\mathbf{T}}_f(t) \begin{bmatrix} \mathbf{J} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n_x} \end{bmatrix} \bar{\mathbf{T}}_f^T(t) &= \\ \bar{\mathbf{T}}_f(t)\mathbf{Rot} \begin{bmatrix} \mathbf{J} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n_x} \end{bmatrix} \mathbf{Rot}^T \bar{\mathbf{T}}_f^T(t) &= \begin{bmatrix} \mathbf{L}_{11} & \mathbf{0} \\ \mathbf{L}_{21} & \mathbf{L}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{J} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n_x} \end{bmatrix} \begin{bmatrix} \mathbf{L}_{11}^T & \mathbf{L}_{21}^T \\ \mathbf{0} & \mathbf{L}_{22}^T \end{bmatrix}. \end{aligned} \quad (5.150)$$

Les égalités suivantes sont alors vérifiées :

$$\Delta(t)\mathbf{J}\Delta^T(t) + \Psi_f^T(t)\mathbf{L}_f(t)\Psi_f(t) = \mathbf{L}_{11}\mathbf{J}\mathbf{L}_{11}^T \quad (5.151)$$

$$\mathbf{L}_f(t-1)\Psi_f(t) = \mathbf{L}_{21}\mathbf{J}\mathbf{L}_{11}^T \quad (5.152)$$

$$\mathbf{L}_f(t-1) = \mathbf{L}_{21}\mathbf{J}\mathbf{L}_{21}^T + \mathbf{L}_{22}\mathbf{L}_{22}^T. \quad (5.153)$$

A partir, de l'expression (5.152) :

$$\mathbf{L}_{21} = \mathbf{L}_f(t-1)\Psi_f(t)\mathbf{L}_{11}^{-T}\mathbf{J}^{-1}. \quad (5.154)$$

Ainsi, avec (5.151), (5.154) et (5.138e), l'équation (5.153) devient :

$$\mathbf{L}_{22}\mathbf{L}_{22}^T = \mathbf{L}_f(t-1) - \mathbf{L}_{21}\mathbf{J}\mathbf{L}_{21}^T = \mathbf{L}_f(t-1) - \mathbf{L}_f(t-1)\Psi_f(t)\mathbf{K}_f(t) = \lambda^2\mathbf{L}_f(t). \quad (5.155)$$

Donc :

$$\mathbf{L}_f^{1/2}(t) = \frac{1}{\lambda}\mathbf{L}_{22}. \quad (5.156)$$

De plus :

$$\mathbf{K}_f(t) = (\mathbf{L}_{11}\mathbf{J}\mathbf{L}_{11}^T)^{-1}\Psi_f^T(t)\mathbf{L}_f^T(t-1) = (\mathbf{L}_{11}\mathbf{J}\mathbf{L}_{11}^T)^{-1}\mathbf{L}_{11}\mathbf{J}\mathbf{L}_{21}^T = \mathbf{L}_{11}^{-T}\mathbf{L}_{21}^T \quad (5.157)$$

donc :

$$\mathbf{K}_f(t) = (\mathbf{L}_{21}\mathbf{L}_{11}^{-1})^T. \quad (5.158)$$

L'algorithme complet devient alors :

$$\mathbf{g}_f(t) = \begin{bmatrix} \mathbf{R}_{\mathbf{z}_{f_2}^+} \xi(t) \xi(t) & \mathbf{z}_{f_2}^+(t) \end{bmatrix} \quad (5.159a)$$

$$\Delta(t) = \begin{bmatrix} \sqrt{\xi^T(t)\xi(t)} & 0 \\ -\frac{\lambda}{\sqrt{\xi^T(t)\xi(t)}} & \frac{\lambda}{\sqrt{\xi^T(t)\xi(t)}} \end{bmatrix} \quad (5.159b)$$

$$\Psi_f(t) = \begin{bmatrix} \mathbf{R}_{\mathbf{z}_{f_1}^+} \xi(t-1) \xi(t) & \mathbf{z}_{f_1}^+(t) \end{bmatrix} \quad (5.159c)$$

Trouver la matrice de rotations **Rot** telle que :

$$\begin{bmatrix} \Delta(t) & \Psi_f^T(t)\mathbf{L}_f^{1/2}(t-1) \\ \mathbf{0} & \mathbf{L}_f^{1/2}(t-1) \end{bmatrix} \mathbf{Rot} = \begin{bmatrix} \mathbf{L}_{11} & \mathbf{0} \\ \mathbf{L}_{21} & \mathbf{L}_{22} \end{bmatrix} \quad (5.159d)$$

$$\mathbf{L}_f^{1/2}(t) = \frac{1}{\lambda}\mathbf{L}_{22} \quad (5.159e)$$

$$\mathbf{K}_f(t) = (\mathbf{L}_{21}\mathbf{L}_{11}^{-1})^T \quad (5.159f)$$

$$\mathbf{P}_f^T(t) = \mathbf{P}_f^T(t-1) + (\mathbf{g}(t) - \mathbf{P}_f^T(t-1)\Psi_f(t))\mathbf{K}_f(t) \quad (5.159g)$$

$$\mathbf{R}_{\mathbf{z}_{f_1}^+} \xi(t) = \lambda\mathbf{R}_{\mathbf{z}_{f_1}^+} \xi(t-1) + \mathbf{z}_{f_1}^+(t)\xi^T(t) \quad (5.159h)$$

$$\mathbf{R}_{\mathbf{z}_{f_2}^+} \xi(t) = \lambda\mathbf{R}_{\mathbf{z}_{f_2}^+} \xi(t-1) + \mathbf{z}_{f_2}^+(t)\xi^T(t). \quad (5.159i)$$

Il sera désormais nommé $\text{EIV}_{\text{sqrtPM}}$. Comme précédemment, aucune contrainte n'est imposée concernant l'initialisation. Utiliser des matrices proportionnelles à l'identité ou construites à partir d'un bloc de cette dernière peut tout à fait être envisagé.

Remarque 5.7. L'application de rotations hyperboliques au problème précédent nécessite de vérifier que $\mathbf{\Lambda}(t) + \mathbf{\Psi}_f^T(t)\mathbf{L}_f(t-1)\mathbf{\Psi}_f(t)$ possède la même signature que $\mathbf{\Lambda}(t)$, condition justifiant sa décomposition $\mathbf{L}_{11}\mathbf{J}\mathbf{L}_{11}^T$. Pour cela, considérons de nouveau l'équation (5.140). En supposant qu'à l'instant $t-1$ la matrice \mathbf{L}_f soit semi-définie positive, $\mathbf{\Lambda}(t) + \mathbf{\Psi}_f^T(t)\mathbf{L}_f(t-1)\mathbf{\Psi}_f(t)$ peut s'écrire comme suit :

$$\mathbf{\Lambda}(t) + \mathbf{\Psi}_f^T(t)\mathbf{L}_f(t-1)\mathbf{\Psi}_f(t) = \begin{bmatrix} -a & b \\ b & c \end{bmatrix}, \quad (5.160)$$

où $a = -\boldsymbol{\xi}^T(t) \left(\mathbf{I} - \mathbf{C}_{z_{f_1}}^T \boldsymbol{\xi}(t-1) \mathbf{L}_f(t) \mathbf{C}_{z_{f_1}} \boldsymbol{\xi}(t-1) \right) \boldsymbol{\xi}(t)$, $b = \lambda + \boldsymbol{\xi}^T(t) \mathbf{C}_{z_{f_1}}^T \boldsymbol{\xi}(t-1) \mathbf{L}_f(t) \mathbf{z}_{f_1}^+(t)$ et $c = \mathbf{z}_{f_1}^{+T}(t) \mathbf{z}_{f_1}^+(t)$ réels positifs. La décomposition suivante est alors valide :

$$\mathbf{\Lambda}(t) + \mathbf{\Psi}_f^T(t)\mathbf{L}_f(t-1)\mathbf{\Psi}_f(t) = \underbrace{\begin{bmatrix} \sqrt{a} & 0 \\ -\frac{b}{\sqrt{a}} & \sqrt{c + \frac{b^2}{a}} \end{bmatrix}}_{\mathbf{L}} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} \sqrt{a} & -\frac{b}{\sqrt{a}} \\ 0 & \sqrt{c + \frac{b^2}{a}} \end{bmatrix}}_{\mathbf{L}^T}. \quad (5.161)$$

Si $\mathbf{L}_f(t-1)$ est semi-définie positive, $\mathbf{\Lambda}(t) + \mathbf{\Psi}_f^T(t)\mathbf{L}_f(t-1)\mathbf{\Psi}_f(t)$ a comme matrice signature \mathbf{J} . L'application des rotations hyperboliques dans de bonnes conditions est ainsi justifiée [PF89].

5.4.3.3 Choix de la variable instrumentale

De nombreuses études sur le choix de la variable instrumentale ont été réalisées dans divers domaines liés au traitement du signal et à l'automatique [WP67, Cai76, War77, SS83a, You84, VSO95]. Certaines se sont plus particulièrement focalisées sur l'influence de la variable instrumentale en identification hors ligne des sous-espaces [OV94, Gus97b, CV97, VWO97, Gus01]. C'est en particulier sur les travaux de Chou et Verhaegen [CV97] que s'articule la réflexion développée dans ce paragraphe.

| | $\tilde{\mathbf{u}}_p^-(t-r)$ | $\mathbf{v}_p^-(t-r)$ | $\mathbf{w}_p^-(t-r)$ | $\tilde{\mathbf{y}}_p^-(t-r)$ | $\mathbf{v}_p^-(t-r)$ |
|-----------------------------|-------------------------------|-----------------------|-----------------------|-------------------------------|-----------------------|
| $\tilde{\mathbf{u}}_f^+(t)$ | 0 | 0 | 0 | 0 | 0 |
| $\mathbf{v}_f^+(t)$ | 0 | 0 | 0 | 0 | 0 |
| $\mathbf{w}_f^+(t)$ | 0 | 0 | 0 | 0 | 0 |
| $\tilde{\mathbf{y}}_f^+(t)$ | ✓ | ✓/0 | ✓ | ✓ | 0 |
| $\mathbf{v}_f^+(t)$ | 0 | 0 | 0 | 0 | 0 |

TAB. 5.4: Corrélacion entre les données d'entrée-sortie futures et passées.

Puisque les données accessibles à la mesure sont de nature temporelle, le moyen le plus simple pour construire une variable instrumentale fiable est d'utiliser des données d'entrée-sortie retardées [SVWW98]. Or, les seules variables accessibles à la mesure sont \mathbf{u} et \mathbf{y} , données

pouvant bien évidemment être bruitées. Il est donc nécessaire d'étudier la corrélation existant entre les signaux passés et futurs mesurés afin de se concentrer sur les données riches en information et décorréélées des perturbations. Pour cela, considérons les relations présentées dans le tableau 5.4, où r , f et p sont des entiers fixés par l'utilisateur tels que $f > n_x$, $p > n_x$ et tels que la longueur de corrélation temporelle de la variable instrumentale soit supérieure à celle du bruit. Analysons plus particulièrement le pire cas, à savoir des bruits blancs en \mathbf{v} , \mathbf{w} et \mathbf{v} .

Notre objectif est de trouver, parmi les données d'entrée-sortie passées, celles qui peuvent être utilisées comme instruments. Supposons donc qu'à l'instant t nous ayons accès aux ensembles $\{\tilde{\mathbf{u}}_p^-, \mathbf{v}_p^-\}$, $\{\tilde{\mathbf{y}}_p^-, \mathbf{w}_p^-, \mathbf{v}_p^-\}$, $\{\tilde{\mathbf{u}}_f^+, \mathbf{v}_f^+\}$ et $\{\tilde{\mathbf{y}}_f^+, \mathbf{w}_f^+, \mathbf{v}_f^+\}$. Puisque nous n'étudions que l'identification des procédés en boucle ouverte, les bruits futurs et les entrées futures sont indépendantes des bruits, entrées et sorties passés. Au contraire, les sorties futures sont liées aux entrées et sorties passés mais sont non tributaires des bruits de mesure en sortie passés $\mathbf{v}_p^-(t-r)$ puisque ces derniers sont supposés blancs. Un cas reste alors à étudier : la dépendance vis-à-vis du bruit en entrée $\mathbf{v}_p^-(t-r)$. Deux situations sont à considérer [CV97] :

- Si $\tilde{\mathbf{u}}$ est un bruit blanc de moyenne nulle, $\mathbf{u}_p^- \equiv \{\tilde{\mathbf{u}}_p^-, \mathbf{v}_p^-\}$ peut être utilisée comme variable instrumentale puisque celle-ci est décorrélée des perturbations présentes et futures et suffisamment corrélée avec $\tilde{\mathbf{y}}_f^+$. Elle permettra ainsi de fournir une estimation consistante du propagateur.
- Si $\tilde{\mathbf{u}}$ n'est pas un bruit blanc, le résultat précédent n'est plus vérifié si $\mathbf{v} \neq \mathbf{0}$. Aucune solution proposée dans la littérature n'est, à notre connaissance, applicable à notre problème. L'estimée obtenue dans ces conditions sera théoriquement asymptotiquement biaisée.

De cette simple analyse de corrélation découle donc la conclusion suivante : choisir comme variable instrumentale une version retardée des entrées (mesurées) du procédé est la solution qui fonctionnera dans la plus grande majorité des cas. Il est cependant fondamental d'être conscient *a priori* que les estimées obtenues dans le cas où $\mathbf{v} \neq \mathbf{0}$ et $\tilde{\mathbf{u}}$ n'est pas un bruit blanc seront inévitablement biaisées.

5.4.4 Synthèse

Les paragraphes précédents nous ont permis de présenter sept nouveaux algorithmes d'estimation de la matrice d'observabilité étendue du système à modéliser. Ces techniques ont en commun d'utiliser un opérateur linéaire particulier : le propagateur. Elles se différencient cependant par un certain nombre de caractéristiques propres, à savoir :

- l'origine du vecteur d'observation utilisé ;
- la forme du critère optimisé ;
- les outils et méthodes employés pour minimiser la fonction coût correspondante.

Ces spécificités sont regroupées au sein du tableau 5.5. Leur analyse permet de mettre en évidence trois familles de méthodes :

Famille 1 : l'algorithme RPM1 qui conduit à des estimées consistantes si et seulement si les perturbations agissant sur le procédé sont nulles :

Famille 2 : les algorithmes RPM2 et COPM qui utilisent les matrices de covariance des observations estimées par les méthodes RPIPO (RPIPO_{im} et RPIPO_{qr}) ;

Famille 3 : les algorithmes COIVPM, IVPM, EIVPM et EIV_{sqrTPM} qui associent aux vecteurs d'observation estimés à l'aide des techniques ROM (ROM_{im} et ROM_{qr}) une variable instru-

mentale permettant d'atténuer les effets des perturbations.

| Nom | Estimation de \mathbf{z}_f^+ | Critère | Outil ou algorithme de minimisation |
|-----------|--------------------------------|--|--|
| RPM1 | ROM | $\mathbb{E} \left\ \mathbf{z}_{f_2}^+ - \mathbf{P}_f^T \mathbf{z}_{f_1}^+ \right\ ^2$ | Moindres carrés récurifs |
| RPM2 | RPIPO | $\left\ \mathbf{R}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+} - \mathbf{P}_f^T \mathbf{R}_{\mathbf{z}_{f_1}^+} \right\ _F^2$ | Deux moindres carrés récurifs combinés |
| COPM | RPIPO | $\sum_{k=1}^t \lambda^{t-k} \left\ \mathbf{r}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+}^i(k) - \mathbf{P}_f^T(t) \mathbf{r}_{\mathbf{z}_{f_1}^+}^i(k) \right\ ^2$ pour $i \in [1, n_x]$ | Moindres carrés récurifs |
| COIVPM | ROM | $\sum_{k=1}^t \lambda^{t-k} \left\ \mathbf{r}_{\mathbf{z}_{f_2}^+ \xi}^i(k) - \mathbf{P}_f^T(t) \mathbf{r}_{\mathbf{z}_{f_1}^+ \xi}^i(k) \right\ ^2$ pour $i \in [1, n_\xi]$ | Variable instrumentale réursive |
| IVPM | ROM | $\left\ \mathbf{R}_{\mathbf{z}_{f_2}^+ \xi}(t) - \mathbf{P}_f^T(t) \mathbf{R}_{\mathbf{z}_{f_1}^+ \xi}(t) \right\ _F^2$ | Variable instrumentale réursive |
| EIVPM | ROM | $\left\ \mathbf{R}_{\mathbf{z}_{f_2}^+ \xi}(t) - \mathbf{P}_f^T(t) \mathbf{R}_{\mathbf{z}_{f_1}^+ \xi}(t) \right\ _F^2$ | Variable instrumentale réursive étendue |
| EIVsqrtPM | ROM | $\left\ \mathbf{R}_{\mathbf{z}_{f_2}^+ \xi}(t) - \mathbf{P}_f^T(t) \mathbf{R}_{\mathbf{z}_{f_1}^+ \xi}(t) \right\ _F^2$ | Rotations hyperboliques et rotations de Givens |

TAB. 5.5: Méthodes d'identification récurives des sous-espaces utilisant le propagateur.

Une étude pratique et théorique des principales propriétés de ces algorithmes doit désormais être menée. C'est ce que nous réalisons dans la suite de ce manuscrit.

5.4.5 Propriétés asymptotiques de la méthode du propagateur

L'objectif de cette section est de présenter les premières analyses des propriétés asymptotiques des méthodes d'identification récurives fondées sur le principe du propagateur. Cette étude, à la fois théorique et expérimentale, repose sur un examen comparé des algorithmes utilisant le propagateur et le projecteur. Ce choix se justifie par le fait que le critère du projecteur, en tout point équivalent au critère de Yang, possède des propriétés asymptotiques connues [Yan96] et qu'il constitue de fait une garantie de bon comportement.

5.4.5.1 Equivalence des critères PT et PM

Le premier résultat mis en évidence dans cette section concerne l'équivalence des critères PT et PM. Ce dernier fait l'objet de la proposition suivante :

Proposition 5.2. *Le critère du projecteur :*

$$\bar{V}(\mathbf{\Pi}) = \mathbb{E} \left\| \mathbf{z}_f^+ - \mathbf{\Pi} \mathbf{z}_f^+ \right\|^2 \quad (5.162)$$

avec :

$$\mathbf{\Pi} = \begin{bmatrix} \mathbf{I}_{n_x} \\ \mathbf{P}_f^T \end{bmatrix} (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} [\mathbf{I}_{n_x} \quad \mathbf{P}_f] \quad (5.163)$$

est une version pondérée du critère du propagateur :

$$J(\mathbf{P}_f) = \mathbb{E} \left\| \mathbf{z}_{f_2}^+ - \mathbf{P}_f^T \mathbf{z}_{f_1}^+ \right\|^2 \quad (5.164)$$

puisque :

$$\bar{V}(\mathbf{P}_f) = \mathbb{E} \left\| \left[\mathbf{I}_{n_{y,f}-n_x} - \mathbf{P}_f^T (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \mathbf{P}_f - (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \mathbf{P}_f \right] \left[\mathbf{z}_{f_2}^+ - \mathbf{P}_f^T \mathbf{z}_{f_1}^+ \right] \right\|^2. \quad (5.165)$$

De plus :

$$\bar{V}(\mathbf{P}_f) = \text{Tr} \left(\mathbf{R}_{\mathbf{z}_f^+} \right) - \text{Tr} \left(\mathbf{R}_{\mathbf{z}_f^+} \begin{bmatrix} \mathbf{I}_{n_x} \\ \mathbf{P}_f^T \end{bmatrix} (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} [\mathbf{I}_{n_x} \quad \mathbf{P}_f] \right). \quad (5.166)$$

La preuve de cette proposition est donnée ci-dessous.

Preuve : Supposons que le système étudié soit observable. Il existe alors une base telle que :

$$\mathbf{\Gamma}_f = \begin{bmatrix} \mathbf{I}_{n_x} \\ \mathbf{P}_f^T \end{bmatrix}. \quad (5.167)$$

Puisque $\mathbf{\Gamma}_f \in \mathbb{R}^{n_{y,f} \times n_x}$, $n_{y,f} > n_x$ est de rang plein, une factorisation QR réduite [GV96] peut lui être appliquée :

$$\begin{bmatrix} \mathbf{I}_{n_x} \\ \mathbf{P}_f^T \end{bmatrix} = \bar{\mathbf{Q}} \bar{\mathbf{R}} \quad (5.168)$$

avec $\bar{\mathbf{Q}} \in \mathbb{R}^{n_{y,f} \times n_x}$ et $\bar{\mathbf{R}} \in \mathbb{R}^{n_x \times n_x}$. Puisque $\bar{\mathbf{R}}$ est inversible :

$$\begin{bmatrix} \mathbf{I}_{n_x} \\ \mathbf{P}_f^T \end{bmatrix} \bar{\mathbf{R}}^{-1} = \bar{\mathbf{Q}}. \quad (5.169)$$

Comme $\bar{\mathbf{Q}}$ est une matrice orthogonale :

$$\bar{\mathbf{Q}}^T \bar{\mathbf{Q}} = \bar{\mathbf{R}}^{-T} (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T) \bar{\mathbf{R}}^{-1} = \mathbf{I}_{n_x} \quad (5.170)$$

et ainsi :

$$(\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T) = \bar{\mathbf{R}}^T \bar{\mathbf{R}}. \quad (5.171)$$

Ces équations permettent d'exprimer le projecteur $\mathbf{\Pi} = \bar{\mathbf{Q}} \bar{\mathbf{Q}}^T$ associé au sous-espace d'observabilité comme une fonction du propagateur \mathbf{P}_f comme suit :

$$\mathbf{\Pi} = \begin{bmatrix} \mathbf{I}_{n_x} \\ \mathbf{P}_f^T \end{bmatrix} (\bar{\mathbf{R}}^T \bar{\mathbf{R}})^{-1} [\mathbf{I}_{n_x} \quad \mathbf{P}_f]. \quad (5.172)$$

Ainsi, en utilisant la relation (5.171) :

$$\mathbf{\Pi} = \begin{bmatrix} \mathbf{I}_{n_x} \\ \mathbf{P}_f^T \end{bmatrix} (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} [\mathbf{I}_{n_x} \quad \mathbf{P}_f] = \begin{bmatrix} (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} & (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \mathbf{P}_f \\ \mathbf{P}_f^T (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} & \mathbf{P}_f^T (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \mathbf{P}_f \end{bmatrix}. \quad (5.173)$$

Il est alors possible d'écrire le critère :

$$\bar{V}(\mathbf{\Pi}) = \mathbb{E} \left\| \mathbf{z}_f^+ - \mathbf{\Pi} \mathbf{z}_f^+ \right\|^2 \quad (5.174)$$

comme une fonction du propagateur \mathbf{P}_f :

$$\begin{aligned}\bar{V}(\mathbf{P}_f) &= \mathbb{E} \left\| \begin{bmatrix} \mathbf{z}_{f_1}^+ \\ \mathbf{z}_{f_2}^+ \end{bmatrix} - \begin{bmatrix} (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} & (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \mathbf{P}_f \\ \mathbf{P}_f^T (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} & \mathbf{P}_f^T (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \mathbf{P}_f \end{bmatrix} \begin{bmatrix} \mathbf{z}_{f_1}^+ \\ \mathbf{z}_{f_2}^+ \end{bmatrix} \right\|^2 \\ &= \mathbb{E} \left\| \begin{bmatrix} \mathbf{z}_{f_1}^+ - (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \mathbf{z}_{f_1}^+ - (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \mathbf{P}_f \mathbf{z}_{f_2}^+ \\ \mathbf{z}_{f_2}^+ - \mathbf{P}_f^T (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \mathbf{z}_{f_1}^+ - \mathbf{P}_f^T (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \mathbf{P}_f \mathbf{z}_{f_2}^+ \end{bmatrix} \right\|^2.\end{aligned}\quad (5.175)$$

En appliquant le lemme d'inversion matricielle [GV96] au bloc $(\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1}$, le critère $\bar{V}(\mathbf{P}_f)$ devient :

$$\bar{V}(\mathbf{P}_f) = \mathbb{E} \left\| \begin{bmatrix} \mathbf{P}_f (\mathbf{I}_{n_y f - n_x} + \mathbf{P}_f^T \mathbf{P}_f)^{-1} \mathbf{P}_f^T \mathbf{z}_{f_1}^+ - (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \mathbf{P}_f \mathbf{z}_{f_2}^+ \\ \mathbf{z}_{f_2}^+ - \mathbf{P}_f^T \mathbf{z}_{f_1}^+ + \mathbf{P}_f^T \left[\mathbf{P}_f (\mathbf{I}_{n_y f - n_x} + \mathbf{P}_f^T \mathbf{P}_f)^{-1} \mathbf{P}_f^T \mathbf{z}_{f_1}^+ - (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \mathbf{P}_f \mathbf{z}_{f_2}^+ \right] \end{bmatrix} \right\|^2.\quad (5.176)$$

De plus, à partir du résultat obtenu à l'aide du lemme d'inversion matricielle, il est facile de montrer que :

$$\mathbf{P}_f (\mathbf{I}_{n_y f - n_x} + \mathbf{P}_f^T \mathbf{P}_f)^{-1} = (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \mathbf{P}_f.\quad (5.177)$$

Ainsi :

$$\begin{aligned}\bar{V}(\mathbf{P}_f) &= \mathbb{E} \left\| \begin{bmatrix} (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \mathbf{P}_f \left[\mathbf{P}_f^T \mathbf{z}_{f_1}^+ - \mathbf{z}_{f_2}^+ \right] \\ \mathbf{z}_{f_2}^+ - \mathbf{P}_f^T \mathbf{z}_{f_1}^+ + \mathbf{P}_f^T (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \mathbf{P}_f \left[\mathbf{P}_f^T \mathbf{z}_{f_1}^+ - \mathbf{z}_{f_2}^+ \right] \end{bmatrix} \right\|^2 \\ &= \mathbb{E} \left\| \left[\mathbf{I}_{n_y f - n_x} - \mathbf{P}_f^T (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \mathbf{P}_f - (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \mathbf{P}_f \right] \left[\mathbf{z}_{f_2}^+ - \mathbf{P}_f^T \mathbf{z}_{f_1}^+ \right] \right\|^2.\end{aligned}\quad (5.178)$$

Enfin, puisque :

$$\bar{V}(\mathbf{\Pi}) = \text{Tr} \left(\mathbf{R}_{z_f^+} \right) - \text{Tr} \left(\mathbf{R}_{z_f^+} \mathbf{\Pi} \right)\quad (5.179)$$

et que :

$$\mathbf{\Pi} = \begin{bmatrix} \mathbf{I}_{n_x} \\ \mathbf{P}_f^T \end{bmatrix} (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \begin{bmatrix} \mathbf{I}_{n_x} & \mathbf{P}_f \end{bmatrix},\quad (5.180)$$

le critère du projecteur peut être écrit comme suit :

$$\bar{V}(\mathbf{P}_f) = \text{Tr} \left(\mathbf{R}_{z_f^+} \right) - \text{Tr} \left(\mathbf{R}_{z_f^+} \begin{bmatrix} \mathbf{I}_{n_x} \\ \mathbf{P}_f^T \end{bmatrix} (\mathbf{I}_{n_x} + \mathbf{P}_f \mathbf{P}_f^T)^{-1} \begin{bmatrix} \mathbf{I}_{n_x} & \mathbf{P}_f \end{bmatrix} \right).\quad (5.181)$$

◀

La relation (5.165) indique que les critères PM et PT sont fortement corrélés. D'un point de vue pratique, l'équation (5.166) est cependant plus intéressante. Elle permet en effet de quantifier l'effet des perturbations sur le biais d'estimation. C'est ce que nous mesurons au sein du paragraphe suivant.

5.4.5.2 Méthode du propagateur sans variable instrumentale : biais asymptotique de l'estimateur

Nous avons montré, au sein du paragraphe 5.4.1.3, que la version « classique » du propagateur fournissait une estimée asymptotiquement biaisée en présence de perturbations et ce, même

au premier ordre (cf. equ. (5.99)). Ce n'est pas le cas du projecteur puisque, au premier ordre, et lorsque la matrice de covariance du bruit est proportionnelle à l'identité :

$$\text{Im}_{\text{col}}(\mathbf{\Pi}) \cong \text{Im}_{\text{col}}(\underline{\underline{\mathbf{U}}}_s) \quad (5.182)$$

où $\underline{\underline{\mathbf{U}}}_s$ contient les vecteurs colonnes du sous-espace d'observabilité.

Preuve : Par définition du projecteur :

$$\text{Im}_{\text{col}}(\mathbf{\Pi}) = \text{Im}_{\text{col}}(\underline{\underline{\mathbf{U}}}_s) \quad (5.183)$$

où $\underline{\underline{\mathbf{U}}}_s$ contient les n_x vecteurs propres dominants de $\mathbf{R}_{z_f^+}$:

$$\mathbf{R}_{z_f^+} = [\underline{\underline{\mathbf{U}}}_s \quad \underline{\underline{\mathbf{U}}}_b] \begin{bmatrix} \underline{\underline{\mathbf{\Lambda}}}_s & \mathbf{0} \\ \mathbf{0} & \underline{\underline{\mathbf{\Lambda}}}_b \end{bmatrix} \begin{bmatrix} \underline{\underline{\mathbf{V}}}_s^T \\ \underline{\underline{\mathbf{V}}}_b^T \end{bmatrix}. \quad (5.184)$$

Lorsque les données sont perturbées par un bruit blanc gaussien, cette matrice de covariance vérifie :

$$\mathbf{R}_{z_f^+} = \mathbf{\Gamma}_f \mathbf{R}_x \mathbf{\Gamma}_f^T + \sigma^2 \mathbf{I}_{n_y f}. \quad (5.185)$$

Appliquons à la DVP de $\mathbf{R}_{z_f^+}$ une analyse des perturbations [FV98]. En posant :

$$\mathbf{M}_f = \mathbf{\Gamma}_f \mathbf{R}_x \mathbf{\Gamma}_f^T \quad (5.186)$$

$$\Delta \mathbf{M}_f = \sigma^2 \mathbf{I}_{n_x}, \quad (5.187)$$

et en notant¹⁵ :

$$\mathbf{M}_f = [\underline{\underline{\mathbf{U}}}_s \quad \underline{\underline{\mathbf{U}}}_b] \begin{bmatrix} \underline{\underline{\mathbf{\Sigma}}}_s & \mathbf{0} \\ \mathbf{0} & \underline{\underline{\mathbf{\Sigma}}}_b \end{bmatrix} \begin{bmatrix} \underline{\underline{\mathbf{U}}}_s^T \\ \underline{\underline{\mathbf{U}}}_b^T \end{bmatrix}, \quad (5.188)$$

la décomposition en valeurs singulières de \mathbf{M}_f , la relation suivante est vérifiée [FV98] :

$$\text{Im}_{\text{col}}(\underline{\underline{\mathbf{U}}}_s) \cong \text{Im}_{\text{col}}\left(\underline{\underline{\mathbf{U}}}_s + \underline{\underline{\mathbf{U}}}_b \underline{\underline{\mathbf{U}}}_b^T \Delta \mathbf{M}_f \underline{\underline{\mathbf{U}}}_s \underline{\underline{\mathbf{\Sigma}}}_s^{-1}\right). \quad (5.189)$$

Ainsi, puisque $\Delta \mathbf{M}_f = \sigma^2 \mathbf{I}_{n_x}$:

$$\text{Im}_{\text{col}}(\mathbf{\Pi}) \cong \text{Im}_{\text{col}}\left(\underline{\underline{\mathbf{U}}}_s + \sigma^2 \underline{\underline{\mathbf{U}}}_b \underline{\underline{\mathbf{U}}}_b^T \underline{\underline{\mathbf{U}}}_s \underline{\underline{\mathbf{\Sigma}}}_s^{-1}\right). \quad (5.190)$$

Or, comme :

$$\underline{\underline{\mathbf{U}}}_b^T \underline{\underline{\mathbf{U}}}_s = \mathbf{0} \quad (5.191)$$

la relation (5.190) devient au premier ordre :

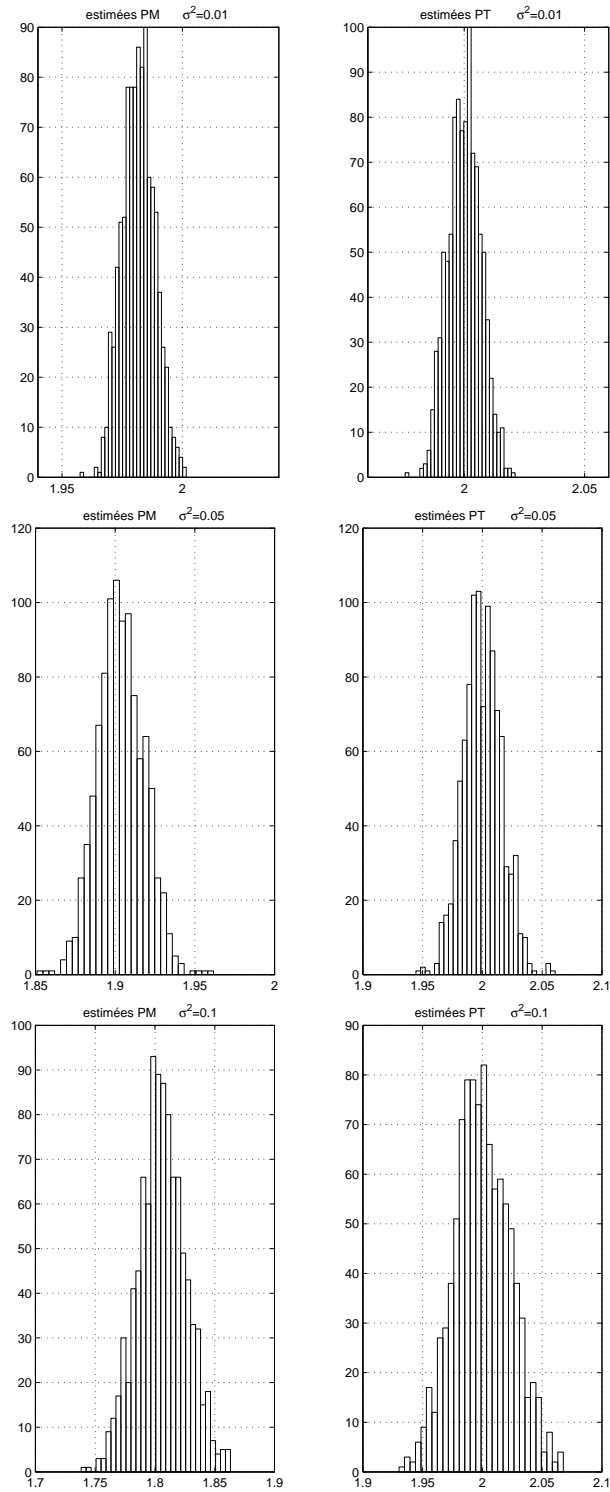
$$\text{Im}_{\text{col}}(\mathbf{\Pi}) \cong \text{Im}_{\text{col}}(\underline{\underline{\mathbf{U}}}_s). \quad (5.192)$$

Cette dernière équation montre que le projecteur n'est pas biaisé au premier ordre. ◀

Il est néanmoins intéressant de quantifier l'altération des performances due au bruit en appliquant la version du propagateur sans variable instrumentale sur l'exemple suivant :

$$\mathbf{z}_f^+(t) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \mathbf{x}(t) + \mathbf{n}_f^+(t) \quad (5.193)$$

où \mathbf{x} et \mathbf{n}_f^+ sont deux bruits blancs indépendants de variance respective 1 et σ^2 .

FIG. 5.1: Histogrammes des estimées de \mathbf{P}_f en utilisant PM et PT.

Une première étude, fondée sur une simulation de Monte Carlo (500 échantillons, 1000 itérations), est réalisée afin de comparer la distribution de l'estimée \mathbf{P}_f obtenue par les deux approches pour différentes valeurs de σ^2 . Les histogrammes de la figure 5.1 montrent clairement que, bien que l'erreur d'estimation du propagateur soit manifeste et que celle-ci augmente avec σ^2 , les distributions obtenues avec PT et PM sont proches. Sur cet exemple, les estimées obtenues avec PT et PM possèdent donc des propriétés du second ordre très similaires.

La seconde analyse repose sur une quantification plus précise de l'altération des performances due au bruit. Nous mesurons l'impact du bruit sur l'estimée $\hat{\mathbf{P}}_f$ en calculant, pour trois valeurs de σ^2 , le critère $\bar{V}(\hat{\mathbf{P}}_f)$ (cf. equ. (5.166)) normalisé sachant que (cf. equ. (5.94)) :

$$\hat{\mathbf{P}}_f^T = \frac{2}{1 + \sigma^2}. \quad (5.194)$$

Les résultats présentés dans le tableau 5.6 montrent que la perte de performance est assez faible,

| Variance du bruit σ^2 | $\frac{\bar{V}(\hat{\mathbf{P}}_f) - \bar{V}(\mathbf{P}_f = 2)}{\bar{V}(\mathbf{P}_f = 2)}$ |
|------------------------------|---|
| 0.01 | 0.008 |
| 0.05 | 0.0392 |
| 0.1 | 0.0768 |

TAB. 5.6: Comparaison des fonctions coût optimales obtenues avec le projecteur et le propagateur.

puisque, même en présence de bruit, le sous-espace estimé à partir du propagateur correspond, relativement à la variance du bruit, à celui obtenu via le projecteur.

5.4.5.3 Consistance de la méthode du propagateur avec variable instrumentale

Bien que la méthode du propagateur sans variable instrumentale fournisse nécessairement une estimée asymptotiquement biaisée de \mathbf{P}_f , il est possible de vérifier que la version avec variable instrumentale est consistante quel que soit le nombre d'instruments utilisés. En effet :

Proposition 5.3. *Les algorithmes du propagateur avec variable instrumentale conduisent à des estimées asymptotiquement non biaisées de \mathbf{P}_f si et seulement si :*

- le système étudié est causal, asymptotiquement stable et observable ;
- l'entrée vérifie la condition d'excitation persistante (cf. §3.3.5) ;
- la variable instrumentale $\boldsymbol{\xi}(t)$ et le bruit $\mathbf{n}(s)$ sont décorrélés pour tous t et s .

Preuve : Considérons le critère quadratique suivant :

$$J_{IV}(\hat{\mathbf{P}}_f) = \left\| \mathbf{R}_{z_2^+} \boldsymbol{\xi} - \hat{\mathbf{P}}_f^T \mathbf{R}_{z_1^+} \boldsymbol{\xi} \right\|_F^2. \quad (5.195)$$

¹⁵Rappelons que la matrice \mathbf{M}_f est symétrique.

La solution au sens des moindres carrés d'une telle fonction coût est donnée par :

$$\hat{\mathbf{P}}_f^T(t) = \mathbf{R}_{z_{f_2}^+} \boldsymbol{\xi}(t) \mathbf{R}_{z_{f_1}^+}^{-1} \boldsymbol{\xi}^T(t) \quad (5.196)$$

lorsque $n_\xi = n_{z_{f_1}^+}$,

$$\hat{\mathbf{P}}_f^T(t) = \mathbf{R}_{z_{f_2}^+} \boldsymbol{\xi}(t) \mathbf{R}_{z_{f_1}^+}^T \boldsymbol{\xi}^T(t) \left(\mathbf{R}_{z_{f_1}^+} \boldsymbol{\xi}(t) \mathbf{R}_{z_{f_1}^+}^T \boldsymbol{\xi}^T(t) \right)^{-1} \quad (5.197)$$

lorsque $n_\xi > n_{z_{f_1}^+}$. Deux cas sont alors à considérer.

– Lorsque $n_\xi = n_{z_{f_1}^+}$, puisque :

$$\mathbf{z}_{f_2}^+(t) = \mathbf{P}_f^T(t) \mathbf{z}_{f_1}^+(t) + \mathbf{n}_{f_2}^+(t) - \mathbf{P}_f^T(t) \mathbf{n}_{f_1}^+(t), \quad (5.198)$$

il est facile de vérifier que :

$$\begin{aligned} \hat{\mathbf{P}}_f^T(t) &= \left(\frac{1}{t} \sum_{k=1}^t \left[\mathbf{P}_f^T(t) \mathbf{z}_{f_1}^+(k) + \mathbf{n}_{f_2}^+(k) - \mathbf{P}_f^T(t) \mathbf{n}_{f_1}^+(k) \right] \boldsymbol{\xi}^T(k) \right) \left(\frac{1}{t} \sum_{k=1}^t \mathbf{z}_{f_1}^+(k) \boldsymbol{\xi}^T(k) \right)^{-1} \\ &= \mathbf{P}_f^T(t) \left(\frac{1}{t} \sum_{k=1}^t \mathbf{z}_{f_1}^+(k) \boldsymbol{\xi}^T(k) \right) \left(\frac{1}{t} \sum_{k=1}^t \mathbf{z}_{f_1}^+(k) \boldsymbol{\xi}^T(k) \right)^{-1} \\ &\quad + \left(\frac{1}{t} \sum_{k=1}^t \left[\mathbf{n}_{f_2}^+(k) - \mathbf{P}_f^T(t) \mathbf{n}_{f_1}^+(k) \right] \boldsymbol{\xi}^T(k) \right) \left(\frac{1}{t} \sum_{k=1}^t \mathbf{z}_{f_1}^+(k) \boldsymbol{\xi}^T(k) \right)^{-1}. \end{aligned} \quad (5.199)$$

En supposant que la variable instrumentale soit décorrélée du bruit \mathbf{n}_f^+ et que le signal d'entrée soit suffisamment riche pour que $\mathbf{R}_{z_{f_1}^+} \boldsymbol{\xi}$ soit inversible, l'estimation du propagateur est consistante puisque :

$$\hat{\mathbf{P}}_f^T(t) = \mathbf{P}_f^T(t) \quad (5.200)$$

– Lorsque $n_\xi > n_{z_{f_1}^+}$, la relation (5.197) s'écrit :

$$\begin{aligned} \hat{\mathbf{P}}_f^T(t) &= \left(\frac{1}{t} \sum_{k=1}^t \left[\mathbf{P}_f^T(t) \mathbf{z}_{f_1}^+(k) + \mathbf{n}_{f_2}^+(k) - \mathbf{P}_f^T(t) \mathbf{n}_{f_1}^+(k) \right] \boldsymbol{\xi}^T(k) \right) \left(\frac{1}{t} \sum_{k=1}^t \mathbf{z}_{f_1}^+(k) \boldsymbol{\xi}^T(k) \right)^T \\ &\quad \left(\left(\frac{1}{t} \sum_{k=1}^t \mathbf{z}_{f_1}^+(k) \boldsymbol{\xi}^T(k) \right) \left(\frac{1}{t} \sum_{k=1}^t \mathbf{z}_{f_1}^+(k) \boldsymbol{\xi}^T(k) \right)^T \right)^{-1} \\ &= \mathbf{P}_f^T(t) \left(\frac{1}{t} \sum_{k=1}^t \mathbf{z}_{f_1}^+(k) \boldsymbol{\xi}^T(k) \right) \left(\frac{1}{t} \sum_{k=1}^t \mathbf{z}_{f_1}^+(k) \boldsymbol{\xi}^T(k) \right)^T \\ &\quad \left(\left(\frac{1}{t} \sum_{k=1}^t \mathbf{z}_{f_1}^+(k) \boldsymbol{\xi}^T(k) \right) \left(\frac{1}{t} \sum_{k=1}^t \mathbf{z}_{f_1}^+(k) \boldsymbol{\xi}^T(k) \right)^T \right)^{-1} \\ &\quad + \left(\frac{1}{t} \sum_{k=1}^t \left[\mathbf{n}_{f_2}^+(k) - \mathbf{P}_f^T(t) \mathbf{n}_{f_1}^+(k) \right] \boldsymbol{\xi}^T(k) \right) \left(\frac{1}{t} \sum_{k=1}^t \mathbf{z}_{f_1}^+(k) \boldsymbol{\xi}^T(k) \right)^T \\ &\quad \left(\left(\frac{1}{t} \sum_{k=1}^t \mathbf{z}_{f_1}^+(k) \boldsymbol{\xi}^T(k) \right) \left(\frac{1}{t} \sum_{k=1}^t \mathbf{z}_{f_1}^+(k) \boldsymbol{\xi}^T(k) \right)^T \right)^{-1}. \end{aligned} \quad (5.201)$$

L'estimation du propagateur est alors consistante si les conditions précédentes sont vérifiées. ◀

5.4.5.4 Propriétés du second ordre de la méthode du propagateur avec variable instrumentale

Le paragraphe précédent nous a permis de montrer que les versions du propagateur avec variable instrumentale conduisent à des estimées asymptotiquement non biaisées de \mathbf{P}_f (et donc du sous-espace d'observabilité). L'introduction d'une variable instrumentale a donc permis de rendre les algorithmes IVPM et EIVPM équivalents (au premier ordre) aux méthodes fondées sur le critère de Yang. Malheureusement, l'étude des propriétés du second ordre des estimateurs basés sur IVPM et EIVPM n'est pas si aisée, et, actuellement, l'unique caution que nous puissions apporter repose sur une simulation de Monte Carlo. Pour cela, nous considérons de nouveau le système :

$$\mathbf{z}_f^+(t) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \mathbf{x}(t) + \mathbf{n}_f^+(t) \quad (5.202)$$

où \mathbf{n}_f^+ est désormais obtenu à l'aide de l'équation suivante :

$$\mathbf{n}_f^+(t+1) = 0.5\mathbf{n}_f^+(t) + \mathbf{e}_f^+(t) \quad (5.203)$$

où \mathbf{e}_f^+ est un bruit blanc gaussien de variance 0.1. Le propagateur est estimé en utilisant les algorithmes IVPM et IVPT. La variable instrumentale est une version filtrée de la séquence \mathbf{x} . Après 1000 itérations de 200 échantillons chacune, les histogrammes des distributions des estimées sont tracés. Il est alors évident (cf. fig. 5.2) que les deux approches sont équivalentes.

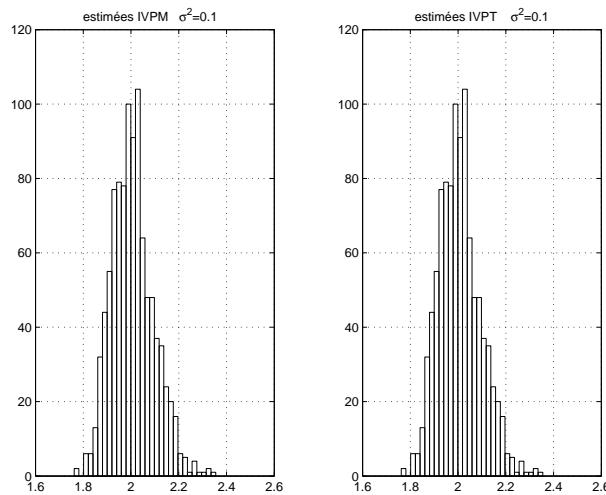


FIG. 5.2: Histogrammes des estimées de \mathbf{P}_f en utilisant IVPM et IVPT.

5.4.6 Complexité numérique des algorithmes d'estimation de la matrice d'observabilité

L'un des objectifs majeurs des techniques récursives d'identification est de pouvoir mettre à jour le modèle du procédé au cours de son utilisation, à chaque fois que de nouvelles données sont mesurées. Cette contrainte nécessite donc que chaque itération des algorithmes proposés dure le moins de temps possible afin d'être applicable, en temps réel, sur des procédés rapides.

Afin de s'assurer que les méthodes développées jusqu'ici vérifient cette condition, il est essentiel de comparer les complexités des algorithmes présentés au sein de ce manuscrit. Le problème de mise à jour du vecteur d'observation a déjà été étudié dans le paragraphe 5.3.3. Nous nous intéressons donc ici uniquement aux algorithmes d'estimation récursive de la matrice d'observabilité. Nous étudierons plus précisément la complexité cumulée de ces deux étapes au sein du paragraphe 5.5.

| Algorithme | Complexité |
|------------------------|---|
| COPAST | $2(n_y f)^2 + 3n_x n_y f + 2n_x^2 + n_x$ |
| COPASTd | $2(n_y f)^2 + 3n_x n_y f + n_x$ |
| COIVPAST | $n_\xi n_y f + 3n_x n_y f + 2n_x^2 + n_x$ |
| COIVPASTd | $n_\xi n_y f + 3n_x n_y f + n_x$ |
| Yang _{grad} | $3n_x (n_y f)^2 + 2(n_y f)^2$ |
| EIVPAST | $2n_\xi n_y f + 2n_\xi n_x + 5n_x n_y f + n_\xi + 4n_x^2 + 8n_x$ |
| UDPAST | $6n_x n_y f + 4n_x^2 + 2n_x$ |
| IVPT | $(n_y f)^3 + n_\xi (n_y f)^2 + 2n_x (n_y f)^2 + n_\xi n_y f - 2n_x n_y f - n_y f$ |
| COPM | $2(n_y f)^2 + 2n_x n_y f + n_x$ |
| COIVPM | $n_\xi n_y f + 2n_x n_y f + n_x$ |
| RPM2 | $4n_x n_y f + 2n_x$ |
| EIVPM | $2n_\xi n_y f + 4n_x n_y f + n_x n_\xi + n_\xi + 8n_x$ |
| EIV _{sqrt} PM | $2n_\xi n_y f + 4n_x n_y f + n_\xi + 2n_x (n_x + 3)^3$ |

TAB. 5.7: Comparaison des complexités numériques de plusieurs algorithmes d'estimation récursive de la matrice d'observabilité.

La première phase de cette analyse consiste à calculer le nombre de multiplications utilisées par chaque algorithme d'estimation récursive de la matrice d'observabilité exposé au sein de ce manuscrit¹⁶. Les résultats de ces calculs sont présentés dans le tableau 5.4.6. La première constatation qui résulte de cette comparaison est que la gamme des complexités est assez large. La plus faible est en effet $\mathcal{O}(4n_x n_y f)$ (RPM2), la plus élevée $\mathcal{O}((n_y f)^3)$ (IVPT). Une analyse un peu plus approfondie permet de regrouper ces algorithmes en trois familles :

- les algorithmes dont le coût numérique est proportionnel à $n_y f$: UDPAST et RPM2 ;
- les algorithmes dont le coût numérique est proportionnel à $(n_y f)^2$: COPAST, COPASTd, COIVPAST, COIVPASTd, COPM, COIVPM, EIVPAST, EIVPM, EIV_{sqrt}PM et Yang_{grad} ;
- les algorithmes dont le coût numérique est proportionnel à $(n_y f)^3$: IVPT.

Ce regroupement a nécessité de supposer que $n_\xi = n_y f$. Ce choix se justifie en pratique puisque :

¹⁶A titre de comparaison, la décomposition en valeurs singulières utilisée par l'algorithme Ordinary MOESP demande environ $n_y n_u^2 f^3$ multiplications et celle de l'algorithme PI/PO MOESP $n_\xi p n_y^2 f^2$.

- en règle générale, les instruments composant la variable instrumentale sont des données d'entrée ou de sortie retardées donc $n_\xi \cong n_y p$;
- par souci de simplicité, on choisit habituellement $p = f$.

Une analyse simplifiée de ces résultats indique que les algorithmes les plus intéressants d'un point de vue calculatoire sont ceux formant la première famille. Il est cependant important de pondérer cette première conclusion puisque :

- les algorithmes UDPAST et RPM2 utilisent la matrice des observations $\mathbf{z}_f^+ \mathbf{z}_f^{+T}$ dont l'estimation (cf. §5.3.2) demande un nombre de multiplications supérieur à celle du vecteur d'observation bruité (cf. §5.3.1) ;
- les algorithmes COIVPAST, COIVPASTd, COIVPM, EIVPAST, EIVPM et EIV_{sqr}tPM font au contraire directement appel au vecteur d'observation dont le calcul demande un coût numérique plus faible.

Cette première analyse nécessite donc d'être complétée par une étude combinée des deux phases des algorithmes d'identification proposés.

5.5 Complexité numérique des algorithmes d'identification récursive : une étude de cas

Bien que nous ayons calculé les complexités théoriques de toutes les méthodes proposées pour estimer récursivement le vecteur d'observation et la matrice d'observabilité (cf. §5.3.3 et 5.4.6), l'analyse directe du regroupement de ces valeurs n'est pas une chose facile. Nous proposons donc de les comparer sur un exemple de simulation simple : un système monovarié d'ordre 2 défini comme suit [MLL04] :

$$\mathbf{x}(t+1) = \begin{bmatrix} 0.7 & 0 \\ 0 & -0.5 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 2 \\ 1 \end{bmatrix} \tilde{u}(t) + \mathbf{w}(t) \quad (5.204a)$$

$$\tilde{y}(t) = [1 \quad 2] \mathbf{x}(t) + 0.05\tilde{u}(t) \quad (5.204b)$$

$$u(t) = \tilde{u}(t) + v(t) \quad (5.204c)$$

$$y(t) = \tilde{y}(t) + \frac{1}{1+0.8q^{-1}}e(t). \quad (5.204d)$$

L'entrée non bruitée \tilde{u} est un bruit blanc gaussien de moyenne nulle et de variance unitaire contaminée par un bruit blanc gaussien de moyenne nulle et de variance 0.1. Les bruits de procédé et de mesure \mathbf{w} et e sont également deux bruits blancs gaussiens de moyenne nulle et de variance 0.1.

Trois études successives vont être réalisées afin de comparer les différentes techniques d'estimation exposées au sein de ce manuscrit. Il est essentiel de préciser que tous les algorithmes utilisés ont fourni des estimées de la matrice d'observabilité du système proches des vraies valeurs.

La première partie de cette étude concerne l'étape d'estimation du vecteur d'observation.

Pour cela, prenons comme variable instrumentale :

$$\xi = \begin{bmatrix} \mathbf{u}(t-p) \\ \vdots \\ \mathbf{u}(t-1) \end{bmatrix} \quad (5.205)$$

et fixons les paramètres suivants :

$$n_u = 1 \quad (5.206)$$

$$n_y = 1 \quad (5.207)$$

$$n_x = 2 \quad (5.208)$$

$$f = p = 8. \quad (5.209)$$

Les complexités théoriques des algorithmes d'estimation des vecteurs d'observation sont présentées dans le tableau¹⁷ 5.8. Ces premiers chiffres indiquent que, sur cet exemple,

| Algorithme | Complexité (nombre de multiplications) | Complexité adimensionnée |
|---------------------|--|--------------------------|
| ROM _{im} | 264 | 1 |
| RPIPO _{im} | 656 | 2.34 |
| ROM _{qr} | 432 | 1.74 |
| RPIPO _{qr} | 1120 | 4.32 |

TAB. 5.8: Complexités théoriques et pratiques des algorithmes d'estimation du vecteur d'observation.

les algorithmes fondés sur la mise à jour des factorisations QR des schémas d'identification Ordinary MOESP et PI/PO MOESP nécessitent deux fois plus de multiplications que leurs homologues utilisant le lemme d'inversion matricielle. Ceci se vérifie en pratique (cf. tab. 5.8) puisque, en appliquant 100 fois de suite ces quatre algorithmes sur 2000 données d'entrée-sortie, il faut environ deux fois moins de temps pour réaliser cette opération avec ROM_{im} (resp. RPIPO_{im}) qu'avec ROM_{qr} (resp. RPIPO_{qr}).

La deuxième étape de cette étude concerne la phase d'estimation de la matrice d'observabilité. En appliquant de nouveau un tirage de Monte Carlo et en analysant conjointement les valeurs théoriques calculées et les complexités pratiques mesurées (cf. tab. 5.9), nous pouvons conclure que :

- les classements des algorithmes sont identiques en pratique et en théorie bien que certaines proportions ne soient pas respectées ;
- dans tous les cas, les algorithmes fondés sur le propagateur sont plus rapides que leurs homologues directs utilisant PAST ;
- la robustesse numérique apportée par EIV_{sqr}tPM se paye au détriment d'une complexité numérique bien supérieure à EIVPM.

La dernière phase de cette analyse consiste à comparer les complexités numériques de toutes les techniques récursives des sous-espaces présentées dans ce rapport en cumulant le nombre de multiplications nécessaires à chacune des deux étapes les composant. Par souci d'homogénéité et afin de ne pas multiplier les combinaisons d'algorithmes, il a été décidé de n'utiliser que les méthodes ROM_{qr} et RPIPO_{qr} pour estimer le vecteur d'observation. Ces techniques peuvent être en effet considérées comme des algorithmes de référence [VD91, LV98, Lov98, LGV00, MLL04].

Cette comparaison montre que :

¹⁷Le calcul des complexités adimensionnées est réalisé en prenant comme référence le plus petit coût numérique.

| Algorithme | Complexité théorique | Complexité pratique adimensionnée |
|------------------------|----------------------|-----------------------------------|
| COPAST | 186 | 1.45 |
| COPASTd | 178 | 1.42 |
| COPM | 162 | 1.31 |
| COIVPAST | 122 | 1.19 |
| COIVPASTd | 114 | 1.15 |
| COIVPM | 98 | 1.08 |
| EIVPAST | 280 | 2.53 |
| EIVPM | 232 | 2.23 |
| EIV _{sqr} tPM | 700 | 7.77 |
| UDPAST | 116 | 1.17 |
| RPM2 | 90 | 1 |
| IVPT | 1304 | 10.9 |
| Yang _{grad} | 512 | 7.11 |

TAB. 5.9: Comparaison des complexités pratiques et théoriques de plusieurs algorithmes d'estimation récursive de la matrice d'observabilité.

- la combinaison la plus rapide est COIVPM+ROM_{qr} ;
- à l'inverse, la plus lente est IVPT+ROM_{qr} ;
- à l'exception de cette dernière combinaison, les algorithmes les plus rapides sont ceux qui utilisent le vecteur d'observation obtenu à l'aide de ROM_{qr} ;
- la phase d'estimation du vecteur d'observation est, en règle général, l'étape dominante dans le calcul des complexités numériques.

Remarque 5.8. *La position relative de l'algorithme EIV_{sqr}tPM+ROM_{qr} vis-à-vis des autres algorithmes est différente en pratique et en théorie. Cette particularité est vraisemblablement due à notre implémentation des rotations de Givens utilisées par EIV_{sqr}tPM (appels de fonctions).*

Cette première étude pratique indique que, d'un point de vue coût calculatoire, l'algorithme le plus intéressant est COIVPM associé à ROM. Ces différentes expériences nous ont permis de constater que, sur cet exemple, il était possible d'atteindre des fréquences d'échantillonnage de l'ordre¹⁸ de $5kHz$. Il faut cependant poursuivre cette analyse en étudiant la consistante et la fiabilité des estimées fournies par ces différentes méthodes. Cet approfondissement sera développé au sein du chapitre 6.

5.6 Estimation des matrices d'état

L'estimation récursive d'une base de la matrice d'observabilité n'est pas une fin en soi. L'objectif des méthodes récursives des sous-espaces est en effet d'estimer en ligne une représentation d'état du système, i.e. de mettre à jour, à chaque nouvelle acquisition, les matrices d'état du procédé. Il est donc nécessaire de proposer des algorithmes récursifs d'estimation des matrices d'état. Puisque les méthodes d'identification récursive des sous-espaces proposées au sein de

¹⁸Cette valeur est bien évidemment relative aux capacités de l'ordinateur employé. Pour notre part, nous avons utilisé un Pentium IV 2.4GHz avec 512Mo de RAM.

| Algorithme | Complexité théorique | Complexité pratique adimensionnée |
|--|----------------------|-----------------------------------|
| COPAST + RPIPO _{qr} | 1306 | 2.16 |
| COPASTd + RPIPO _{qr} | 1298 | 2.15 |
| COPM + RPIPO _{qr} | 1282 | 2.13 |
| COIVPAST + ROM _{qr} | 554 | 1.06 |
| COIVPASTd + ROM _{qr} | 546 | 1.05 |
| COIVPM + ROM _{qr} | 530 | 1 |
| EIVPAST + ROM _{qr} | 712 | 1.47 |
| EIVPM + ROM _{qr} | 644 | 1.38 |
| EIV _{sqr} tPM + ROM _{qr} | 1132 | 3.07 |
| UDPAST + RPIPO _{qr} | 1236 | 2.11 |
| RPM2 + RPIPO _{qr} | 1210 | 2.08 |
| IVPT + ROM _{qr} | 1736 | 4.07 |
| Yang _{grad} + RPIPO _{qr} | 1632 | 3.92 |

TAB. 5.10: Comparaison des complexités pratiques et théoriques de plusieurs algorithmes d'estimation récursive des sous-espaces.

ce manuscrit (et dans la littérature) ne cherchent en aucun cas à estimer le vecteur d'état du système, les techniques permettant d'avoir accès en ligne aux matrices d'état du système utilisent exclusivement la matrice d'observabilité estimée et les données d'entrée-sortie mesurées. Deux étapes consécutives sont ainsi développées :

1. Comme dans le cas hors ligne, l'estimation des matrices \mathbf{A} et \mathbf{C} est réalisée en utilisant la propriété d'A-invariance de $\mathbf{\Gamma}_f$:

$$\hat{\mathbf{C}} = \hat{\mathbf{\Gamma}}_f(1 : n_y, :) \quad (5.210)$$

$$\hat{\mathbf{A}} = \left\{ \hat{\mathbf{\Gamma}}_f^\uparrow \right\}^\dagger \hat{\mathbf{\Gamma}}_f^\downarrow. \quad (5.211)$$

2. La mise à jour des matrices $\hat{\mathbf{B}}$ et $\hat{\mathbf{D}}$ est obtenue à partir de l'erreur de simulation :

$$\mathbf{y}(t) = \sum_{k=0}^{t-1} \hat{\mathbf{C}} \hat{\mathbf{A}}^{t-k-1} \hat{\mathbf{B}} \mathbf{u}(k) + \hat{\mathbf{D}} \mathbf{u}(t) + \mathbf{n}(t) \quad (5.212)$$

qui peut être facilement réécrite à l'aide de la régression linéaire suivante (cf. §3.3.4.2) :

$$\mathbf{y}(t) = \boldsymbol{\phi}^T(t) \boldsymbol{\eta} + \mathbf{n}(t) \quad (5.213)$$

où :

$$\boldsymbol{\phi}^T(t) = [\boldsymbol{\phi}_1^T(t) \quad \boldsymbol{\phi}_2^T(t)] = [\mathbf{u}^T(t) \otimes \mathbf{I}_{n_y} \quad -(\sum_{k=0}^{t-1} \mathbf{u}^T(k) \otimes \hat{\mathbf{C}} \hat{\mathbf{A}}^{t-k-1})] \quad (5.214)$$

$$\boldsymbol{\eta} = \begin{bmatrix} \text{vec}(\mathbf{D}) \\ \text{vec}(\mathbf{B}) \end{bmatrix}. \quad (5.215)$$

Sachant de plus que [McK95] :

$$\boldsymbol{\phi}_2^T(t) = \boldsymbol{\phi}_2^T(t-1) (\mathbf{I}_{n_u} \otimes \hat{\mathbf{A}}) + (\mathbf{u}^T(t-1) \otimes \hat{\mathbf{C}}). \quad (5.216)$$

L'estimation du vecteur de paramètres $\boldsymbol{\eta}$ peut alors être réalisée en utilisant, par exemple, un moindres carrés récursif classique. L'application de la méthode de la variable instrumentale récursive peut également être envisagée puisque, théoriquement, le bruit et le régresseur sont asymptotiquement décorrés [Gus97a].

Ces étapes possèdent respectivement des complexités en $\mathcal{O}(n_y f n_x^2)$ et $\mathcal{O}(n_u^3 (n_x + n_y)^3)$ [Lov98].

5.7 Conclusion

Dans ce chapitre, de nouveaux algorithmes d'identification récursive des sous-espaces ont été développés. Ces derniers se décomposent en trois étapes consécutives :

1. la mise à jour du vecteur d'observation à chaque nouvelle acquisition de données d'entrée-sortie ;
2. l'estimation d'une base de la matrice d'observabilité en mettant à jour récursivement un opérateur linéaire particulier nommé propagateur ;
3. l'estimation des matrices d'état du système.

Toutes ces techniques ont en commun d'être fondées sur l'adaptation, au problème d'identification récursive, d'une méthode particulière de traitement d'antennes : la méthode du propagateur. Cet ajustement, combiné à l'introduction d'une variable instrumentale, permet d'éviter d'utiliser une décomposition en valeurs singulières pour séparer le sous-espace bruit du sous-espace signal dans un contexte stochastique. Deux hypothèses sont nécessaires pour estimer de manière consistante la matrice d'observabilité du système :

- connaître *a priori* l'ordre du système ;
- identifier un système stable.

Ces dernières sont en effet essentielles à la construction des critères quadratiques, sans contrainte et sans approximation utilisés pour estimer le propagateur. Une étude des premières propriétés asymptotiques de ces algorithmes a montré que la méthode du propagateur avec variable instrumentale était consistante si et seulement si :

- le système étudié est causal, asymptotiquement stable et observable ;
- l'entrée vérifie la condition d'excitation persistante ;
- la variable instrumentale $\boldsymbol{\xi}(t)$ et le bruit $\mathbf{n}(s)$ sont décorrés pour tous t et s .

Cette analyse a également permis de révéler, tout au moins sur un exemple, que les propriétés du premier et du second ordre de la méthode du propagateur avec variable instrumentale sont équivalentes à celles des techniques utilisant l'approche de Yang. Une analyse comparée de la complexité des algorithmes a finalement été exposée. Cette dernière démontre que les méthodes utilisant le propagateur sont, dans leur globalité, plus rapides que leurs homologues utilisant l'approximation PAST.

Bien que plusieurs exemples de simulation aient déjà été présentés dans ce chapitre, il est important de proposer au lecteur une analyse critique des performances des techniques d'identification récursive fondées sur le propagateur à la fois sur des exemples de simulation plus complexes mais également sur des données réelles. Ce travail est exposé au sein du chapitre suivant.

Chapitre 6

Applications

Sommaire

| | |
|---|------------|
| 6.1 Exemples de simulation | 140 |
| 6.1.1 Étude d'un système invariant. Comparaison des algorithmes | 140 |
| 6.1.2 Étude d'un système variant | 153 |
| 6.1.3 Conclusion | 162 |
| 6.2 Étude de systèmes réels | 162 |
| 6.2.1 Étude d'un système de transmission flexible | 162 |
| 6.2.2 Étude d'un système électronique | 169 |
| 6.3 Conclusion | 180 |

Les études de complexité réalisées à la fin du chapitre 5 (cf. §5.4.6 et 5.5) ont permis de mettre en avant le faible coût calculatoire des algorithmes fondés sur le propagateur. Ces méthodes présentent également des qualités théoriques intéressantes puisqu'elles reposent exclusivement sur des critères quadratiques, sans contrainte et sans approximation. Bien que ces nombreux bénéfices les conduisent naturellement à être considérées comme une très bonne alternative aux techniques récursives de sous-espaces utilisant le critère de Yang et ses dérivées (cf. chapitre 4), il est fondamental de vérifier que ces algorithmes fournissent des estimées tout au moins aussi fiables que celles obtenues en appliquant les méthodes d'identification employant l'approximation PAST. Ce chapitre a pour objectif de mettre en évidence les performances des techniques d'estimation récursive utilisant le propagateur. L'approche envisagée repose sur :

- une étude à l'aide de données simulées ;
- une analyse sur données réelles (système mécanique et électronique).

Ces examens vont ainsi permettre :

- de comparer les algorithmes d'identification PAST et PM ;
- de caractériser l'influence des paramètres de construction n_x , λ , f et p nécessaires à toute méthode d'identification récursive des sous-espaces ;
- de mettre en évidence les capacités de poursuite de ces techniques adaptatives.

6.1 Exemples de simulation

Considérons le système linéaire, causal et invariant d'ordre 3 suivant¹ :

$$\mathbf{x}(t+1) = \begin{bmatrix} 0.8 & -0.4 & 0.2 \\ 0 & 0.3 & -0.5 \\ 0 & 0 & 0.5 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 & 0 \\ 0 & -0.6 \\ 0.5 & 0 \end{bmatrix} \tilde{\mathbf{u}}(t) + \begin{bmatrix} 0.055 \\ 0.04 \\ 0.045 \end{bmatrix} v(t) \quad (6.1a)$$

$$\tilde{\mathbf{y}}(t) = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0.025 \\ 0.03 \end{bmatrix} w(t) \quad (6.1b)$$

$$\mathbf{u}(t) = \tilde{\mathbf{u}}(t) + \mathbf{v}(t) \quad (6.1c)$$

où $\tilde{\mathbf{u}} \in \mathbf{R}^{2 \times 1}$ et $\tilde{\mathbf{y}} \in \mathbf{R}^{2 \times 1}$ sont les entrées et sorties du système sans bruit de mesure, w le bruit d'état du système et $\{\mathbf{v}, v\}$ les bruits de mesure en entrée et en sortie du procédé (cf. fig. 6.1).

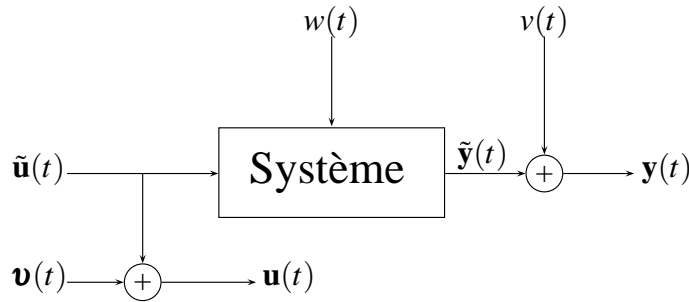


FIG. 6.1: Représentation schématique du système étudié.

L'objectif de cette section consiste à mettre en évidence les performances des algorithmes fondés sur le propagateur (désignés par l'acronyme PM dans la suite de ce chapitre) en appliquant ces derniers à l'exemple de simulation décrit précédemment. Plusieurs situations vont être successivement examinées :

- la première étape se compose d'une étude comparée des techniques PM et des méthodes d'identification reposant sur l'approximation PAST (cf. §4.3) ;
- la deuxième phase de l'étude a pour objet l'analyse de l'influence des paramètres de construction f , p , n_x et λ sur la fiabilité des estimées des algorithmes PM ;
- un examen des capacités de poursuite des techniques PM est finalement proposé en les appliquant à trois versions lentement variantes du procédé (6.1).

6.1.1 Étude d'un système invariant. Comparaison des algorithmes

Afin de valider les capacités des algorithmes PM, il est nécessaire de les confronter à d'autres techniques récursives des sous-espaces. Les études proposées dans la littérature montrent que les résultats les plus intéressants ont été obtenus, jusqu'à aujourd'hui, à l'aide des méthodes utilisant l'approche PAST [Lov98, Gus99, LPV01, ONVV01, OK02]. Il semble donc essentiel de comparer les estimées produites par ces deux classes d'algorithmes d'identification récursive.

¹Ce système a été initialement proposé par H. Oku [OK99]. Il est complété par un bruit de mesure en entrée \mathbf{v} .

6.1.1.1 Données d'entrée-sortie et paramètres

Considérons donc le système (6.1) tel que :

- l'entrée non bruitée \mathbf{u} , le bruit d'état w et le bruit de mesure en sortie v soient des bruits blancs gaussiens de moyenne nulle et de variance unitaire ;
- le bruit de mesure en entrée \mathbf{v} soit un bruit blanc gaussien de moyenne nulle et de variance² égale à 0.1.

Ces perturbations sont relativement faibles puisqu'elles correspondent à un rapport signal sur bruit de 46dB. Les signaux d'entrée choisis permettent quant à eux de s'assurer que la condition d'excitation persistante [Lju99] ainsi que les contraintes décrites dans le paragraphe 5.4.3.3 pour la construction de la variable instrumentale soient vérifiées. Supposons de plus que nous ayons accès *a priori* à l'ordre du système³. Trois paramètres doivent alors être choisis :

- l'indice p ;
- l'indice f ;
- le facteur d'oubli λ .

Les indices p et f doivent vérifier la condition suivante :

$$p > n_x \text{ et } f > n_x \quad (6.2)$$

pour assurer, par exemple, que $\text{rang}(\mathbf{\Gamma}_f) \geq n_x$. Ces indices ne doivent cependant pas être trop élevés afin de ne pas excessivement augmenter le temps de calcul. Nous poserons donc :

$$p = f = 8, \quad (6.3)$$

les conditions précédentes étant ainsi vérifiées. Le facteur d'oubli est quant à lui choisi pour oublier la phase initiale d'estimation⁴ puis conserver au maximum les informations acquises sur le procédé. Rappelons en effet que le système étudié est stationnaire. Un facteur d'oubli tout à fait adapté à ce type de problématique est (cf. §2.3.1.1) :

$$\lambda(t) = \min \{ \lambda_0 \lambda(t-1) + 1 - \lambda_0, \lambda_{final} \}. \quad (6.4)$$

Nous posons alors :

$$\lambda_0 = 0.99 \quad (6.5)$$

$$\lambda(0) = 0.99 \quad (6.6)$$

$$\lambda_{final} = 0.9995. \quad (6.7)$$

La valeur finale est alors atteinte après 300 itérations. Une fois ces paramètres fixés, il est nécessaire de choisir une variable instrumentale. Comme nous l'avons vu dans le paragraphe 5.4.3.3, sélectionner comme instruments des entrées bruitées passées :

$$\boldsymbol{\xi} = \begin{bmatrix} \mathbf{u}(t-p) \\ \vdots \\ \mathbf{u}(t-1) \end{bmatrix} \quad (6.8)$$

se révèle être un choix tout à fait approprié à notre problème d'estimation.

²Remarquez que les bruits d'état et de sortie sont filtrés par des matrices dont les composantes sont 10 fois plus faibles que celles des autres matrices. Par souci d'homogénéité, puisque le bruit de mesure en entrée est appliqué directement, il est nécessaire de fixer la variance de \mathbf{v} à 0.1.

³Une analyse expérimentale de l'influence d'une surestimation et d'une sous-estimation de ce paramètre sera exposée dans la suite de ce paragraphe.

⁴Les matrices d'état initiales seront choisies totalement au hasard.

6.1.1.2 Analyse comparative

Les méthodes d'identification récursive des sous-espaces sont composées de deux grandes étapes (cf. §5.2) :

- l'estimation récursive du vecteur d'observation ;
- la mise à jour de la matrice d'observabilité.

Bien que ces deux phases soient intimement liées, il est intéressant d'étudier l'influence de chacune d'entre-elles. Une analyse séparée de ces deux étapes d'estimation est donc proposée au sein des deux paragraphes suivants.

Algorithmes de mise à jour du vecteur d'observation Nous avons vu, tout au long de ce manuscrit, qu'il existait un grand nombre de méthodes récursives des sous-espaces et que les combinaisons des étapes « mise à jour du vecteur d'observation » / « estimation de la matrice d'observabilité » étaient très nombreuses. La première phase de cette étude consiste donc à comparer la fiabilité des estimées du vecteur d'observation obtenues par mise à jour des décompositions QR des schémas d'identification Ordinary MOESP et PI/PO MOESP (cf. §5.3.1.2 et 5.3.2.2) à celles produites en utilisant le lemme d'inversion matricielle (cf. §5.3.1.1 et 5.3.2.1). Puisqu'il est difficile de comparer directement les vecteurs d'observation estimés, l'orientation choisie consiste à utiliser, dans chaque cas, un algorithme d'estimation de la matrice d'observabilité commun à ROM_{qr} et ROM_{im} (resp. RPIPO_{qr} et RPIPO_{im}) puis de comparer le résultat final, i.e. le modèle de procédé estimé. Considérons donc successivement les couples EIVPM+ROM_{qr}, EIVPM+ROM_{im}, RPM2+RPIPO_{qr} et RPM2+RPIPO_{im} et comparons les deux à deux (cf. fig. 6.2 et 6.3). Par souci d'homogénéité, les matrices initiales, dont les composantes

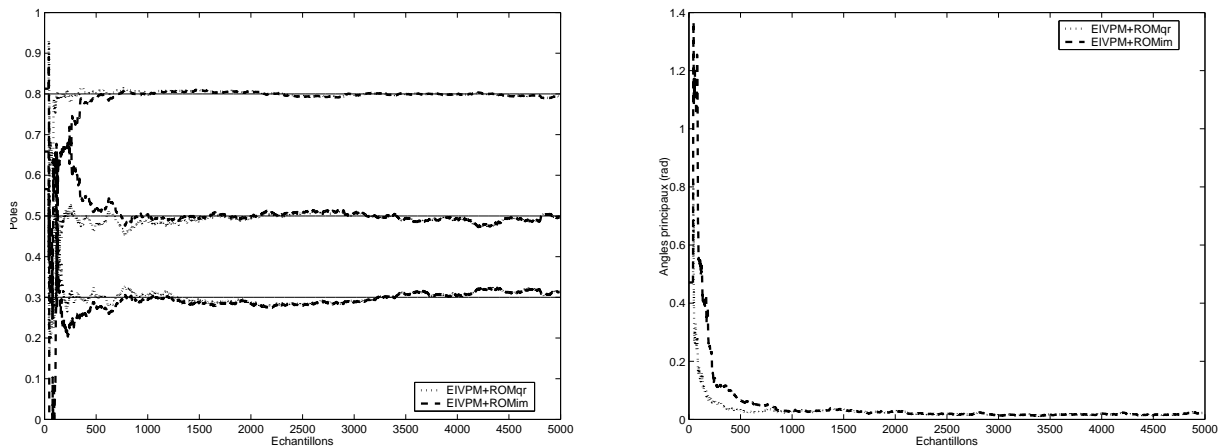


FIG. 6.2: Comparaison des algorithmes ROM_{im} et ROM_{qr}.

sont calculées au hasard à l'aide la fonction `randn` de Matlab sous la contrainte d'avoir un modèle d'état initial stable, sont communes aux paires EIVPM+ROM_{qr} et EIVPM+ROM_{im} (resp. RPM2+RPIPO_{qr} et RPM2+RPIPO_{im}). Les résultats sont illustrés dans chaque cas à l'aide des tracés suivants :

- les pôles estimés à partir de la matrice d'état du modèle ;
- l'angle entre la matrice d'observabilité théorique et la matrice d'observabilité construite à partir des matrices **A** et **C** estimées.

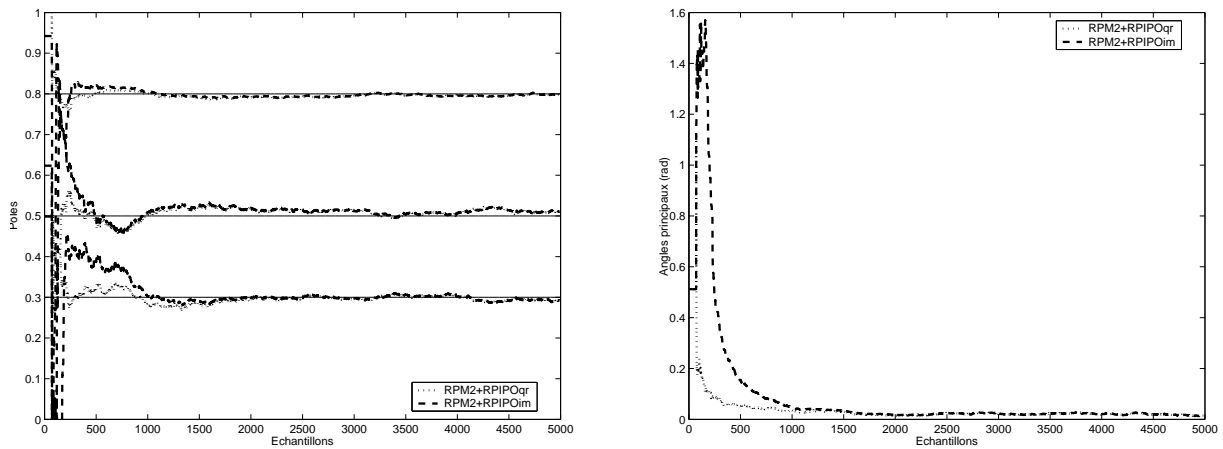


FIG. 6.3: Comparaison des algorithmes $RPIPO_{im}$ et $RPIPO_{qr}$.

Le premier enseignement que nous pouvons tirer de ces figures est que les quatre techniques d'estimation analysées produisent des estimées fiables du vecteur d'observation puisque les pôles calculés par les quatre couples d'algorithmes convergent vers les valeurs théoriques. Il est cependant intéressant de constater que les estimées obtenues en appliquant les algorithmes fondés sur la mise à jour des décompositions QR des schémas d'identification Ordinary et PI/PO MOESP sont légèrement plus rapides à converger que celles calculées en utilisant le lemme d'inversion. Ceci est particulièrement flagrant sur les tracés des angles entre matrices d'observabilité. Il faut en effet, en moyenne, 800 échantillons pour que les angles obtenus avec ROM_{im} et $RPIPO_{im}$ prennent de valeurs équivalentes à celles estimées à l'aide des méthodes ROM_{qr} et $RPIPO_{qr}$. Ce résultat a été confirmé à l'aide d'un tirage de Monte Carlo (100 paires de données d'entrée-sortie de 5000 échantillons chacune couplées à 100 initialisations aléatoires des matrices d'état). Sur cet exemple de simulation, les techniques utilisant les rotations de Givens conduisent donc plus rapidement à des estimées fiables du vecteur d'observation. Cette caractéristique, associée à une robustesse numérique démontrée, laissent à penser que les algorithmes ROM_{qr} et $RPIPO_{qr}$ sont à privilégier malgré un coût calculatoire supérieur (cf. §5.3.3). Dans la suite de cette étude comparative, ces méthodes seront donc exclusivement appliquées pour estimer le vecteur d'observation. Il est en effet nécessaire d'étudier la vitesse de convergence des techniques d'estimation récursive de la matrice d'observabilité. Or, pour éviter de cumuler les temps de convergence des deux étapes d'identification, il est essentiel d'utiliser les algorithmes de mise à jour du vecteur d'observation les plus rapides.

Algorithmes d'estimation récursive de la matrice d'observabilité Cette seconde phase d'étude concerne l'étape d'estimation de la matrice d'observabilité. L'objectif de cette analyse réside dans la comparaison des performances des algorithmes récursifs utilisant l'approximation PAST et le propagateur. Cinq cas vont être ainsi successivement considérés :

- une analyse comparée des algorithmes $EIVPAST$, $EIVPM$ et $EIV_{sqr}tPM$;
- une comparaison des méthodes $UDPAST$ et $RPM2$;
- une étude des techniques $COPAST$ et $COPM$;
- un examen des algorithmes $COIVPAST$ et $COIVPM$;
- une confrontation des méthodes $COPASTd$ et $COPM$.

Ce classement est légitimé par le fait que, dans chaque cas, l'architecture des algorithmes proposés est commune aux deux approches examinées.

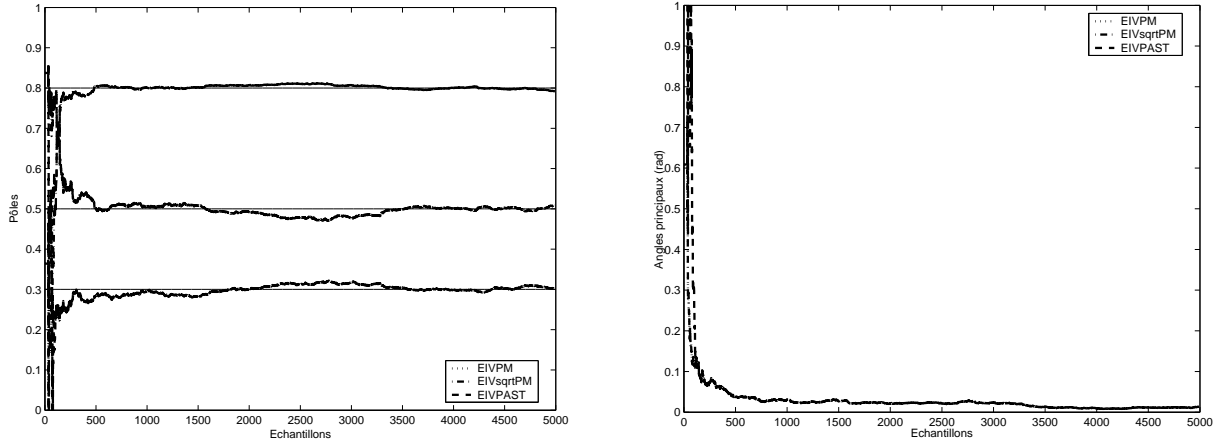


FIG. 6.4: Comparaison des algorithmes EIVPAST, EIVPM et EIV_{sqrt}PM.

Considérons donc, dans un premier temps, les techniques EIVPAST, EIVPM et EIV_{sqrt}PM et traçons respectivement les évolutions, au cours de la récursion, des pôles calculés à partir de la matrice d'état du modèle estimé et des angles entre matrice d'observabilité théorique et matrices d'observabilité estimées (cf. fig. 6.4). Le premier constat est que les trois méthodes sont capables de modéliser efficacement la dynamique du système dans un contexte bruité puisque les pôles estimés convergent vers les pôles du procédé identifié. Il est également manifeste de noter que les courbes de la figure 6.4 sont pratiquement superposées. Un zoom sur les 120 premiers échantillons des angles entre matrices d'observabilité (cf. fig. 6.5) indiquent cependant quelques différences dans la phase initiale d'estimation. Ces constatations renforcent les remarques faites dans les chapitres 4 et 5 ainsi que les résultats obtenus dans l'étude de cas du paragraphe 5.4.5.4 :

- Les premières itérations des algorithmes exploitant l'approximation PAST conduisent à un sous-espace différent de celui qui serait obtenu en utilisant la fonction coût originelle de Yang. La matrice d'observabilité estimée est par conséquent distincte de la matrice théorique. Cette différence s'estompe au cours de la récursion puisque les estimées obtenues avec EIVPAST rejoignent celles fournies par EIVPM et EIV_{sqrt}PM après 200 itérations.
- Les techniques fondées sur le critère de Yang (ou une approximation de ce dernier) utilisant une variable instrumentale et celles basées sur le propagateur associé également à une variable instrumentale possèdent des propriétés du premier et du second ordre équivalentes.

Ces résultats sont confirmés par l'analyse comparée des méthodes UDPAST et RPM2 (cf. fig. 6.6), COPAST et COPM (cf. fig. 6.7) et COIVPAST et COIVPM (cf. fig. 6.8). Pour chaque couple étudié, les courbes d'évolution des pôles et angles principaux sont en effet très rapidement superposées. Quelques différences apparaissent néanmoins dans la phase d'initialisation. En moyenne, il faut 70 échantillons pour que l'angle principal entre la matrice d'observabilité théorique et les matrices d'observabilité estimées à l'aide des techniques fondées sur l'approximation PAST rejoignent les valeurs des angles principaux calculés à partir des modèles PM.

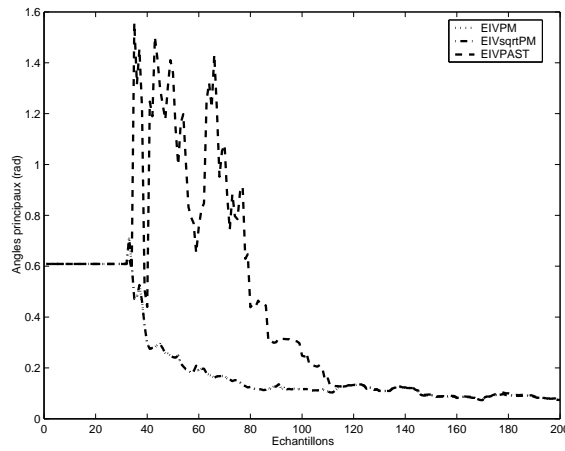


FIG. 6.5: Comparaison des algorithmes EIVPAST, EIVPM et EIV_{sqrt}PM. Les 200 premiers échantillons.

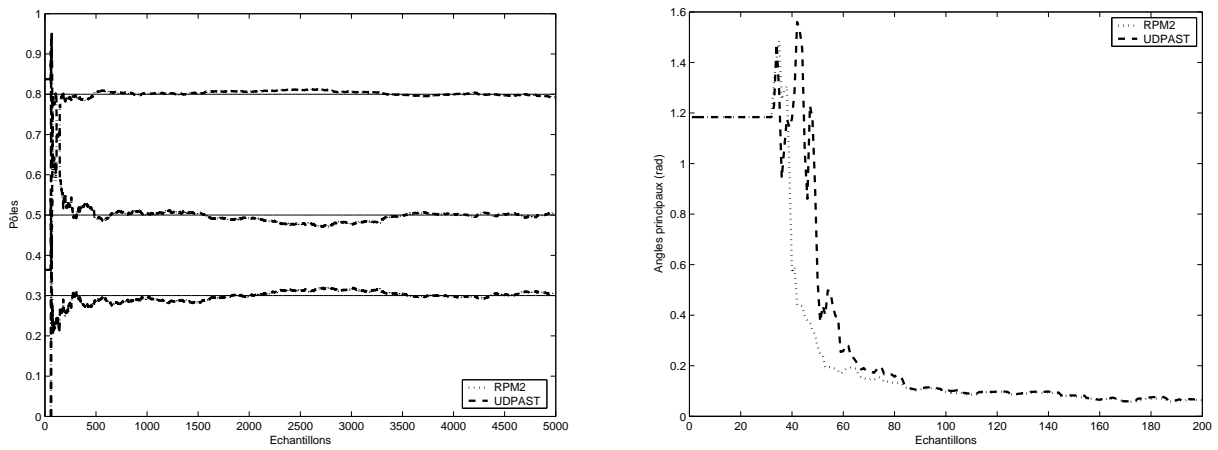


FIG. 6.6: Comparaison des algorithmes UDPAST et RPM2.

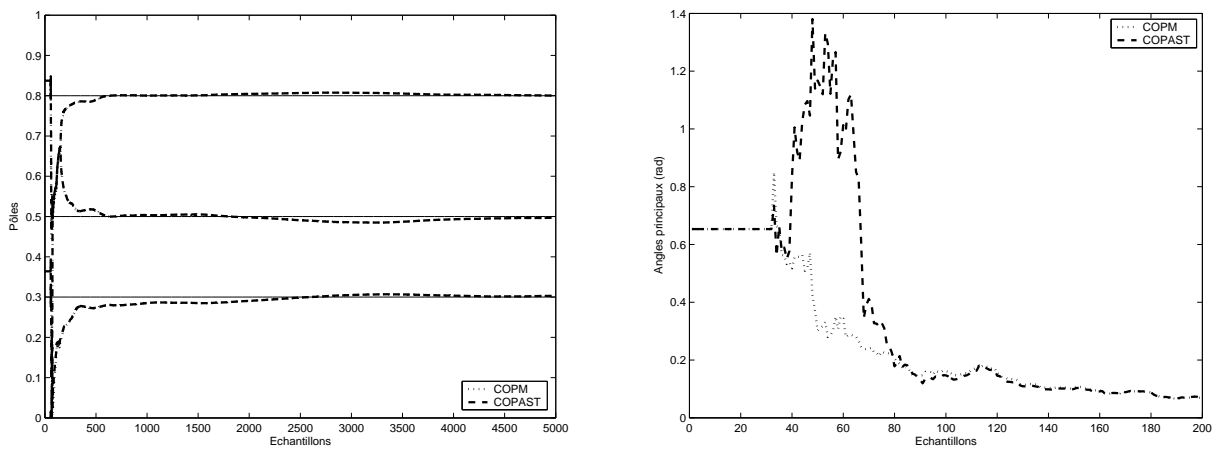


FIG. 6.7: Comparaison des algorithmes COPAST et COPM.

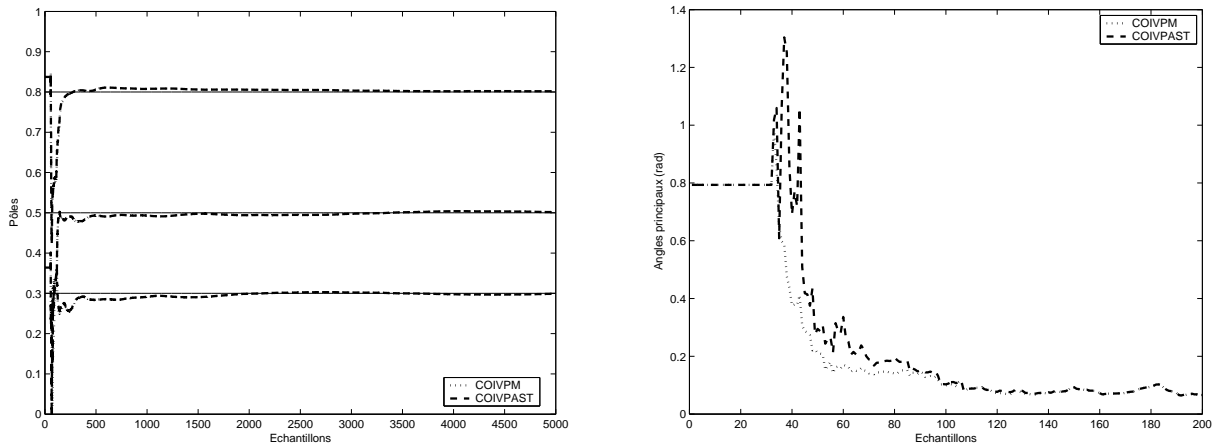


FIG. 6.8: Comparaison des algorithmes COIVPAST et COIVPM.

La dernière étape de cette étude a pour objectif d'analyser les performances des algorithmes utilisant la technique de déflation. Afin de ne pas multiplier les combinaisons de méthodes à comparer, nous ne considérons que les algorithmes COPM et COPASTd. Des résultats équivalents sont obtenus avec COIVPM et COIVPASTd. Au vu des courbes de la figure 6.9, il est flagrant que les algorithmes utilisant la technique de déflation sont bien plus lents à converger que leurs homologues n'employant pas cet artifice mathématique. Cette caractéristique les rend ainsi plus difficilement utilisables sur des procédés dont la dynamique peut évoluer relativement rapidement.

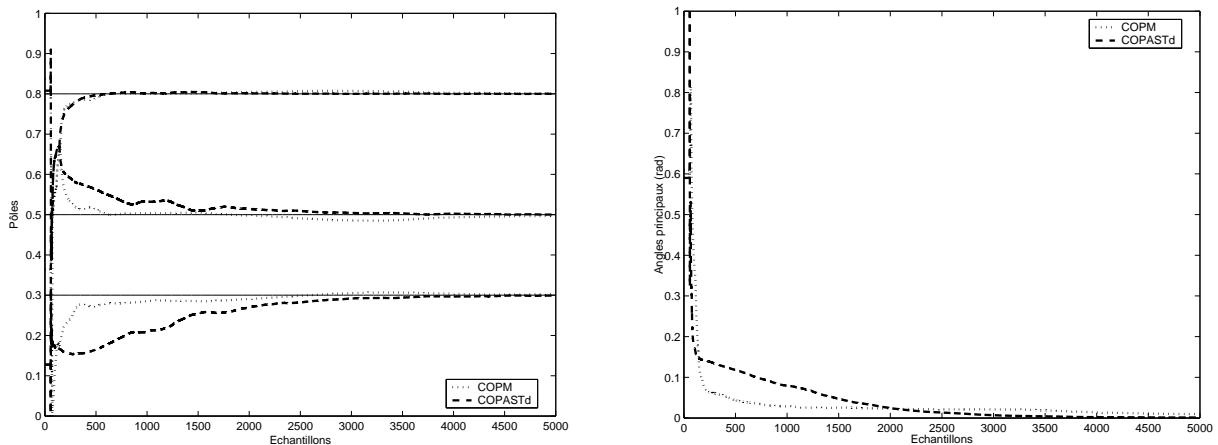


FIG. 6.9: Comparaison des algorithmes COPM et COPASTd.

Conclusion

Cette étude comparative a permis de mettre en évidence que :

- les algorithmes récursifs d'estimation du vecteur d'observation fondés sur la mise à jour des décompositions QR des schémas Ordinary et PI/PO MOESP produisent plus rapidement des estimées fiables de ce vecteur que ceux employant le lemme d'inversion matricielle ;
- les méthodes utilisant la technique de déflation possèdent une dynamique de convergence lente qui les rend difficilement applicables sur des systèmes variants.
- tous les algorithmes PM conduisent plus rapidement aux estimées des pôles théoriques du système que leurs homologues fondés sur l'approximation PAST ;

Cette dernière constatation est d'autant plus intéressante que les méthodes PM possèdent un coût calculatoire légèrement inférieur aux techniques PAST (cf. §5.5). Or, cette complexité numérique est fortement liée aux paramètres de construction f , p et n_x . Il est donc important d'analyser l'influence de ces derniers sur la fiabilité des estimées afin de s'assurer, par exemple, qu'il n'est pas nécessaire de les surévaluer pour obtenir un modèle de procédé consistant. Cette étude est réalisée dans le paragraphe suivant.

6.1.1.3 Influence des paramètres de construction

Les méthodes d'identification récursive développés dans le chapitre 5 dépendent d'un certain nombre de paramètres :

- l'ordre du système ;
- les indices f et p ;
- le facteur d'oubli λ .

Bien que quelques algorithmes et certains principes aient été proposés pour estimer l'ordre du système et choisir les indices f et p en identification hors ligne des sous-espaces (cf. §3.3.5), aucune règle et aucune technique d'estimation n'ont été développées dans le cas récursif. Seules quelques contraintes sont mises en avant :

$$f > n_x \text{ et } p > n_x. \quad (6.9)$$

Il semble donc important d'éprouver la fiabilité des algorithmes d'estimation PM aux choix des paramètres de construction n_x , f , p et λ . Cette analyse est plus précisément réalisée sur un exemple de simulation.

Influence de l'ordre Le paramètre le plus important à déterminer pour la construction des techniques PM est l'ordre du système. Puisque ces méthodes n'estiment nullement les éléments propres du procédé, il est impossible de calculer l'ordre du système au cours de la récursion à l'aide d'algorithmes classiques d'estimation (NIC, SVC...). Il est donc primordial d'examiner le comportement de ces techniques d'identification récursive lors d'une faible sous-estimation ou surestimation de l'ordre du système. Considérons donc de nouveau l'exemple de simulation (6.1) et testons les algorithmes EIVPM, RPM2, COIVPM et COPM sur ce dernier en posant successivement $n_x = 2$ puis $n_x = 4$. Afin d'examiner la fiabilité des modèles identifiés, traçons, dans chaque cas, l'évolution des pôles estimés ainsi que l'angle entre la matrice d'observabilité théorique d'ordre 3 et les matrices d'observabilité d'ordre 2 et d'ordre 4 estimées.

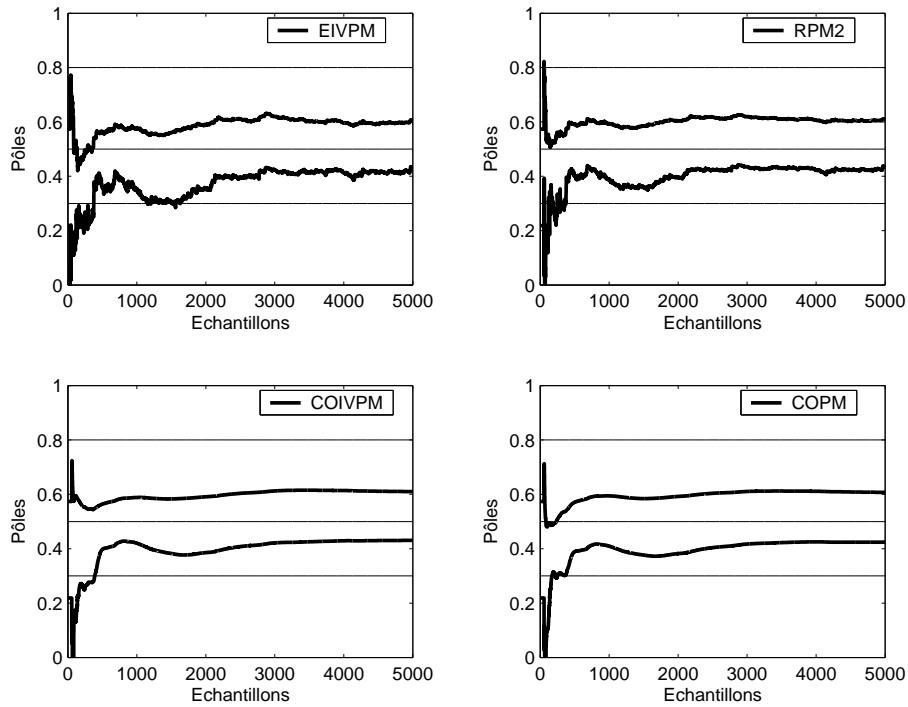


FIG. 6.10: Pôles estimés à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM.
 $n_x = 2$.

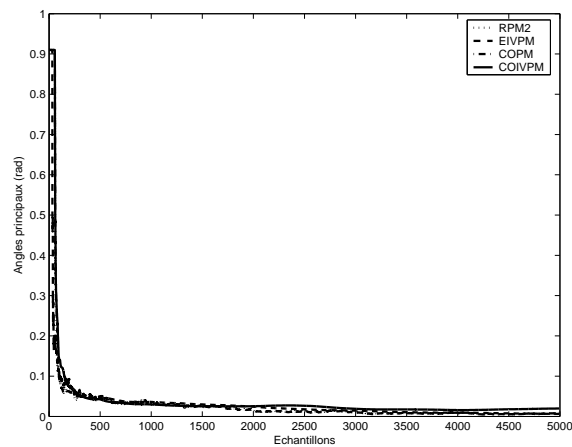


FIG. 6.11: Angles entre la matrice d'observabilité théorique d'ordre 3 et les matrices d'observabilité estimées à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM.
 $n_x = 2$.

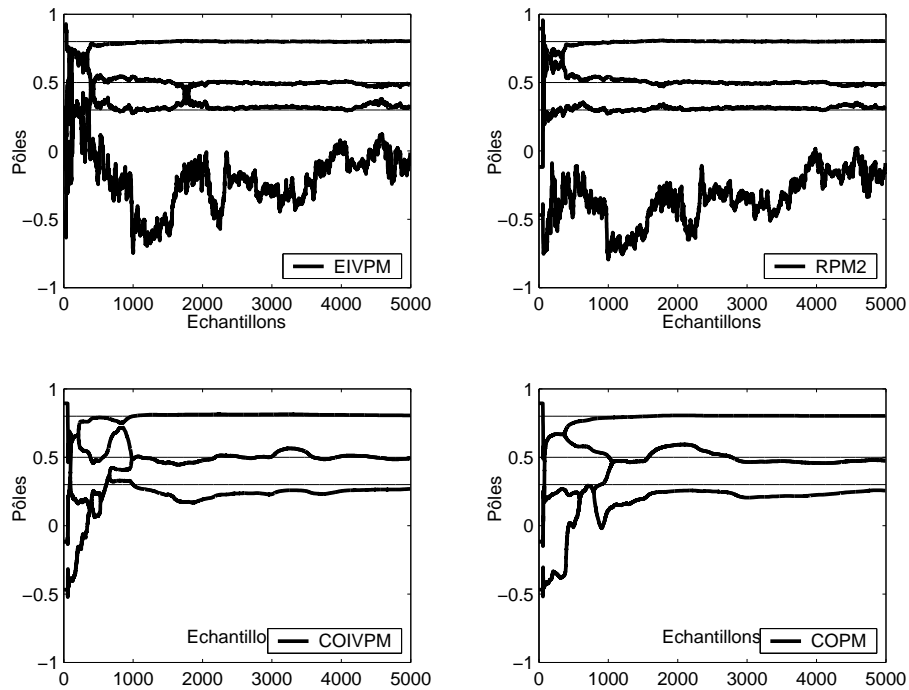


FIG. 6.12: Pôles estimés à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM.
 $n_x = 4$.

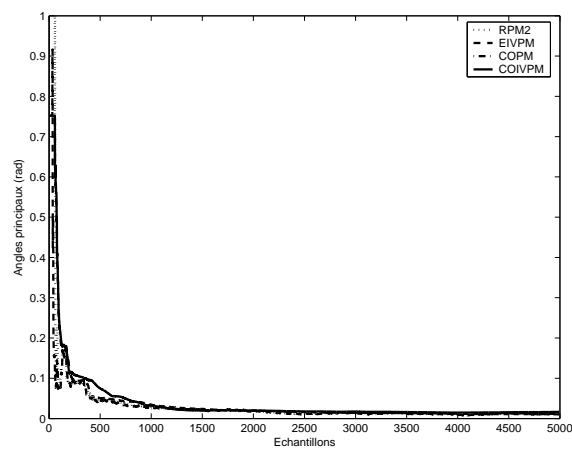


FIG. 6.13: Angles entre la matrice d'observabilité théorique d'ordre 3 et les matrices d'observabilité estimées à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM.
 $n_x = 4$.

Dans le cas d'une faible sous-estimation de l'ordre, les courbes de la figure 6.10 indiquent que les quatre algorithmes étudiés estiment des modèles moyens d'ordre 2 relativement stables. Leur comportement est proche de celui du système recherché puisque les angles entre la matrice d'observabilité théorique d'ordre 3 et les matrices d'observabilité estimées d'ordre 2 représentés figure 6.11 sont rapidement très proches de zéro.

Dans le cas d'une faible surestimation ($n_x = 4$), la figure 6.13 révèle que les méthodes récursives analysées produisent des modèles d'ordre 4 ayant également un comportement proche de celui du système d'ordre 3. Il est important de noter que, pour chaque modèle, trois des pôles estimés sont voisins des pôles théoriques. Les méthodes considérées se différencient cependant par le fait que :

- les algorithmes EIVPM et RPM2 estiment un pôle supplémentaire possédant une variance importante ;
- les algorithmes COIVPM et COPM produisent un modèle ayant un pôle multiple.

Des constatations analogues ont été réalisées lors d'une surestimation plus conséquente (ordre 6). Les 4^e, 5^e et 6^e pôles des modèles d'ordre 6 sont soit de variance importante (EIVPM et RPM2), soit confondus avec les trois autres pôles calculés (COIVPM et COPM). Ces constats conduisent à penser qu'il serait possible de retrouver les pôles théoriques du système à partir de modèles surdimensionnés estimés par les méthodes PM.

Influence du facteur d'oubli Le facteur d'oubli est un paramètre qui permet de pondérer les dernières informations acquises relativement aux connaissances cumulées depuis le début de la procédure d'identification. Ce dernier est généralement utilisé pour tenir compte des probables évolutions du processus. Avant de caractériser son rôle lorsqu'un système variant est considéré, il est intéressant d'analyser son influence sur l'estimation du modèle de procédé lorsque le processus à identifier est à dynamique invariante. Considérons donc de nouveau le système stationnaire (6.1) et estimons le modèle de processus à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM en fixant successivement, toute chose égale par ailleurs, $\lambda_{final} = 0.9995$ et $\lambda_{final} = 0.995$.

Il est aisé de constater (cf. fig. 6.14 et 6.15) que les meilleurs résultats sont obtenus pour $\lambda_{final} = 0.9995$. La variance des pôles estimés est en effet bien plus faible que lorsque $\lambda_{final} = 0.995$. Ce constat peut se justifier par le fait que, plus le facteur d'oubli est faible, plus l'algorithme considéré devient sensible aux variations du bruit agissant sur le procédé. Cette propriété a été démontrée dans le paragraphe 2.3.1.1 pour les méthodes des moindres carrés récursifs. Il semble en être de même pour les techniques EIVPM, RPM2, COIVPM et COPM. L'introduction d'une variable instrumentale ne permet donc pas d'annihiler totalement l'effet des perturbations. Elle permet cependant de fournir des estimées asymptotiquement non biaisées. Ce n'aurait bien évidemment pas été le cas si nous avions employé, dans un contexte aussi bruité, les moindres carrés récursifs. Cette analyse indique de plus que les estimées calculées par les algorithmes fondés sur une approche séquentielle (COIVPM et COPM) possèdent une dynamique de variation différente de celles obtenues à l'aide de leurs homologues utilisant directement les matrices de covariance \mathbf{R}_z et $\mathbf{R}_{z\xi}$. La variance des estimées produites par COIVPM et COPM est en effet plus lente que celle des techniques EIVPM et RPM2. Ceci peut s'expliquer par la nature séquentielle des algorithmes COIVPM et COPM. Cette dernière conduit à un effet de moyennage temporel des estimations se traduisant par une variance moins rapide des estimées. Il sera intéressant d'étudier l'influence de ce moyennage temporel sur un système variant.

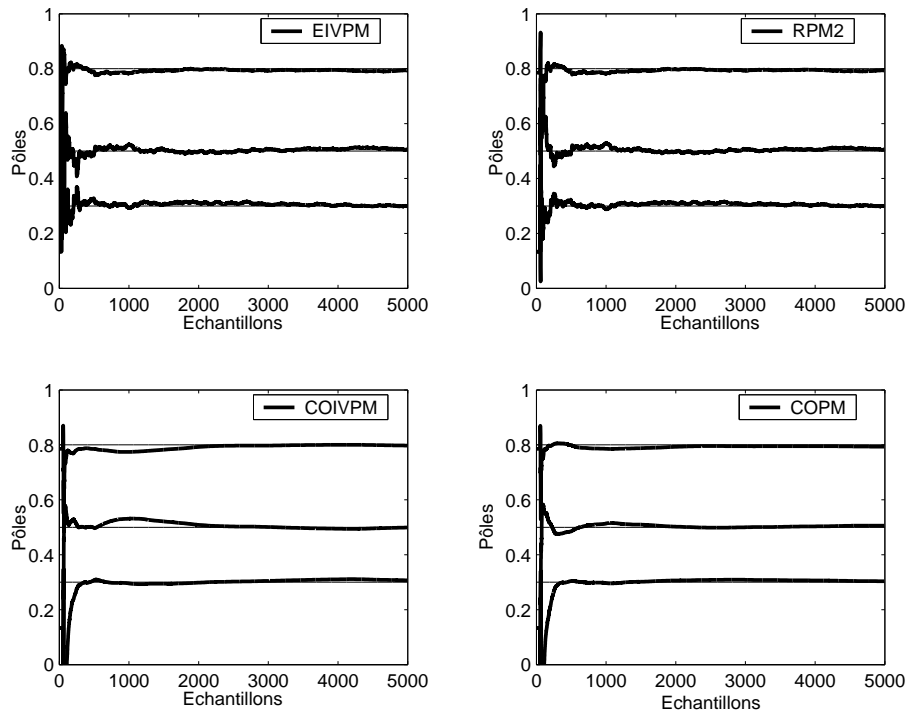


FIG. 6.14: Pôles estimés à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM.

$$\lambda_{final} = 0.9995.$$

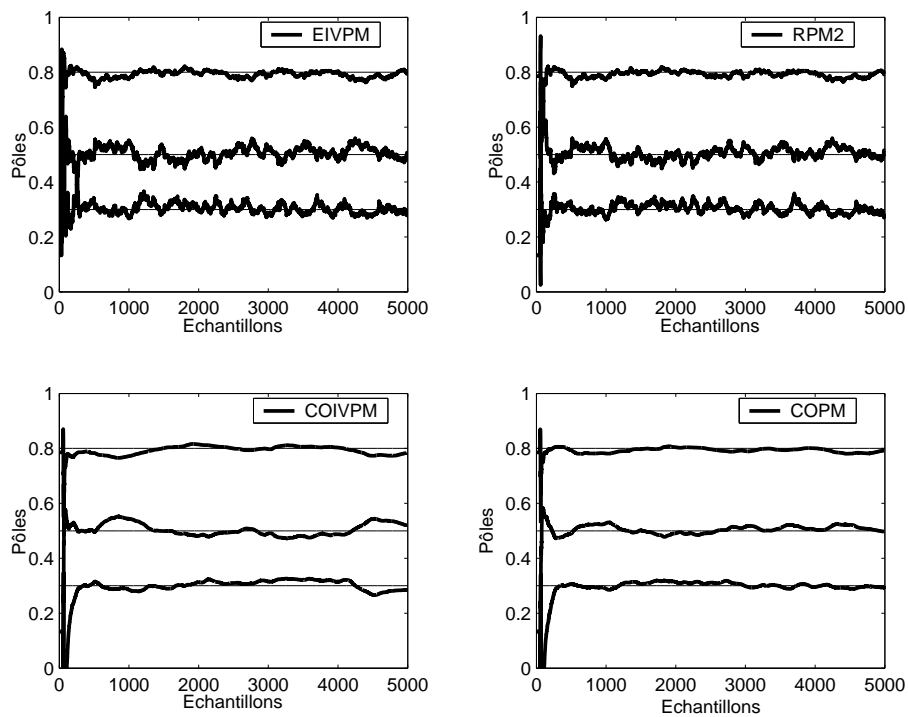


FIG. 6.15: Pôles estimés à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM.

$$\lambda_{final} = 0.995.$$

Influence des paramètres f et p Le choix des paramètres f et p est un problème récurrent en identification des sous-espaces. Des études dans le cas non récursif [Chi97, Bau03a] (cf. §3.3.5) ont permis de fixer quelques contraintes sans pour autant donner de solutions universelles. Or, en identification récursive des sous-espaces, ces indices interviennent invariablement au niveau de la complexité numérique des algorithmes d'estimation (cf. §5.3.3 et 5.4.6). Il est donc important de vérifier si l'utilisation de paramètres f et p élevés améliorent la consistance des estimées. Deux cas sont ainsi considérés : $f = p = 8$ et $f = p = 20$. Les pôles des modèles calculés par EIVPM, RPM2, COIVPM et COPM dans ces deux situations sont alors tracés (cf. fig. 6.16 et 6.17). Ces courbes indiquent que la surévaluation des paramètres f et p n'apportent aucune amélioration de la fiabilité des estimées. Un constat équivalent est réalisé lorsque la valeur finale du facteur d'oubli est choisie égale à 0.995. Il ne semble donc pas intéressant, pour cet exemple de simulation, de fixer des indices f et p trop élevés.

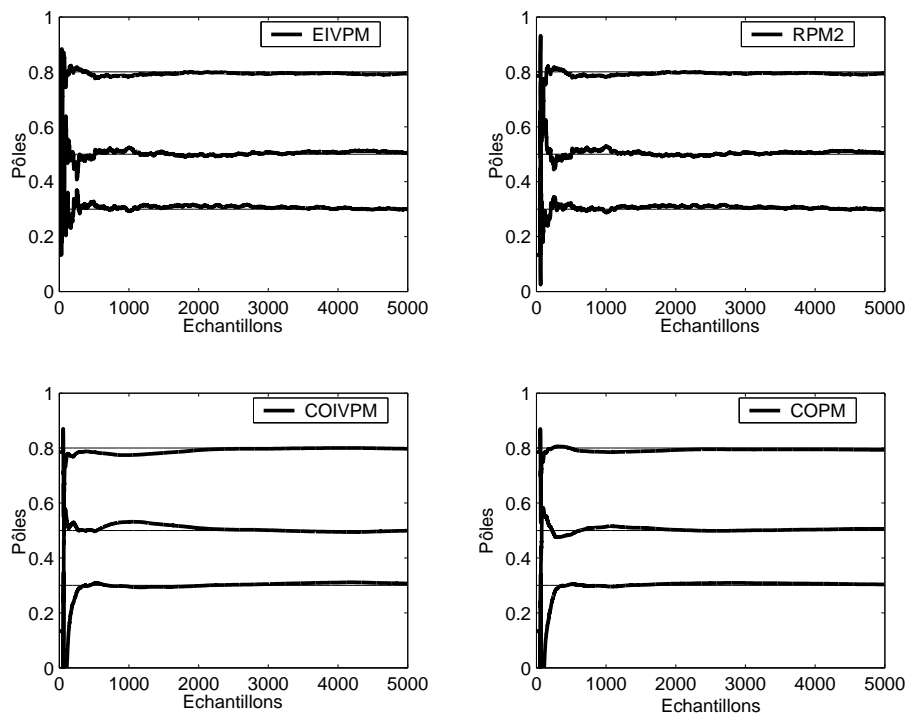


FIG. 6.16: Pôles estimés à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM.

$$f = p = 8.$$

Conclusion

Cette analyse de l'influence des paramètres de construction réalisée à l'aide d'un système simulé indique que :

- les algorithmes PM sont capables de produire, lors d'une faible sous-estimation ou surestimation de l'ordre, des modèles stables ayant un comportement proche de celui du système à identifier ;
- dans le cas d'une surestimation de l'ordre, parmi l'ensemble des pôles calculés, les pôles

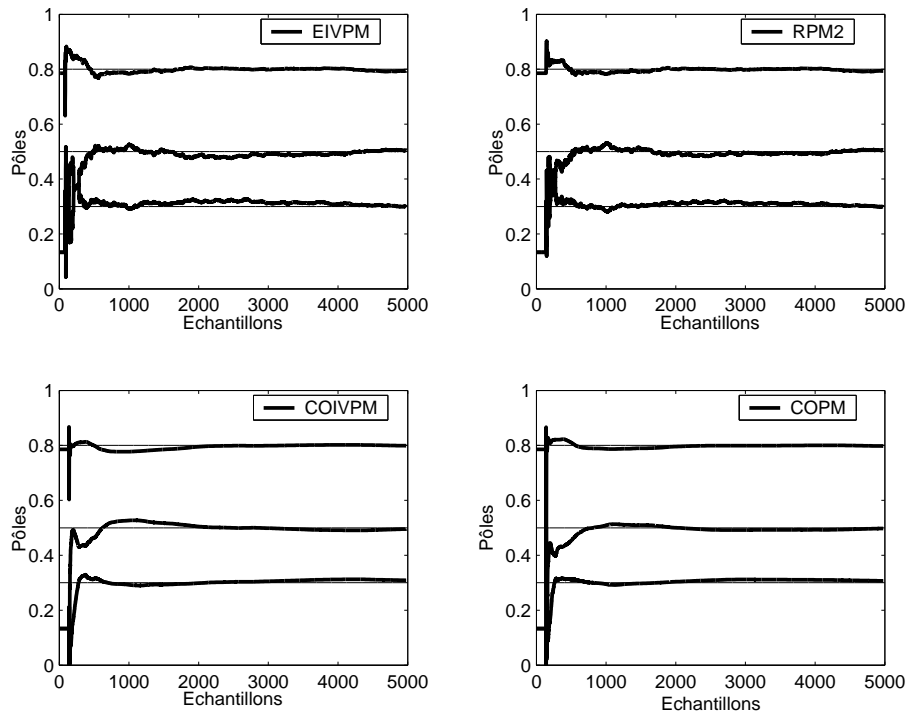


FIG. 6.17: Pôles estimés à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM.
 $f = p = 20$.

réels du système sont relativement bien estimés⁵ ;

- un facteur d'oubli faible augmente grandement la sensibilité des méthodes récursives vis-à-vis des perturbations ;
- la surévaluation des paramètres de construction f et p ne conduit en aucun cas à un modèle plus fiable.

Bien qu'encourageantes, ces conclusions doivent également être validées sur des données non stationnaires. L'étude de systèmes de simulation variants est proposée dans le paragraphe suivant.

6.1.2 Étude d'un système variant

Les algorithmes d'identification récursive sont souvent employés pour fournir des modèles capables de suivre les évolutions de la dynamique du procédé à identifier. La dernière étape de cette étude sur données simulées consiste donc à comparer les méthodes EIVPM, RPM2, COPM et COIVPM sur des processus variants. Deux types d'évolution vont être successivement considérés :

- une variation lente des pôles du système ;
- un saut de la valeur de l'un des pôles du procédé.

⁵Une réflexion sur la détermination de critères de sélection de ces pôles pourrait être engagée pour tenter d'estimer l'ordre du système à partir des modèles surdimensionnés produits par les algorithmes PM.

Trois matrices d'état vont ainsi être envisagées :

$$\mathbf{A}_{var_1} = \begin{bmatrix} 0.8 + 0.15 \frac{\exp\left(-\frac{t-3333}{6667}\right) - 1}{\exp(-1) - 1} & -0.4 & 0.2 \\ 0 & 0.3 - 0.5 \frac{\exp\left(-\frac{t-3333}{6667}\right) - 1}{\exp(-1) - 1} & -0.5 \\ 0 & 0 & 0.5 + 0.2 \frac{\exp\left(-\frac{t-3333}{6667}\right) - 1}{\exp(-1) - 1} \end{bmatrix} \quad (6.10)$$

et

$$\mathbf{A}_{var_2} = \begin{bmatrix} 0.8 - 0.3 \frac{\exp\left(-\frac{t-3333}{6667}\right) - 1}{\exp(-1) - 1} & -0.4 & 0.2 \\ 0 & 0.3 - 0.5 \frac{\exp\left(-\frac{t-3333}{6667}\right) - 1}{\exp(-1) - 1} & -0.5 \\ 0 & 0 & 0.5 + 0.2 \frac{\exp\left(-\frac{t-3333}{6667}\right) - 1}{\exp(-1) - 1} \end{bmatrix} \quad (6.11)$$

dans le cas d'une variation lente ;

$$\mathbf{A}_{var_3} = \begin{cases} \begin{bmatrix} 0.8 & -0.4 & 0.2 \\ 0 & 0.3 & -0.5 \\ 0 & 0 & 0.5 \end{bmatrix} & \text{pour } t < t_{change} \\ \begin{bmatrix} 0.8 & -0.4 & 0.2 \\ 0 & 0.45 & -0.5 \\ 0 & 0 & 0.5 \end{bmatrix} & \text{pour } t \geq t_{change} \end{cases} \quad (6.12)$$

dans le cas d'une variation brutale. La particularité de la matrice \mathbf{A}_{var_2} réside dans le fait que les valeurs des deux plus grands pôles se croisent durant l'expérience. Une analyse des capacités de poursuite des algorithmes PM va ainsi pouvoir être menée. Une étude de l'influence des paramètres λ , f et p dans le cas de l'identification de systèmes variants sera également proposée.

6.1.2.1 Cas 1 : $\mathbf{A} = \mathbf{A}_{var_1}$

La première phase de cette étude consiste à étudier le comportement des algorithmes EIVPM, RPM2, COPM et COIVPM lorsque le système évolue lentement. Trois paramètres peuvent *a priori* agir sur les capacités de poursuite des techniques d'identification récursive considérées : le facteur d'oubli λ et les indices f et p . Ces derniers fixent en effet la fenêtre temporelle utilisée pour estimer, à chaque itération, les matrices d'état du système.

Influence du facteur d'oubli Considérons donc le système de simulation (6.1) pour lequel $\mathbf{A} = \mathbf{A}_{var_1}$. Le nombre d'échantillons utilisé pour l'expérience vaut 10000. Les variations des paramètres du système apparaissent après 3333 échantillons. Cette convention permet d'avoir un modèle de procédé stable avant que ce dernier évolue. Trois facteurs d'oubli sont alors envisagés : $\lambda_{final} = 0.9995$, $\lambda_{final} = 0.995$ et $\lambda_{final} = 0.99$. Ces choix conduisent respectivement aux estimées des figures 6.18, 6.19 et 6.20.

- D'un point de vue général, ces courbes mettent en évidence le compromis « capacité de poursuite / sensibilité aux perturbations » relatif au choix du facteur d'oubli. L'analyse combinée des figures 6.18 et 6.19 indiquent en effet qu'il est nécessaire de diminuer le facteur d'oubli pour être capable de suivre les évolutions des pôles du procédé. La figure

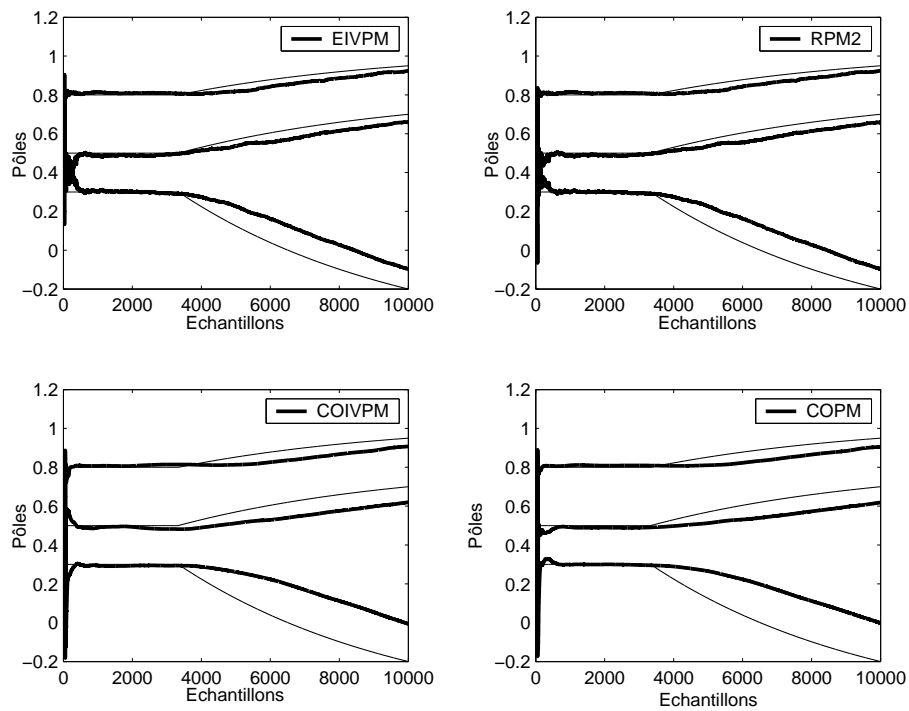


FIG. 6.18: Pôles estimés à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM.
 $\lambda_{final} = 0.9995$.

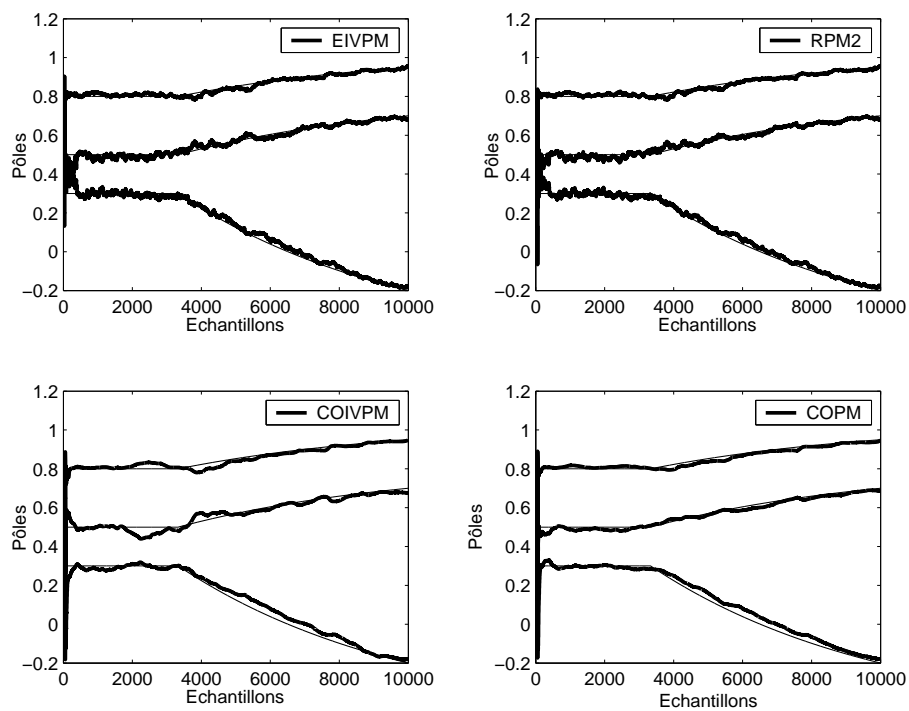


FIG. 6.19: Pôles estimés à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM.
 $\lambda_{final} = 0.995$.

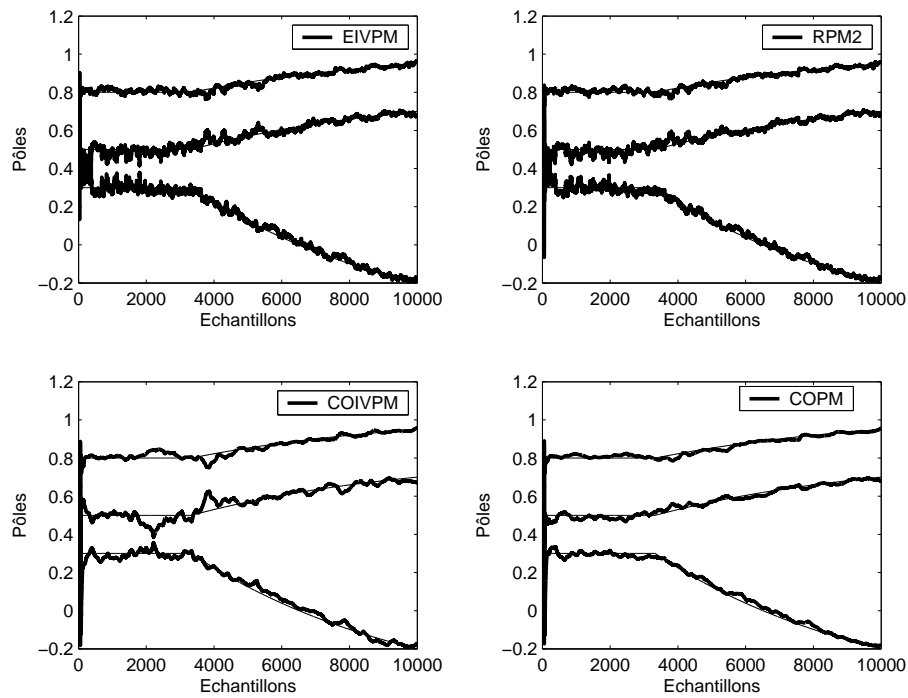


FIG. 6.20: Pôles estimés à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM.
 $\lambda_{final} = 0.99$.

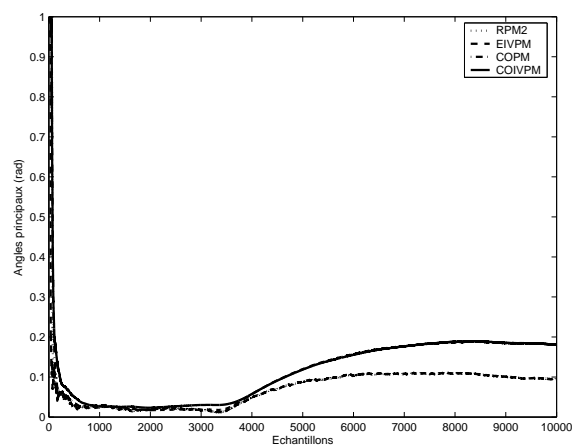


FIG. 6.21: Angles entre la matrice d'observabilité théorique et les matrices d'observabilité estimées à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM.
 $\lambda_{final} = 0.9995$.

6.20 pondèrent cette constatation en montrant qu'une baisse trop importante de la valeur de ce facteur conduit les algorithmes étudiés à devenir relativement sensibles aux bruits agissant sur le procédé. Il est donc nécessaire de choisir un facteur d'oubli plus faible que dans le cas invariant pour que le modèle soit susceptible de reproduire au mieux la dynamique du système évolutif sans pour autant le diminuer trop fortement afin qu'il ne soit pas trop affecté par les perturbations. Un tel ajustement doit être réalisé *a priori*, en fonction des connaissances acquises sur le procédé.

- D'un point de vue plus spécifique, il est à noter que le comportement des algorithmes COIVPM et COPM diffère légèrement de celui des techniques EIVPM et RPM2. La dynamique des estimées calculées à partir de ces dernières est en effet légèrement plus chaotique que celle des pôles déterminés à l'aide de COIVPM et COPM. Elle est également plus rapide puisque, pour un même facteur d'oubli (ici 0.9995), les angles principaux obtenus avec EIVPM et RPM2 tendent plus rapidement vers zéro qu'avec COIVPM et COPM (cf. fig. 6.21). Ces propriétés sont probablement dues, comme dans le cas invariant, au caractère séquentiel des algorithmes COIVPM et COPM. Cette spécificité les rend moins sensibles aux effets des perturbations. La méthode COPM semble plus particulièrement produire les meilleures estimées des pôles du système. Cette technique serait donc à privilégier lorsque les variations du procédé sont lentes.

Influence des paramètres f et p Il a été constaté au sein du paragraphe 6.1.1.3 qu'une surévaluation des indices f et p n'améliorait aucunement l'estimation du modèle de procédé. Il est important de vérifier s'il en est de même pour les systèmes variants. Pour cela, comparons, pour $\lambda_{final} = 0.995$, les pôles calculés par les algorithmes EIVPM, RPM2, COIVPM et COPM lorsque $f = p = 8$ puis lorsque $f = p = 20$.

Les courbes des figures 6.22 et 6.23 montrent qu'une survalorisation des paramètres f et p ne modifie nullement la fiabilité des estimées. L'introduction d'indices de construction élevés semble donc n'apporter aucun avantage. Au contraire, ces derniers conduisent à un accroissement du coût calculatoire des algorithmes d'identification. Choisir des indices relativement faibles vérifiant les contraintes $f > n_x$ et $p > n_x$ s'avère donc être une solution raisonnable.

Remarque 6.1. *Les algorithmes d'identification récursive sont de plus en plus utilisés pour la surveillance et le diagnostic [Ger98, DBL⁺01]. Une étude récente en identification hors ligne des sous-espaces a montré que l'utilisation des angles principaux entre matrices d'observabilité pouvait conduire à la détection de certaines défaillances [DD02, De 02]. Il est alors intéressant d'adapter ces résultats en identification récursive des sous-espaces. L'angle principal entre la matrice d'observabilité du modèle identifié stable et celle du modèle évolutif permet en effet, comme l'indique la figure 6.24, de fournir un bon indicateur d'évolution et, par extrapolation, de défaillance.*

6.1.2.2 Cas 2 : $\mathbf{A} = \mathbf{A}_{var_2}$

Les algorithmes d'identification récursive analysés au sein de ce manuscrit s'inspirent de techniques développées originellement en traitement d'antennes. Lorsque ces dernières sont utilisées pour détecter les directions d'arrivée, il est important qu'elles soient capables de séparer des sources se croisant au cours de l'expérience. En identification, un problème similaire peut

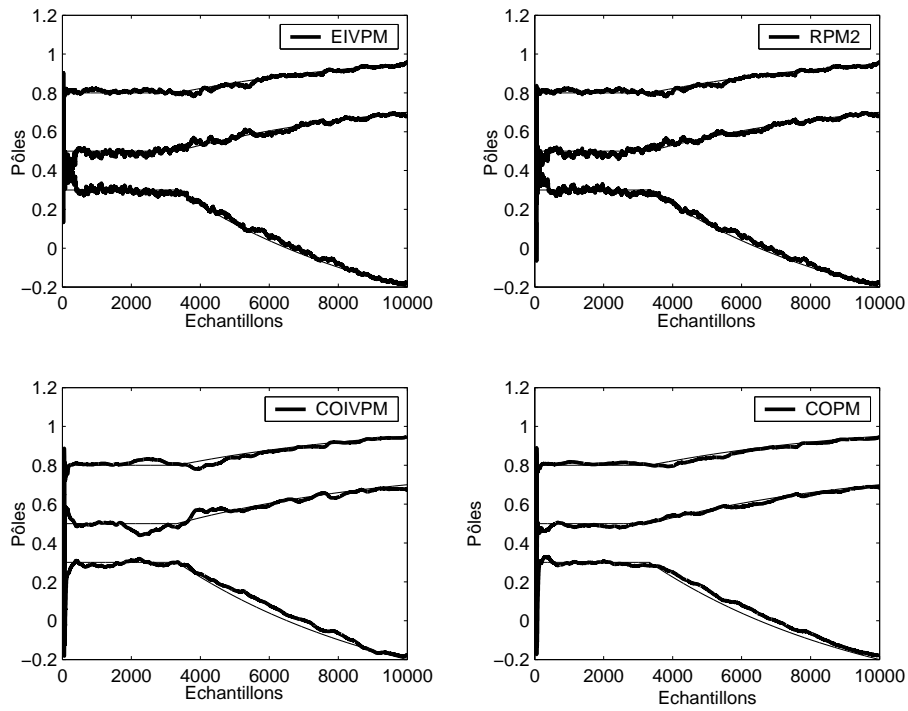


FIG. 6.22: Pôles estimés à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM.
 $f = p = 8$. $\lambda_{final} = 0.995$.

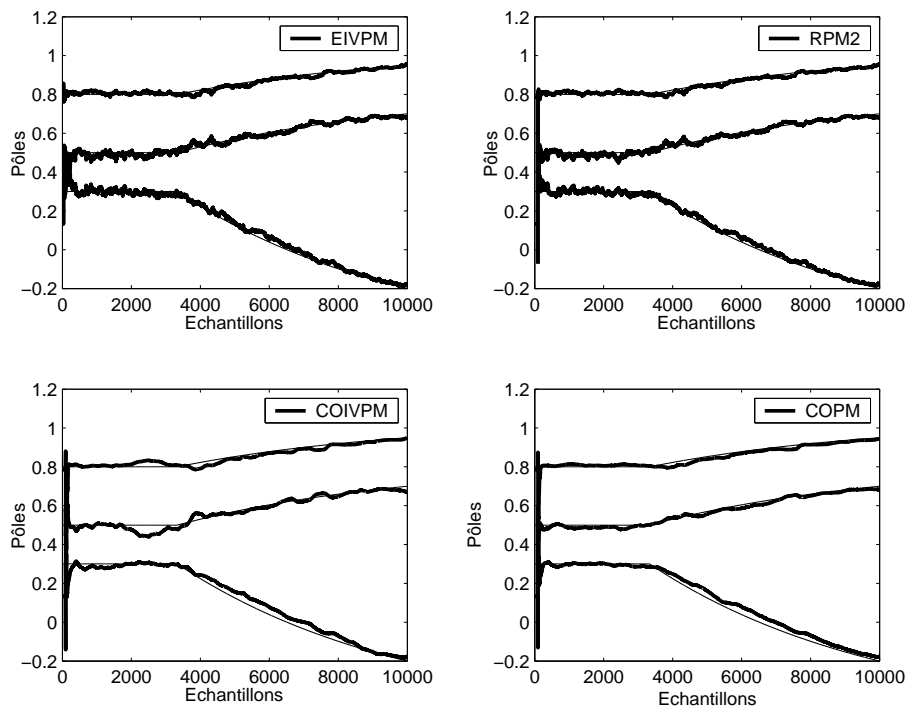


FIG. 6.23: Pôles estimés à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM.
 $f = p = 20$. $\lambda_{final} = 0.995$.

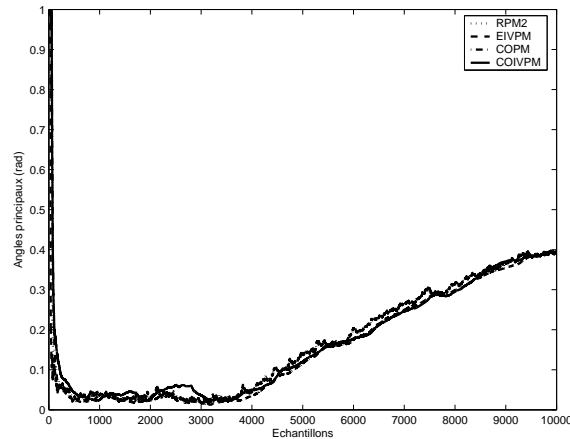


FIG. 6.24: Angles entre la matrice d'observabilité théorique initiale et les matrices d'observabilité estimées à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM.

être rencontré lorsque, par exemple, les valeurs des pôles du système se chevauchent, ne serait-ce que ponctuellement. Une étude du comportement des méthodes EIVPM, RPM2, COIVPM et COPM sous ce type de contraintes est proposée dans ce paragraphe. Le système (6.1) est modifié de telle sorte que les deux plus grands pôles du procédé valent tous deux 0.62 aux alentours du 6500^e échantillon. Les pôles estimés à l'aide de chaque algorithme récursif sont présentés figure 6.25. Celle-ci indique clairement que les techniques COIVPM et COPM sont plus à même de gérer ce type de situation. Il est cependant important de noter qu'une fois le chevauchement passé, les méthodes étudiées ont tendance à considérer que le pôle est toujours double et ce, pendant près de 1000 échantillons. Ces courbes mettent ainsi en évidence les limites des capacités d'oubli des algorithmes PM. COIVPM permet toutefois, sur cet exemple, de détecter l'instant de croisement.

6.1.2.3 Cas 3 : $A = A_{var_3}$

Bien que certaines hypothèses nécessaires à la construction des algorithmes d'identification récursive ne soient plus vérifiées lorsque les évolutions subies par le système à identifier sont brutales, il est important d'analyser le comportement des techniques EIVPM, RPM2, COIVPM et COPM lorsque l'un des pôles du système croît, par exemple, de 0.15 unité en une période d'échantillonnage. Les courbes de la figure 6.26 indiquent que, pour $\lambda = 0.995$, les quatre méthodes sont capables de détecter le changement rapide de valeur du second pôle du processus. Comme précédemment les algorithmes COIVPM et COPM réagissent différemment des techniques EIVPM et RPM2. La vitesse de convergence de ces dernières, mise en évidence sur le tracé des angles principaux de la figure 6.27, se paye au détriment d'une variance plus forte. La méthode fournissant dans cette situation la meilleure estimation d'un point de vue juste est COPM. Il est possible d'accélérer la vitesse de convergence de cette dernière en diminuant le facteur d'oubli à 0.99 tout en conservant une certaine qualité des estimées (cf. fig. 6.28).

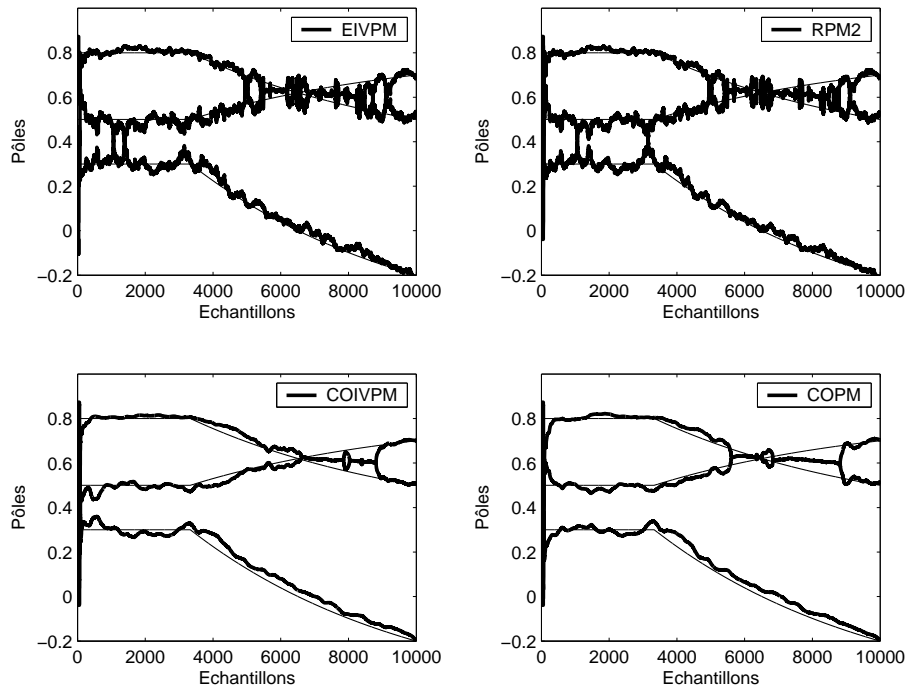


FIG. 6.25: Pôles estimés à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM.
 $\lambda = 0.995$.

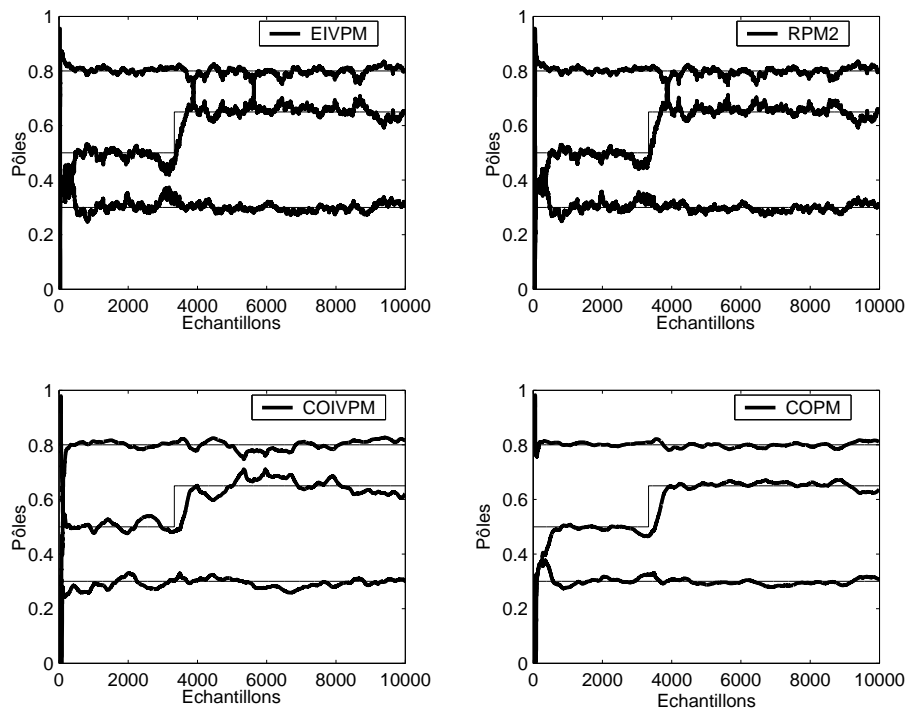


FIG. 6.26: Pôles estimés à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM.
 $\lambda = 0.995$.

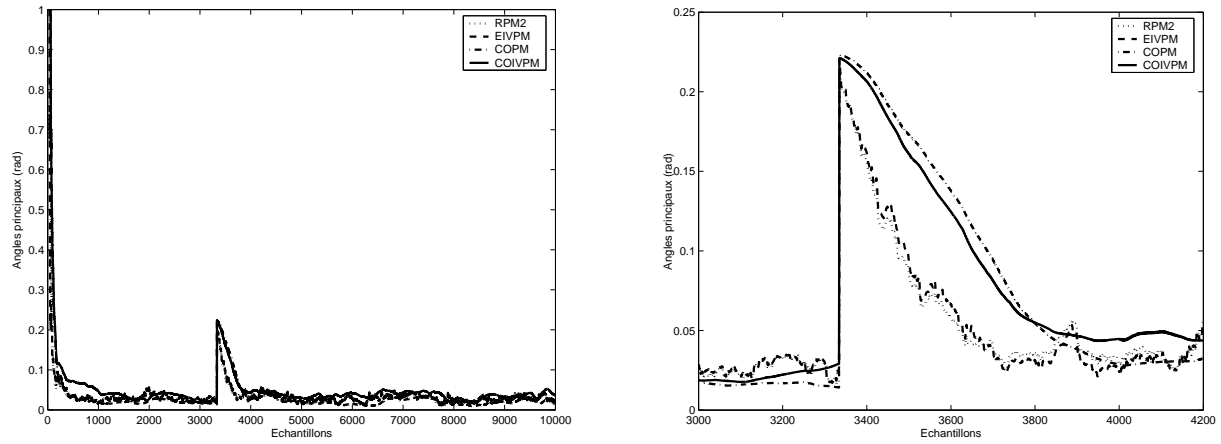


FIG. 6.27: Angles entre la matrice d'observabilité théorique et les matrices d'observabilité estimées à l'aide des algorithmes EIVPM, RPM2, COIVPM et COPM.
 $\lambda_{final} = 0.995$.

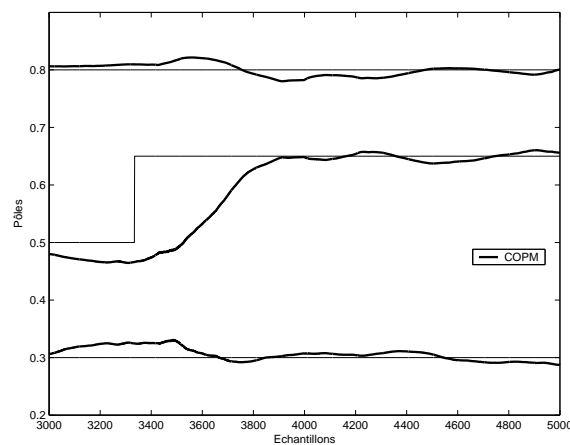


FIG. 6.28: Pôles estimés à l'aide de l'algorithme COPM. $\lambda = 0.99$.

6.1.3 Conclusion

Cette étude sur système simulé a eu pour objectif d'illustrer les performances des algorithmes PM sur des données stationnaires et non stationnaires. Une comparaison des modèles estimés par les méthodes PAST et PM ainsi qu'un examen de l'influence des paramètres de construction λ , f , p et n_x ont été proposés. Les enseignements de cet analyse sont les suivants :

- les algorithmes PM fournissent des modèles aussi fiables que les méthodes fondées sur l'approximation PAST ;
- la vitesse de convergence des techniques PM est légèrement supérieure à celle des algorithmes PAST pour un coût numérique plus faible ;
- les algorithmes PM sont capables de produire des modèles stables ayant un comportement proche de celui du système à identifier lors d'une faible sous-estimation ou surestimation de l'ordre ;
- le facteur d'oubli joue un rôle fondamental dans la capacité de poursuite et la sensibilité aux perturbations des méthodes récursives des sous-espaces ;
- les paramètres de construction f et p sont peu influents sur la consistance des estimées ;
- les méthodes PM sont capables de fournir des modèles à paramètres évolutifs pour suivre des systèmes lentement variants.

Il nous semble désormais primordial d'analyser le comportement de ces nouvelles techniques d'identification récursive sur des données réelles. Une étude du comportement des algorithmes PM à l'aide des mesures acquises sur deux maquettes de procédé est donc présentée dans la section suivante.

6.2 Étude de systèmes réels

Les méthodes PM développées dans cette thèse ont montré des possibilités intéressantes sur le système simulé précédent. Nous proposons de compléter cet examen en appliquant ces algorithmes sur les données des systèmes suivants :

- une maquette de transmission flexible représentative de nombreux procédés mécaniques ;
- un circuit électronique de type passe bas ayant la propriété de pouvoir changer de régime, continûment ou par saut, à l'aide d'une tension de commande variable.

Ces systèmes vont ainsi permettre de mettre en évidence les performances ainsi que les limites des techniques PM sur des données réelles.

6.2.1 Étude d'un système de transmission flexible

Le système de transmission flexible étudié est constitué de trois poulies reliées par deux courroies métalliques comportant chacune deux ressorts (système du Laboratoire d'Automatique et Procédés LAP de Caen [Ton00] (cf. fig. 6.29). La première poulie est commandée par un moteur à courant continu asservi en position et en vitesse par une boucle locale. La position de la poulie 3 est mesurée à l'aide d'un potentiomètre. Ce système a un comportement oscillatoire et possède deux modes résonants [LLM97, MC00]. Son principal intérêt réside dans le fait que cette maquette est représentative des systèmes stables à comportement oscillatoire ou pendulaire, systèmes souvent rencontrés en mécanique (grue par exemple). La séquence d'entrée appliquée sur ce procédé est une séquence binaire pseudo aléatoire (SBPA) de longueur 255. Les



FIG. 6.29: Un exemple de système de transmission flexible.

données d'entrée-sortie utilisées⁶ pour l'identification sont présentées figure 6.30. La période d'échantillonnage est fixée à 0.05s.

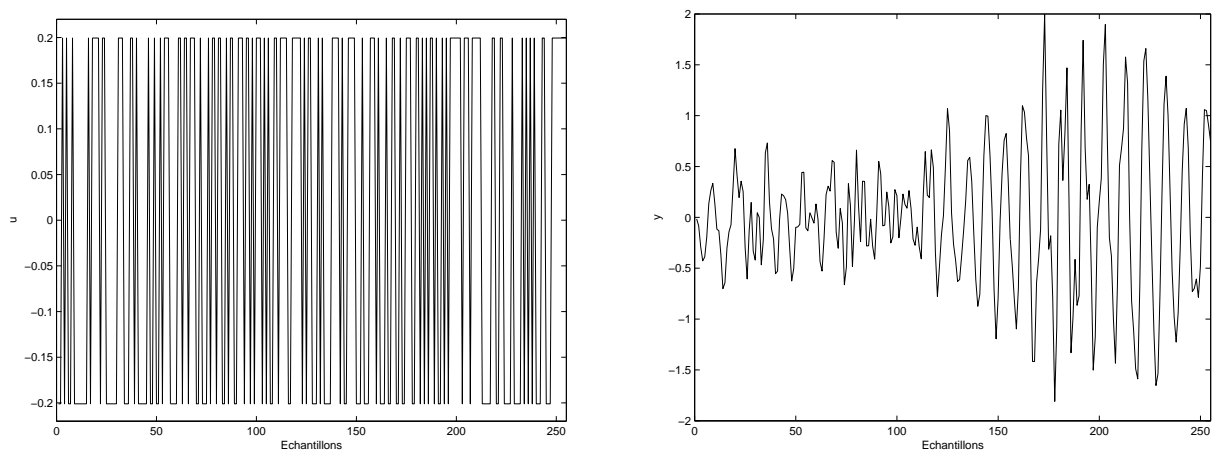


FIG. 6.30: Données d'entrée-sortie du système flexible.

La première partie de l'étude consiste à comparer les estimées obtenues en utilisant les méthodes d'identification récursive EIVPM, RPM2, COPM et COIVPM. Comme précédemment, il est nécessaire de fixer un certain nombre de paramètres pour effectuer l'identification.

- L'ordre du système est *a priori* connu puisque, par simple modélisation physique (cf. annexe B), le processus considéré possède deux modes oscillants. n_x vaut donc 4.
- L'analyse du système de simulation de la section 6.1 a montré que les indices f et p devaient être choisis légèrement supérieurement à n_x . Nous posons $f = p = 8$.
- Le facteur d'oubli est ici utilisé pour oublier le plus rapidement possible les informations liées aux matrices d'état initiales construites au hasard. Il est donc choisi comme suit :

$$\lambda(t) = \min \{ \lambda_0 \lambda(t-1) + 1 - \lambda_0, \lambda_{final} \}. \quad (6.13)$$

⁶Les algorithmes récursifs sont appliqués aux données d'un fichier. Il ne s'agit pas d'une identification en temps réel réalisée simultanément à l'acquisition.

avec :

$$\lambda_0 = 0.99 \quad (6.14)$$

$$\lambda(0) = 0.99 \quad (6.15)$$

$$\lambda_{final} = 0.999. \quad (6.16)$$

– La variable instrumentale est quant à elle une version retardée des entrées du procédé :

$$\xi = \begin{bmatrix} \mathbf{u}(t-p) \\ \vdots \\ \mathbf{u}(t-1) \end{bmatrix} = \mathbf{u}_p^- \quad (6.17)$$

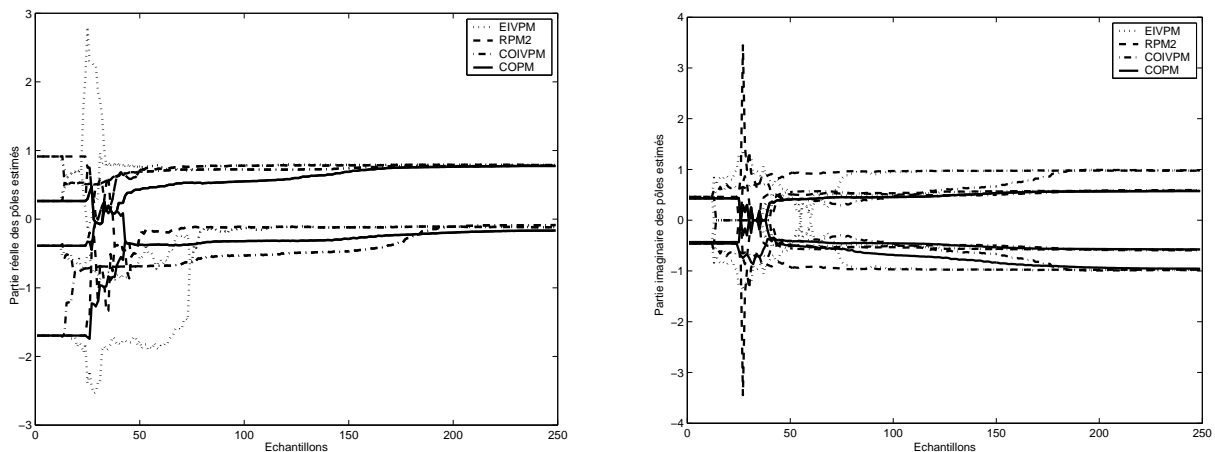


FIG. 6.31: Pôles estimés par les algorithmes EIVPM, RPM2, COPM et COIVPM.

Une fois ces paramètres choisis, comparons⁷ les estimées des pôles calculés à l'aide des algorithmes EIVPM, RPM2, COPM et COIVPM. Puisque le système mécanique considéré possède deux pôles complexes conjugués, il est nécessaire de représenter séparément les évolutions des parties réelles et parties imaginaires des pôles estimés afin de mettre en évidence la dynamique de convergence des méthodes étudiées. Les courbes de la figure 6.31 indiquent que les pôles estimés par les techniques analysées convergent tous vers les deux mêmes pôles complexes conjugués. La fiabilité de ces estimées ainsi que leur sensibilité à l'initialisation sont confirmées à l'aide d'un tirage de Monte Carlo (100 initialisations aléatoires). La dispersion des pôles déterminés à l'aide des algorithmes EIVPM, RPM2, COPM et COIVPM est en effet relativement faible (cf. fig. 6.32). Il est également à noter que la convergence des méthodes EIVPM et RPM2 est bien plus rapide que celle de COPM et COIVPM puisque les estimées obtenues avec EIVPM et RPM2 atteignent $0.8 \pm 0.55j$ et $-0.1 \pm 0.98j$ en 70 pas d'échantillonnage alors qu'il faut plus de 150 échantillons pour que COPM et COIVPM atteignent ces mêmes valeurs. Cette première analyse comparée montre donc que, sur ce système mécanique, les techniques EIVPM et RPM2 sont à privilégier.

⁷Comme nous l'avons précisé dans le paragraphe 6.1, le vecteur d'observation est estimé à l'aide des algorithmes ROM_{qr} et RPIPO_{qr}.

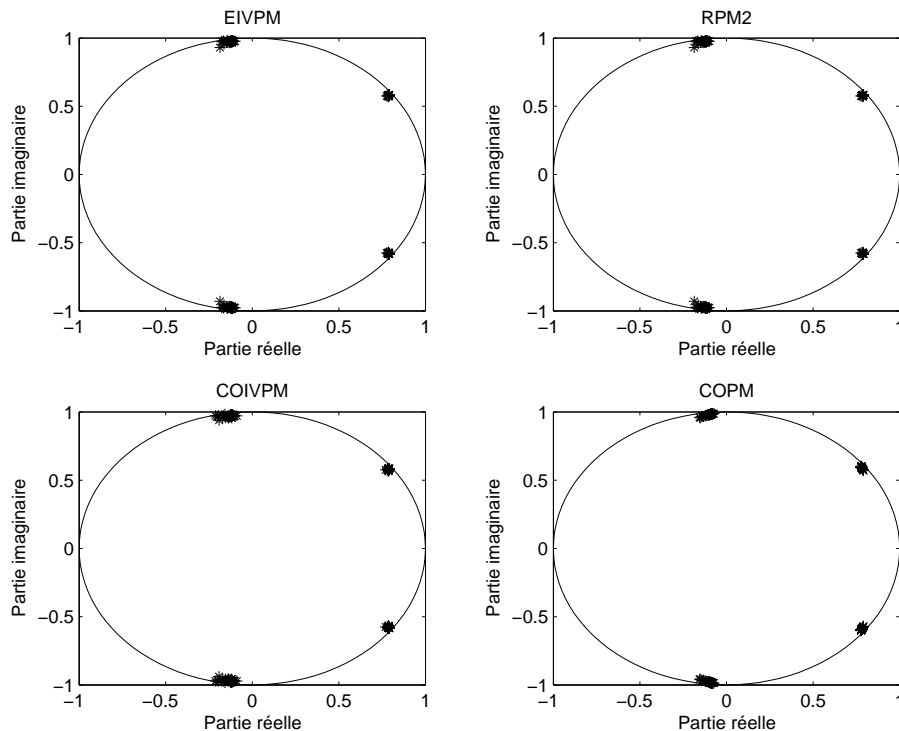


FIG. 6.32: Pôles estimés par les algorithmes EIVPM, RPM2, COPM et COIVPM.

Bien que toutes les techniques récursives analysées fournissent des estimées équivalentes, cette première étude ne permet pas d'affirmer que les résultats obtenus soient exacts. Afin de valider ces estimations, il est essentiel de comparer le modèle calculé à l'aide de ces algorithmes récursifs aux modèles estimés par d'autres méthodes d'identification plus classiques. Notre choix s'est plus particulièrement porté sur la méthode récursive d'erreur de prédiction⁸ RPEM [Lju99] et la technique des sous-espaces hors ligne MOESP. Il nous semble en effet intéressant de confronter les algorithmes récursifs des sous-espaces à une méthode adaptative classique ainsi qu'aux techniques hors ligne d'estimation dont ils s'inspirent. Par souci d'homogénéité, puisque la variable instrumentale utilisée par les algorithmes récursifs des sous-espaces est une version retardée des données d'entrée, la version MOESP choisie est PI MOESP. Le modèle estimé est alors calculé à l'aide des fonctions de la boîte à outils SMI 1.0 de B. Haverkamp et M. Verhaegen [HV97]. La fonction `rpem` de la Toolbox Identification de Matlab est quand à elle appliquée pour estimer un modèle récursif SISO. Enfin, afin de ne pas multiplier les comparaisons, il est choisi de n'utiliser qu'un seul algorithme récursif des sous-espaces. En raison des résultats mis en évidence par l'analyse comparée des techniques EIVPM, RPM2, COPM et COIVPM précédente, nous ne considérerons, dans la suite de ce paragraphe, que la méthode RPM2.

Il est important, dans un premier temps, de s'assurer de la justesse des pôles estimés à l'aide des méthodes récursives d'identification PM. Cette vérification peut être facilement réalisée en comparant les pôles déterminés respectivement par RPM2 et PI MOESP. La figure 6.33 indique que les pôles estimés à partir de RPM2 après tirage de Monte Carlo sont équivalents de

⁸RPEM : Recursive Prediction Error Method.

ceux fournis par PI MOESP. Les algorithmes d'identification récursive étudiés semblent donc produire des modèles consistants et très peu sensibles à l'initialisation des matrices d'état.

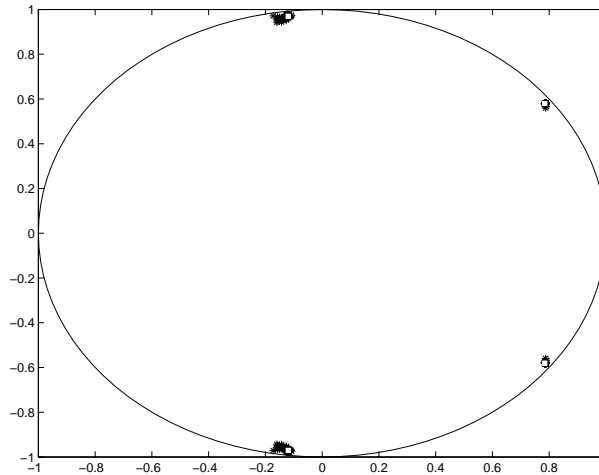


FIG. 6.33: Comparaison des pôles estimés par les algorithmes RPM2 (*) et PI MOESP (□).

Notre objectif étant de développer des algorithmes récursifs ayant une vitesse de convergence relativement rapide, il est important, dans un deuxième temps, de comparer la dynamique d'évolution des pôles des modèles estimés par RPM2 et RPEM. La figure 6.34 présente respectivement les variations temporelles des parties réelles et imaginaires de ces pôles. Ces courbes révèlent que, sur ce système mécanique, la vitesse de convergence de l'algorithme RPM2 est très proche de celle de la méthode RPEM.

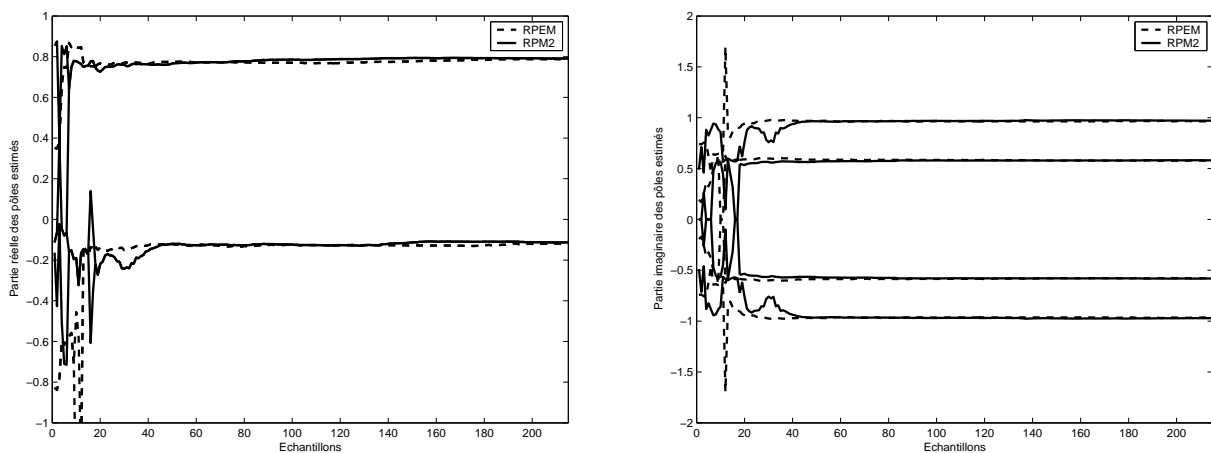


FIG. 6.34: Pôles estimés par les algorithmes RPM2 et RPEM.

Une façon simple de conforter ces premières conclusions consiste à tracer sur un même diagramme les réponses fréquentielles des modèles RPEM, RPM2 et PI MOESP. Ce tracé permet en effet de comparer les trois modèles sur toute la bande de fréquences utile du système. La figure 6.35 indique que leurs réponses fréquentielles sont similaires pour quasiment toutes les

pulsations calculées. Ces courbes permettent ainsi de conclure que le modèle estimé par l'algorithme récursif RPM2 est relativement⁹ fiable pour presque toutes les fréquences accessibles.

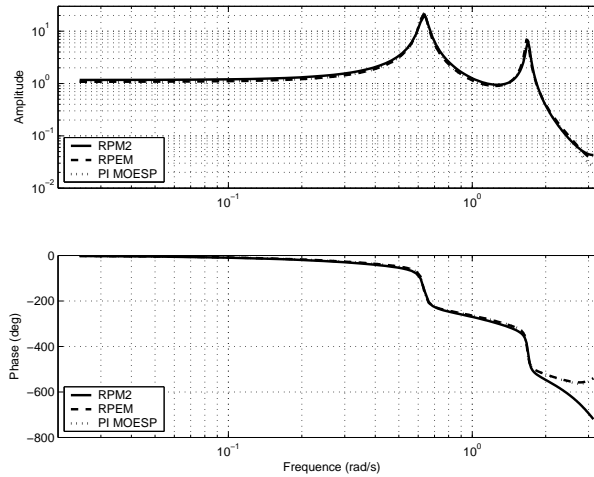


FIG. 6.35: Comparaison des réponses fréquentielles des modèles estimés par RPM2, RPEM et PI MOESP.

Une dernière technique de validation qu'il est intéressant de proposer consiste à comparer la sortie réelle du procédé aux sorties simulées des modèles RPM2, RPEM et PI MOESP. Cette comparaison est présentée sur la figure 6.36. Bien que les courbes d'erreur $y - \hat{y}_m$ indiquent que le modèle RPM2 connaît plus de difficultés à reproduire la sortie réelle du système que les modèles RPEM et PI MOESP, il est intéressant de noter que le pourcentage de correspondance défini par :

$$100 * \left(1 - \frac{\|y - \hat{y}_m\|}{\|y - \frac{1}{N} \sum_{t=1}^N y(t)\|} \right) \quad (6.18)$$

est relativement élevé¹⁰ pour RPM2 puisque ce dernier vaut 82%.

Remarque 6.2. *Il est à noter que les résidus de la figure 6.36 présentent, dans les trois cas, un caractère non stationnaire très marqué. Ces sauts de variance sont vraisemblablement dus à la présence de non linéarités dans le système étudié.*

Conclusion

Cette première étude sur système réel a permis de mettre en évidence que :

- les algorithmes PM sont capables de fournir un modèle consistant du système mécanique considéré avec une certaine reproductibilité ;
- les méthodes basées sur des algorithmes séquentiels sont plus lents à converger que les autres techniques ;
- l'algorithme récursif RPM2 conduit à des modèles presque aussi performants que ceux obtenus à l'aide de techniques hors ligne des sous-espaces et ce, avec un coût calculatoire bien plus faible.

⁹... aux modèles RPM2 et PI MOESP.

¹⁰Il vaut 90% pour le modèle PI MOESP et 92% pour le modèle RPEM.

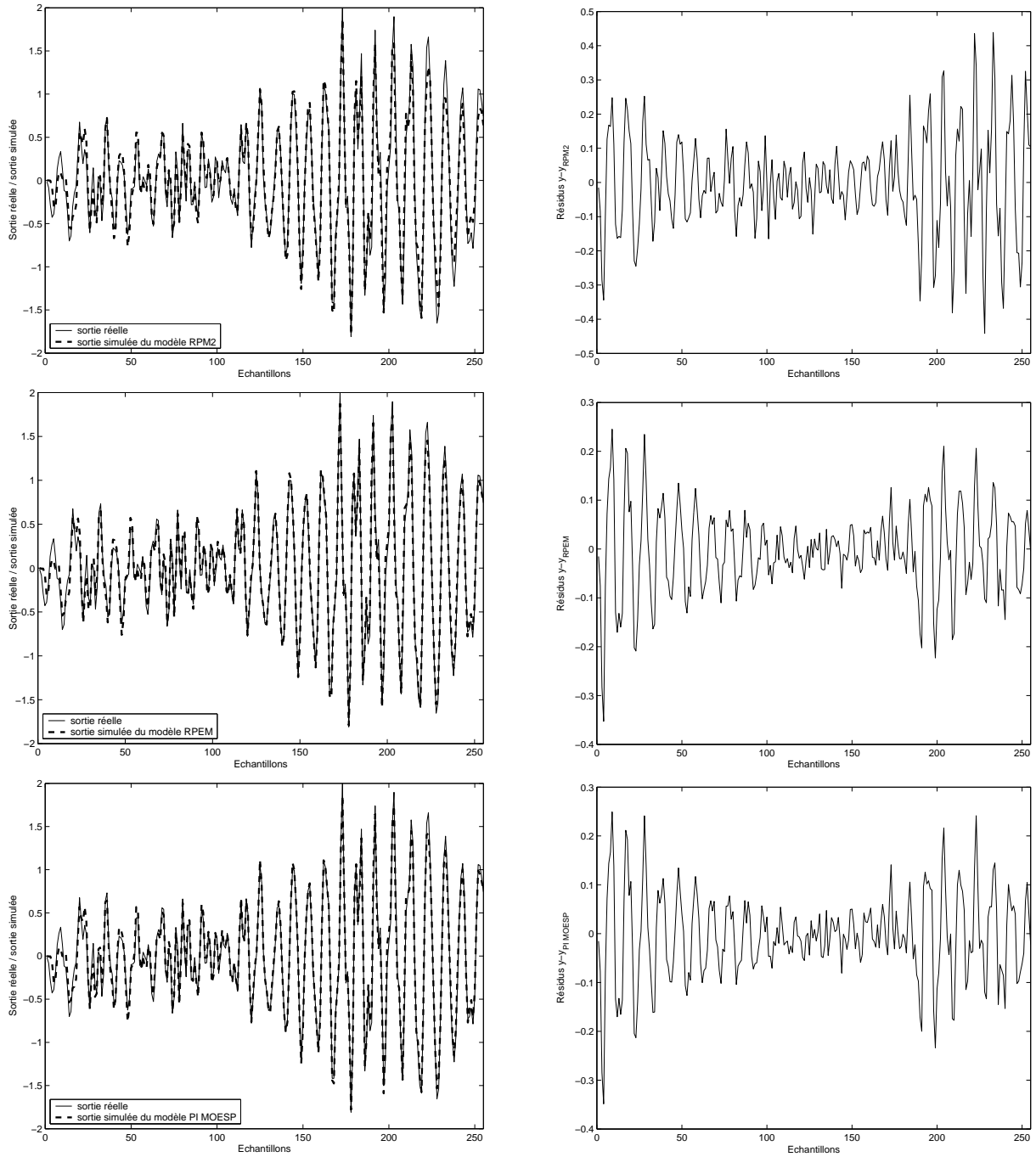


FIG. 6.36: Comparaison des sorties réelles et simulées à l'aide des modèles RPM2, RPEM et PI MOESP.

6.2.2 Étude d'un système électronique

Le système considéré (cf. fig. 6.37) est un circuit électronique passe bas du second ordre de type Sallen Key [Bil97] dont la dynamique est commandée à l'aide d'une tension continue $0 - 10V$. Cette propriété constitue le principal intérêt de ce montage. En choisissant correctement les composants passifs de ce circuit, il est en effet possible de faire évoluer, de manière continue, le comportement du système d'un régime aperiodique à un régime pseudo periodique. Les pôles de la fonction de transfert passe alors continûment de pôles réels à pôles complexes conjugués.

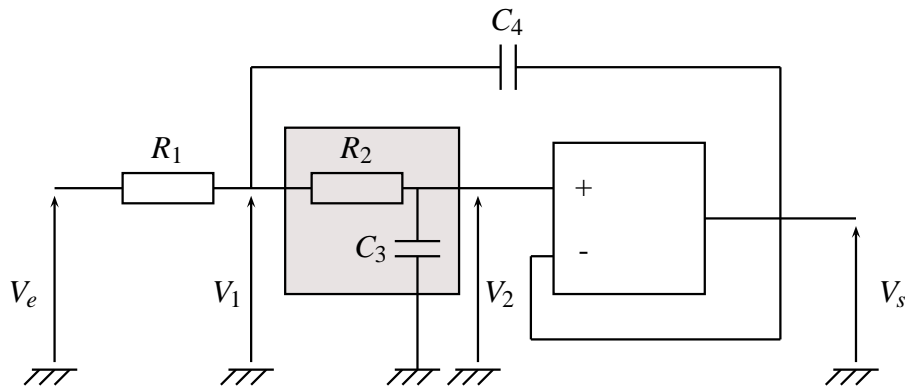


FIG. 6.37: Circuit électronique du second ordre

La tension commandant cette évolution intervient au niveau du filtre $\{R_2, C_3\}$. La technique utilisée pour réaliser cette commande consiste à associer au filtre $\{R_2, C_3\}$ un multiplieur AD633. Le schéma électronique de ce circuit est présenté sur la figure 6.38. La relation liant les tensions V_s et V_e du circuit 6.38 s'écrit alors :

$$V_2(p) = \frac{1}{1 + \frac{10R_2C_3}{E_c} p} V_1(p) \quad (6.19)$$

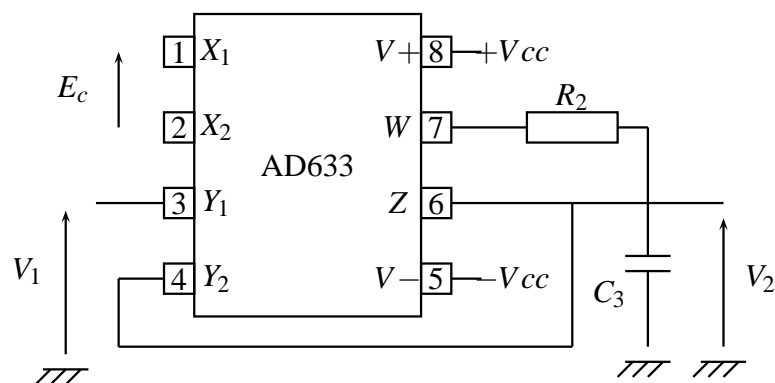


FIG. 6.38: Filtre passe bas commandé en tension

La constante de temps du filtre passe bas $\{R_2, C_3\}$ peut ainsi être commandée via E_c . Bien que, théoriquement, nous puissions avoir accès à cette constante et, par extrapolation, aux valeurs R_2

et C_3 , l'introduction du circuit AD633 au sein du filtre de Sallen Key modifie grandement ces propriétés. Il devient alors impossible de calculer *a priori* les valeurs théoriques de la pulsation propre et du facteur d'amortissement du circuit considéré. Il est toutefois toujours possible de modifier la dynamique d'un circuit du second ordre de manière continue à l'aide de la tension de commande E_c .

Avant de tenter d'identifier ce circuit électronique, il est important d'en connaître certaines caractéristiques pour correctement établir la procédure d'acquisition des données d'entrée-sortie. Une façon simple d'obtenir, par exemple, le temps de réponse du système, consiste à tracer les réponses indicielles du procédé pour différentes valeurs de E_c . Ces dernières sont présentées figure 6.39. Nous pouvons ainsi constater que le système étudié peut présenter indifféremment un comportement apériodique ou pseudo périodique. Le tableau 6.1 regroupe les temps de réponse à 95% du circuit électronique considéré pour E_c variant de 1 à 10V par pas de 0.5V (noté désormais $E_c \in [1 : 0.5 : 10]$). Ces valeurs permettront de fixer la durée des paliers de la séquence d'excitation.

| Tension de commande E_c (V) | Temps de réponse à 95% (ms) |
|-------------------------------|-----------------------------|
| 1 | 11.5 |
| 1.5 | 7.6 |
| 2 | 5.5 |
| 2.5 | 4.3 |
| 3 | 3.5 |
| 3.5 | 3 |
| 4 | 2.5 |
| 4.5 | 2.2 |
| 5 | 3.2 |
| 5.5 | 3.1 |
| 6 | 3.2 |
| 6.5 | 3.1 |
| 7 | 3.1 |
| 7.5 | 3 |
| 8 | 2.9 |
| 8.5 | 2.8 |
| 9 | 2.7 |
| 9.5 | 2.6 |
| 10 | 2.6 |

TAB. 6.1: Temps de réponse à 95% du circuit électronique pour $E_c \in [1 : 0.5 : 10]$

La première phase de l'étude consiste à étudier le comportement des algorithmes EIVPM, RPM2, COPM et COIVPM sur des données stationnaires. Nous excitons pour cela le système électronique à l'aide d'une séquence binaire pseudo aléatoire d'amplitude unitaire dont la durée maximale de palier est choisie égale au temps de réponse à 95% du procédé pour la tension de commande considérée. La période d'échantillonnage employée (en émission et en acquisition) vaut 0.0002s. Il s'agit de la plus petite valeur que nous puissions obtenir avec la carte d'acqui-

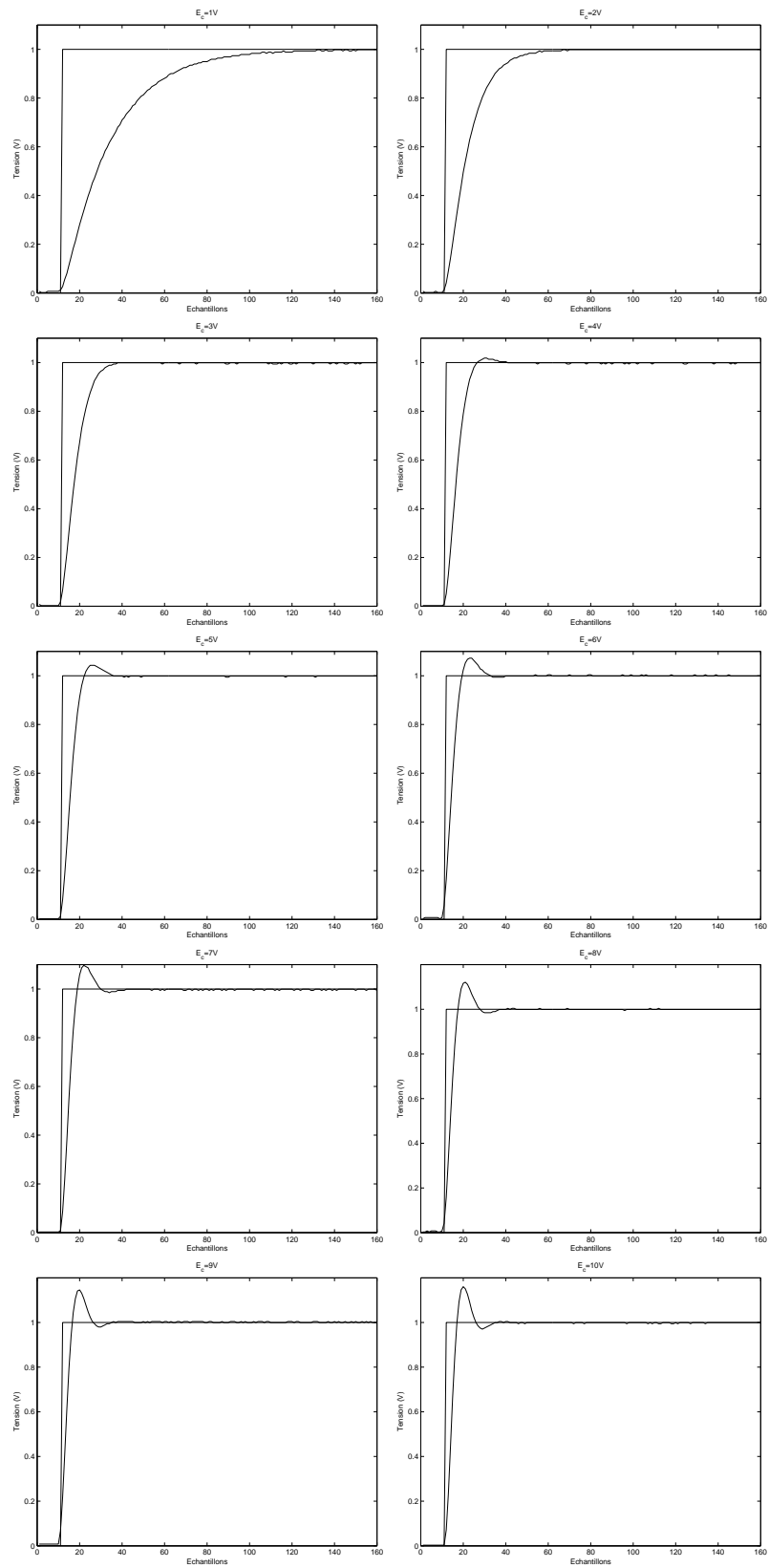


FIG. 6.39: Réponses indicielles du circuit électronique pour $E_c \in [1 : 1 : 10]$.
Système invariant.

sition utilisée. Les paramètres des algorithmes d'identification récursive sont les suivants :

$$n_x = 2 \quad (6.20)$$

$$f = p = 8 \quad (6.21)$$

$$\lambda(t) = \min \{ \lambda_0 \lambda(t-1) + 1 - \lambda_0, \lambda_{final} \} \text{ avec } \lambda_0 = \lambda(0) = 0.99 \text{ et } \lambda_{final} = 0.999 \quad (6.22)$$

$$\xi = \mathbf{u}_p^- \quad (6.23)$$

Puisqu'il nous est malheureusement impossible de connaître *a priori* les valeurs théoriques des pôles du circuit, nous proposons parallèlement d'identifier ce système, pour $E_c \in [1 : 0.25 : 10]$, à l'aide d'une méthode non récursive afin d'obtenir un ensemble de modèles référents. Comme dans le cas du système de transmission flexible, la méthode choisie est PI MOESP. Les modèles utilisés sont des modèles moyens obtenus après un tirage de Monte Carlo (tirage aléatoire des initialisations et des couples de données d'entrée-sortie).

L'expérience réalisée pour s'assurer de la fiabilité des algorithmes EIVPM, RPM2, COPM et COIVPM consiste à :

1. exciter le système électronique à l'aide d'une SBPA et enregistrer sa réponse pour $E_c \in [1 : 0.25 : 10]$;
2. appliquer à chacun de ces 37 enregistrements 100 fois les quatre algorithmes étudiés avec initialisation aléatoire ;
3. reporter sur un même graphique la gamme des valeurs prises par les parties réelles (resp. imaginaires) des pôles estimés pour chaque tension E_c considérée ;
4. comparer ses échelles de valeurs aux pôles des modèles référents estimés par PI MOESP.

Les courbes ainsi obtenues sont présentées figure 6.40. Elles indiquent que :

- les quatre techniques présentent une sensibilité relative à l'initialisation de l'ordre de 0.03 ;
- la méthode COPM possède une variance légèrement plus élevée pour les parties réelles des pôles les plus faibles correspondant à $E_c < 2.5$;
- les pôles obtenus à l'aide de l'algorithme PI MOESP sont inclus dans les valeurs estimées par EIVPM, RPM2, COPM et COIVPM.

Ces premiers résultats montrent que les algorithmes considérés sont fiables puisqu'ils sont capables de produire des estimés proches de la valeur théorique probable avec une bonne reproductibilité.

Cette analyse doit cependant être complétée par une étude de leur régime transitoire. Il est en effet important de proposer des méthodes pouvant fournir, en peu d'échantillons, un bon modèle de procédé. La figure 6.41 présente l'évolution, au cours de la récursion, des parties réelles des pôles estimés par les techniques EIVPM, RPM2, COPM et COIVPM. Ces courbes montrent que les méthodes les plus rapides à converger sont EIVPM et RPM2. Ceci est particulièrement flagrant pour des valeurs de E_c faible puisqu'il faut, pour $E_c = 1V$, 1800 échantillons pour que EIVPM et RPM2 atteignent 0.7 alors que 5000 et 11500 itérations sont nécessaires pour que les pôles estimés par COIVPM et COPM égalent cette valeur. Ce constat est vraisemblablement la conséquence du caractère séquentiel des algorithmes COIVPM et COPM. Cette différence s'estompe lorsque la tension de commande augmente. Un tel inconvénient les rend, sur ce système, inexploitable pour des données non stationnaires. Ils seront donc écartés pour la deuxième phase de l'analyse.

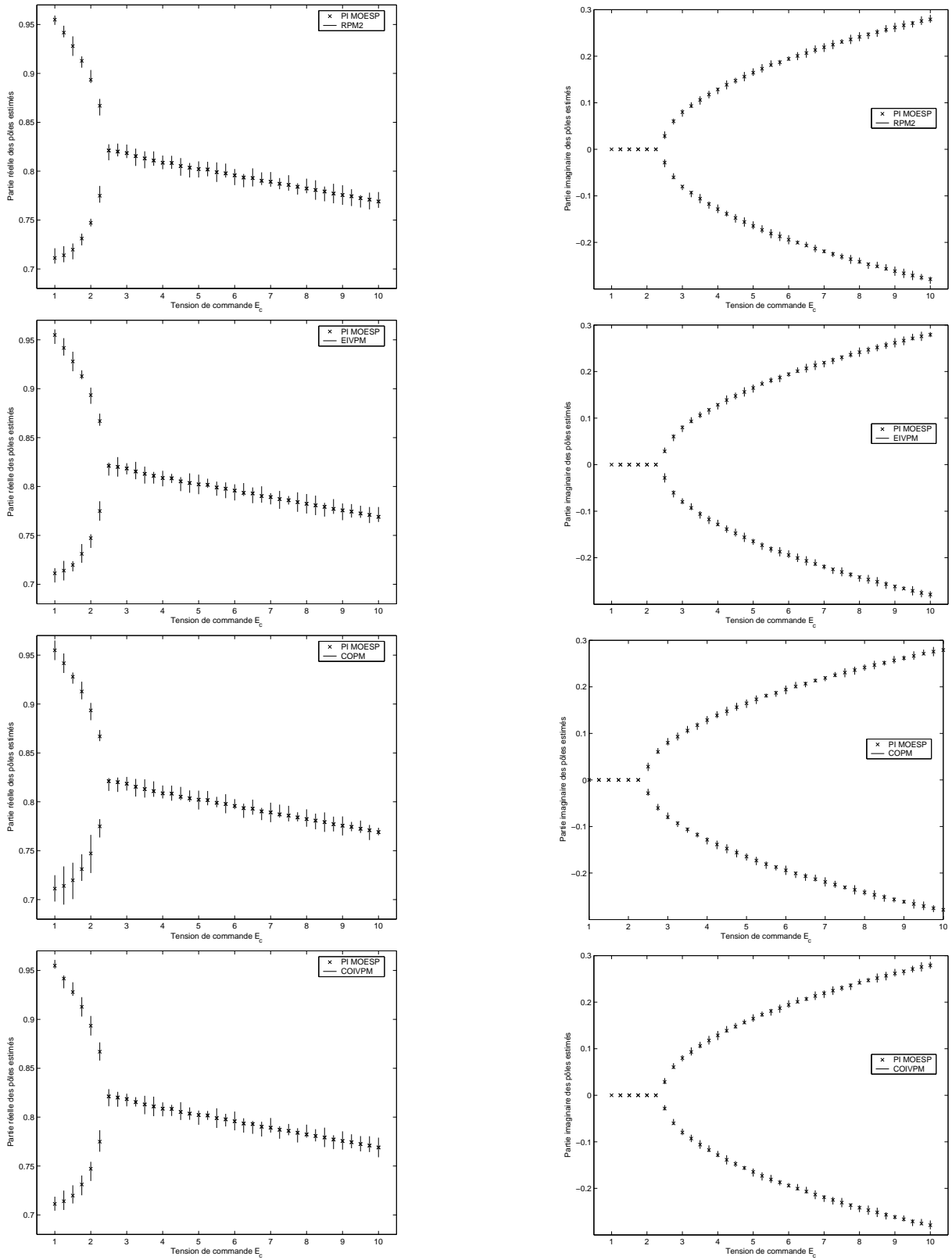


FIG. 6.40: Variance des pôles estimés par les algorithmes EIVPM, RPM2, COPM et COIVPM. Système invariant.

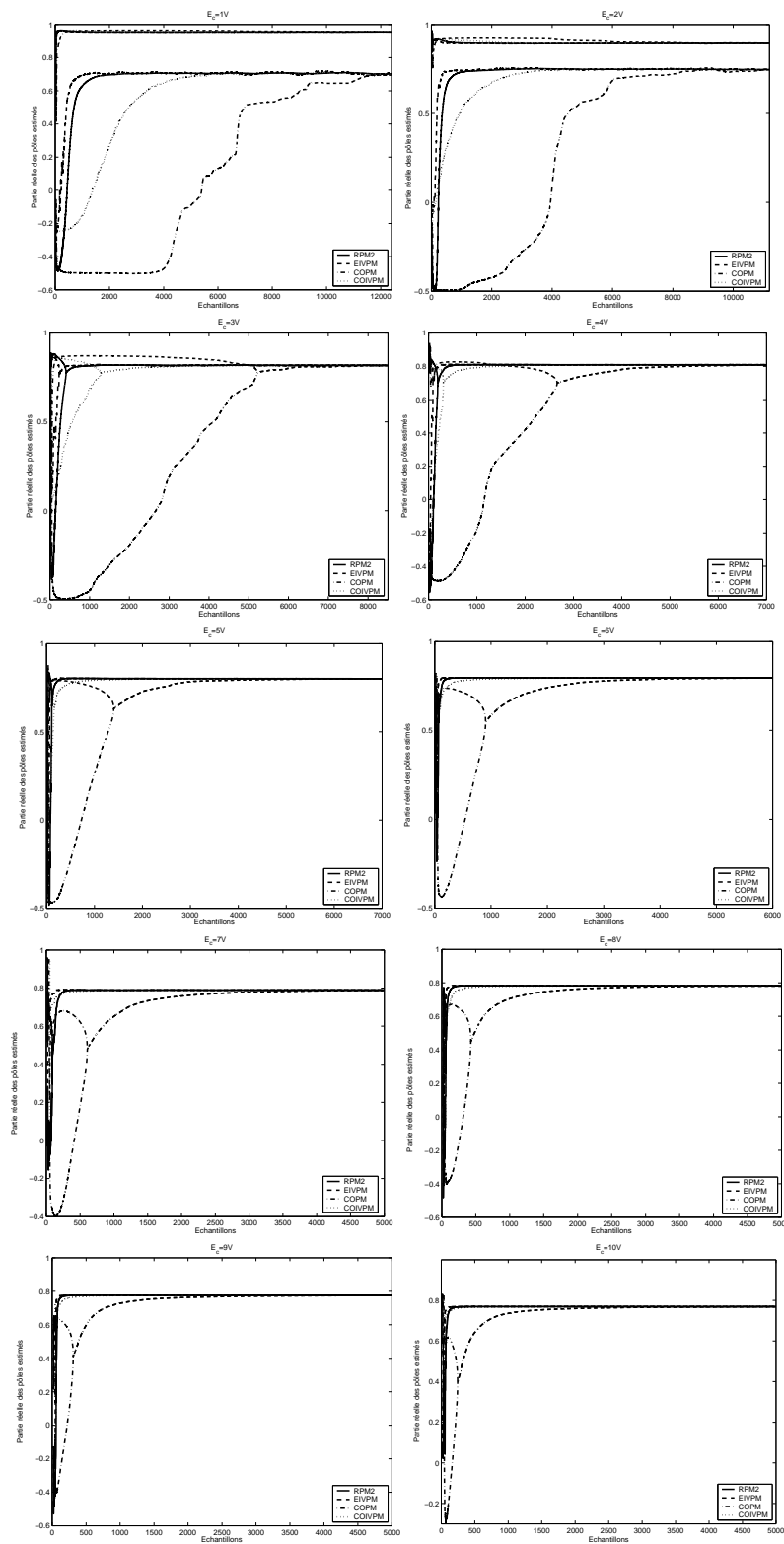


FIG. 6.41: Partie réelle des pôles estimés par les algorithmes EIVPM, RPM2, COPM et COIVPM $E_c \in [1 : 1 : 10]$. Système invariant.

La dernière étape de cette étude expérimentale a pour objectif de mettre en évidence les capacités de poursuite des algorithmes EIVPM et RPM2 ainsi que leurs limites lorsque ceux-ci sont employés sur des données non stationnaires. La génération de ces dernières est réalisée en appliquant aux bornes $\{X_1, X_2\}$ du multiplieur AD633 (cf. fig. 6.38) des tensions sinusoïdales d'amplitude variant entre 1 et 10V et de fréquence respective 0.05Hz, 0.1Hz et 0.5Hz. Ces fréquences d'évolution ont été choisies en fonction des résultats obtenus sur les données stationnaires. Les expériences réalisées précédemment ont en effet montré que les deux méthodes étudiées fournissaient, pour une initialisation aléatoire, des estimées fiables des pôles du système en moins de 0.4s (2000 échantillons). Les durées d'évolutions proposées (resp. 10, 5 et 1s) pour balayer l'ensemble de la gamme des tensions de commande applicables au procédé peuvent donc être considérées comme relativement rapides. Comme dans le cas stationnaire, la tension d'entrée du circuit électronique est une SBPA dont la durée de palier maximum est fixée à 14ms. Cette valeur permet ainsi de s'assurer que la bande passante du système soit correctement excitée quel que soit la tension E_c appliquée. Les données d'entrée-sortie pour une tension de commande E_c de fréquence 0.5Hz sont représentées figure 6.42. La courbe de tension de sortie illustre clairement les variations de régime subies par le circuit électronique. La valeur finale du facteur d'oubli, désormais égale à 0.995, est modifiée pour permettre de suivre ces changements. Les techniques EIVPM et RPM2 sont ensuite appliquées à ces mesures. Les figures 6.43, 6.44 et 6.45 présentent les variations des parties réelles des pôles estimés par ces deux méthodes pour $F_{sin} = 0.05Hz$, $F_{sin} = 0.1Hz$ et $F_{sin} = 0.5Hz$. Ces courbes indiquent que les techniques considérées sont capables de suivre, dans une certaine mesure, des variations rapides de dynamique. Le tracé des évolutions des parties réelles des pôles estimés suivent en effet assez bien les caractéristiques mises en évidence durant l'étude des données stationnaires (cf. fig. 6.40) puisque :

- pour $E_c \in [1 : 2.5]$, les pôles calculés sont réels et varient au sein des gammes $[0.82 : 0.95]$ et $[0.67 : 0.82]$;
- pour $E_c > 2.5$, les pôles estimés sont complexes conjugués et évoluent dans la gamme $[0.77 : 0.82]$.

Il est cependant important de noter que cette expérience souligne certaines de leurs limites. Les figures 6.43, 6.44 et 6.45 montrent en effet que les algorithmes EIVPM et RPM2 connaissent plus de difficultés à estimer correctement le pôle réel de faible valeur que son symétrique. L'estimation de ce pôle est en effet bien plus sujette aux perturbations que son opposé. Cette différence est accentuée avec la fréquence de la tension de commande. La figure 6.45 met particulièrement bien cette difficulté en évidence. La courbe d'évolution du pôle réel de faible valeur y est en effet très chaotique tandis que celle de son symétrique est tout à fait acceptable. Elle indique également que cet inconvénient est plus prononcé pour la technique EIVPM. Nous ne sommes malheureusement pas capables de justifier un tel comportement. Aucune de nos études sur la relation entre la variable instrumentale et les pôles du système à identifier n'a pour l'instant abouti. La difficulté est d'autant plus corsée qu'il n'est pas rare que ces algorithmes fournissent des estimées fiables de systèmes à pôles proches de zéro ! Nous pouvons seulement remarquer que ce désagrément n'est pas uniquement propre aux méthodes PM. Des courbes similaires sont obtenues lorsque les techniques EIVPAST et UDAPST sont appliquées à ces mêmes données (cf. fig. 6.46, 6.47 et 6.48). Il en est également de même avec des techniques d'identification récursive plus classiques telles que RPEM. Les courbes des figures 6.49, 6.50 et 6.48 révèlent effectivement que cet algorithme adaptatif estime péniblement le pôle de plus petite partie réelle. Bien que moins bruitées, les valeurs calculées semblent *a priori* en-

tachées d'erreur puisque la symétrie mise en évidence sur les figures précédentes disparaît. Ces constatations relativisent ainsi les difficultés rencontrées par les méthodes PM.

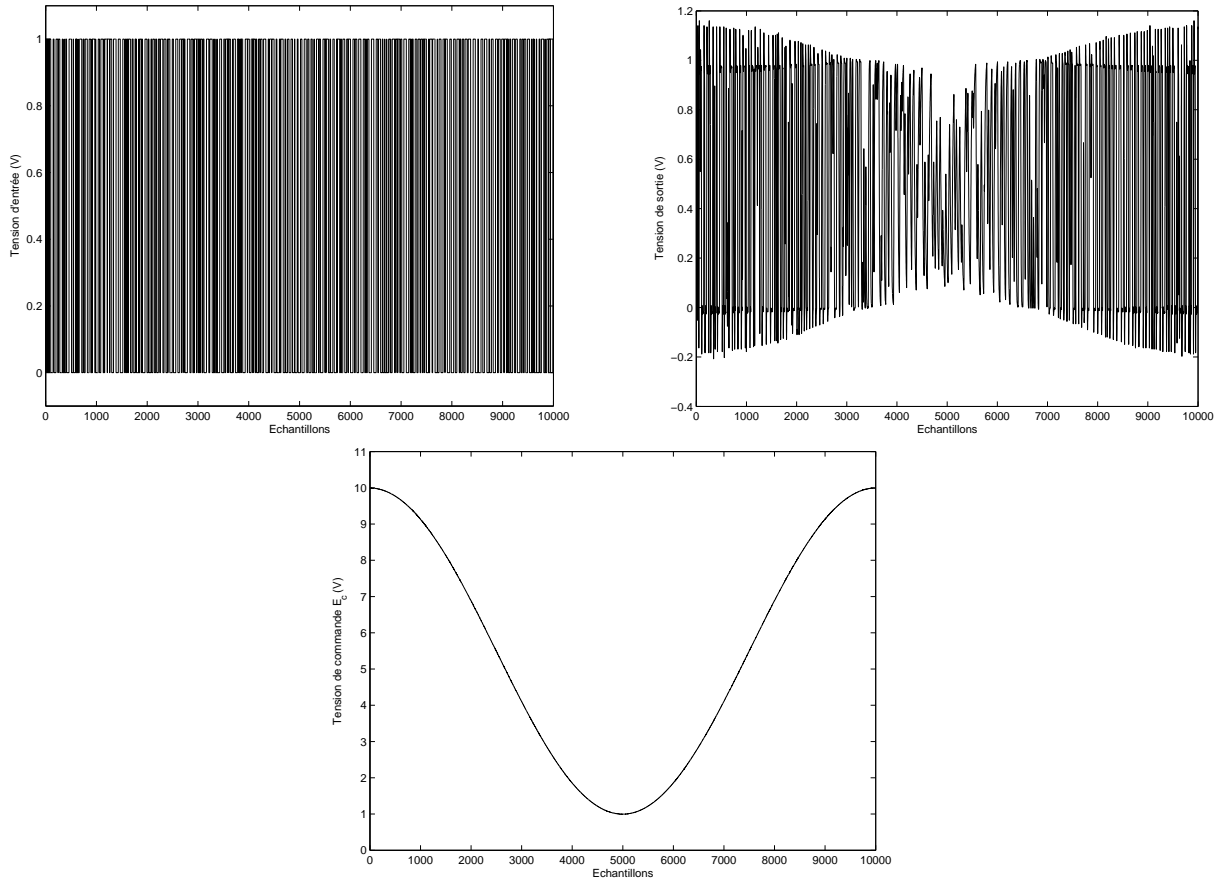
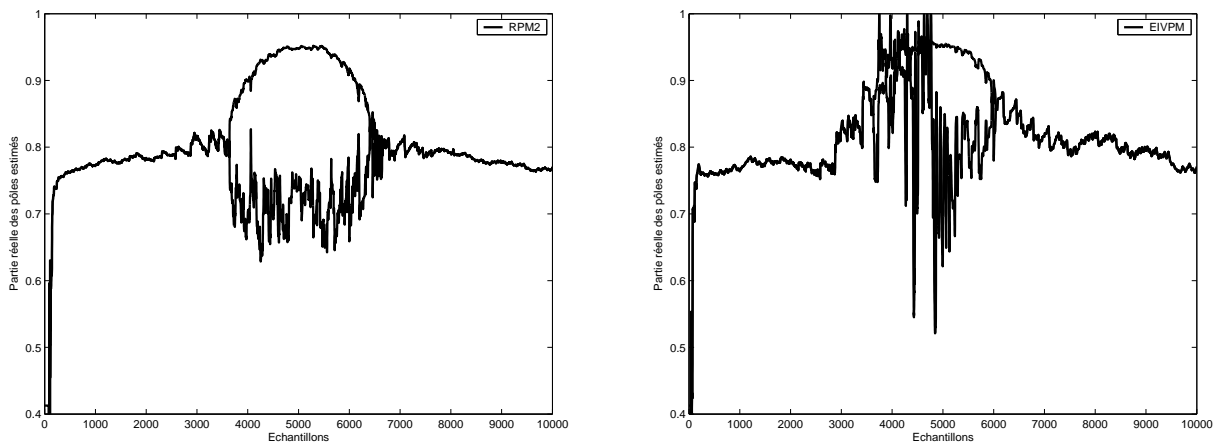
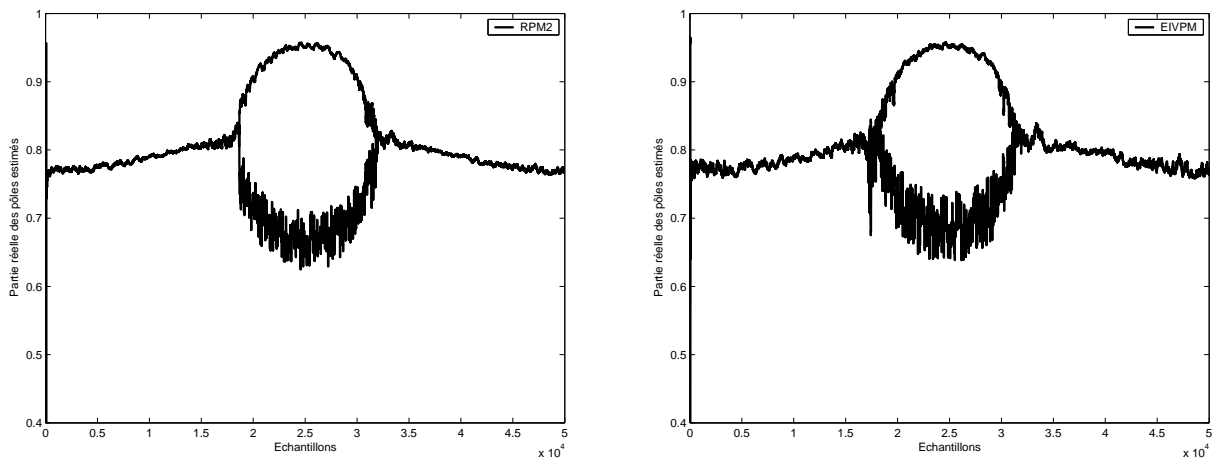
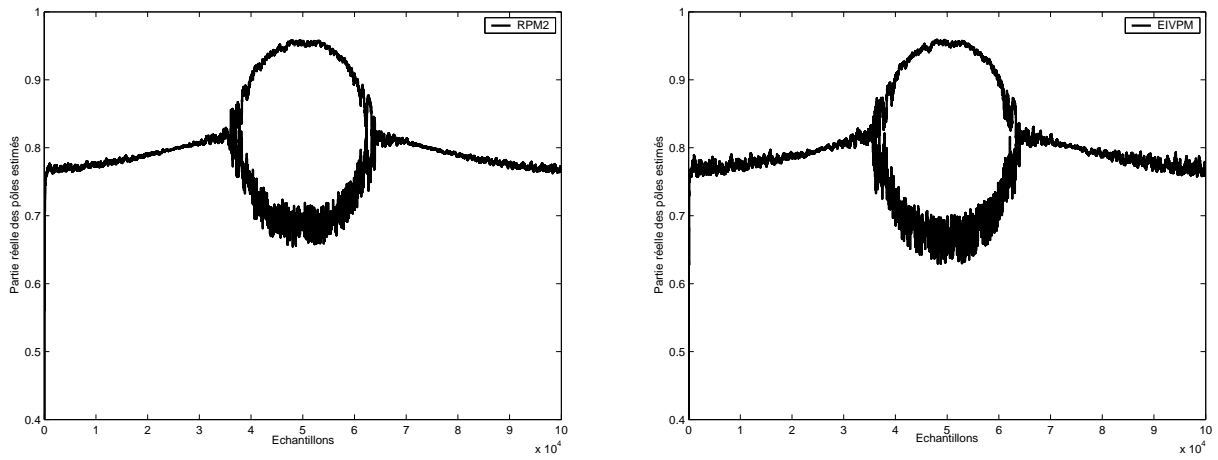


FIG. 6.42: Données d'entrée-sortie du circuit électronique. $F_{sin} = 0.5Hz$.

Conclusion

Le problème de l'identification d'un modèle de filtre passe bas a été considéré au sein de cette section. Le principal intérêt de ce montage réside dans le fait que sa dynamique peut être facilement modifiée à l'aide d'une tension continue. Plusieurs régimes de fonctionnement ont ainsi pu être étudiés. L'examen de ce système à l'aide de données stationnaires a montré que les algorithmes PM produisaient des modèles très proches de ceux obtenus en utilisant des méthodes hors lignes des sous-espaces. L'analyse du procédé variant a permis de mettre en évidence certaines limitations des méthodes d'identification récursive des sous-espaces.



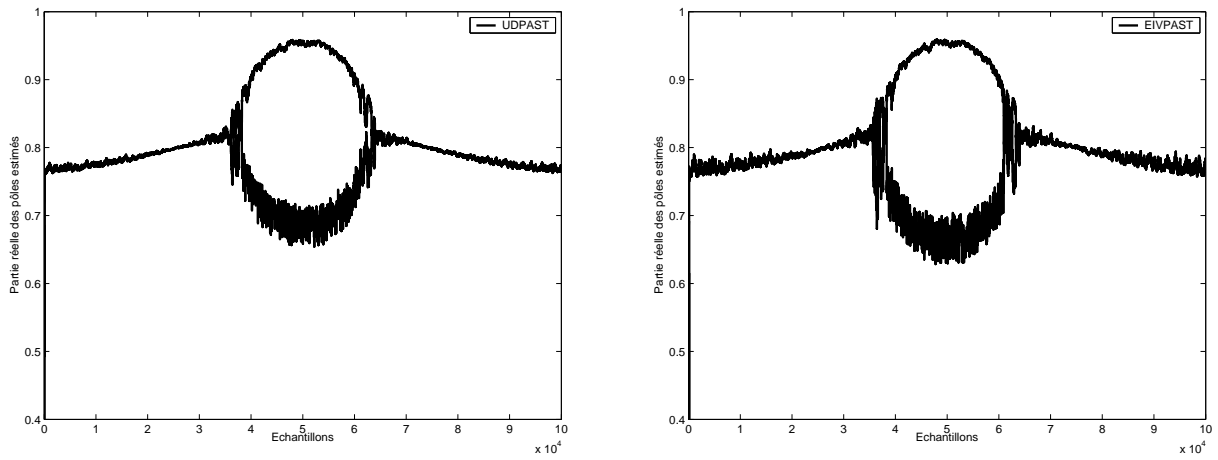


FIG. 6.46: Parties réelles des pôles estimés par les algorithmes EIVPAST et UDPAST. $F_{sin} = 0.05Hz$.

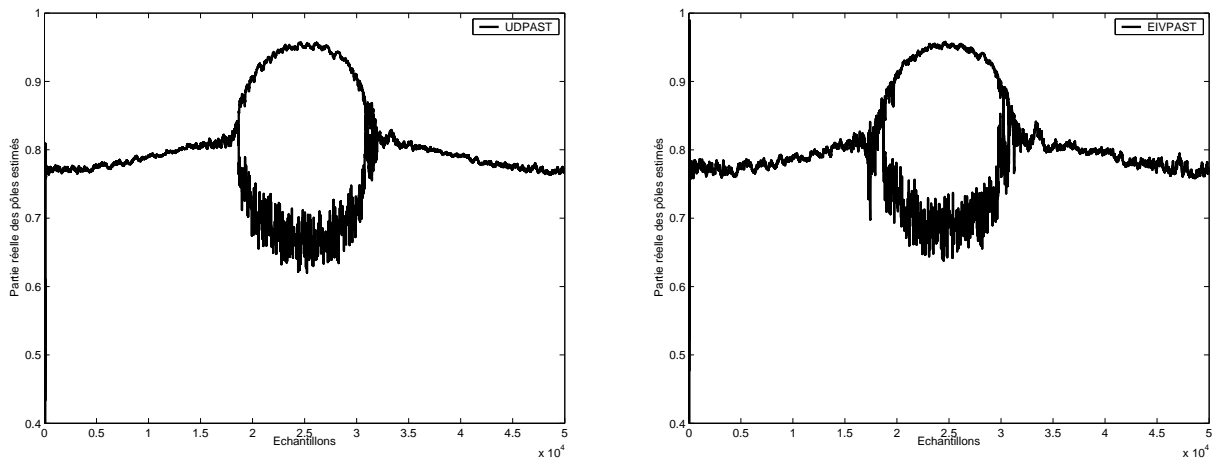


FIG. 6.47: Parties réelles des pôles estimés par les algorithmes EIVPAST et UDPAST. $F_{sin} = 0.1Hz$.

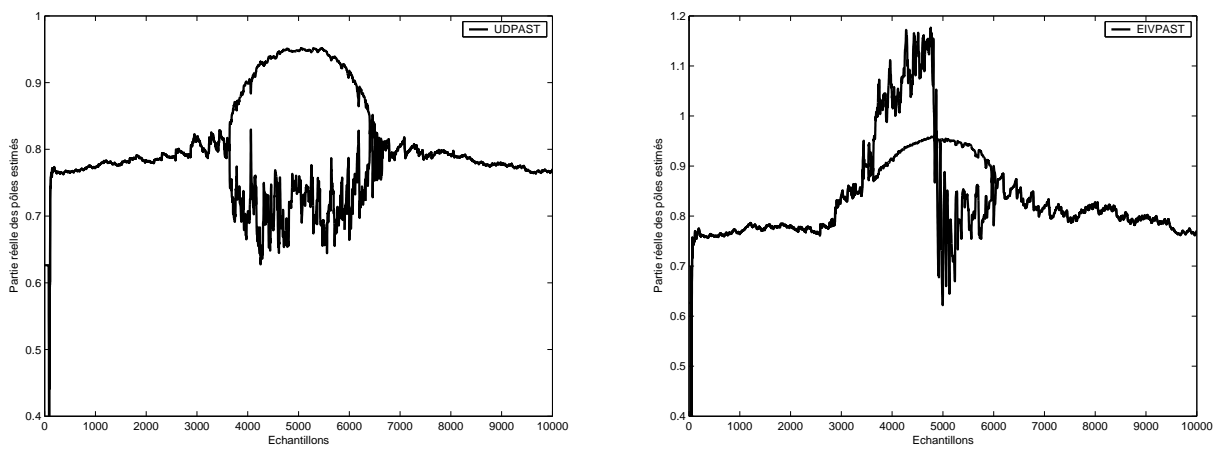


FIG. 6.48: Parties réelles des pôles estimés par les algorithmes EIVPAST et UDPAST. $F_{sin} = 0.5Hz$.

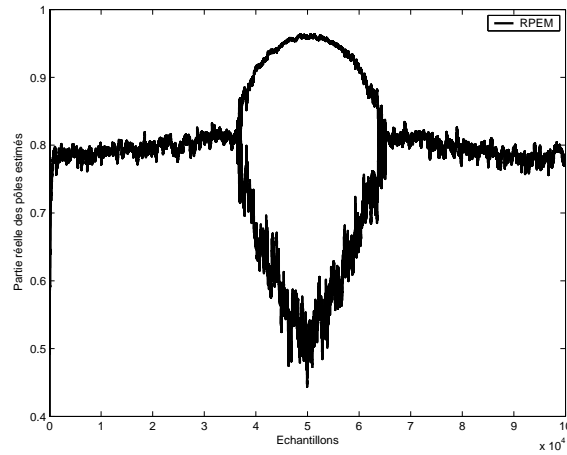


FIG. 6.49: Parties réelles des pôles estimés par l'algorithme RPEM. $F_{sin} = 0.05Hz$.

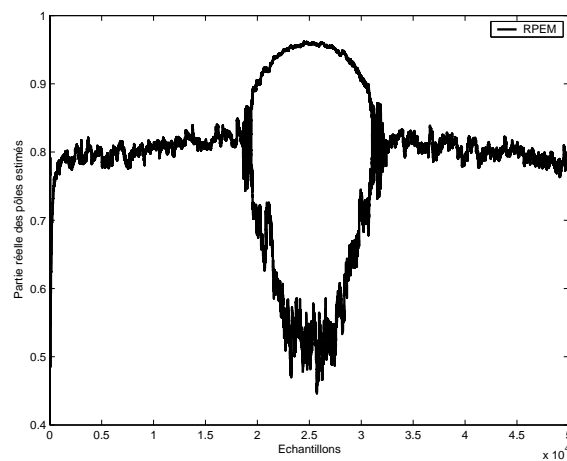


FIG. 6.50: Parties réelles des pôles estimés par l'algorithme RPEM. $F_{sin} = 0.1Hz$.

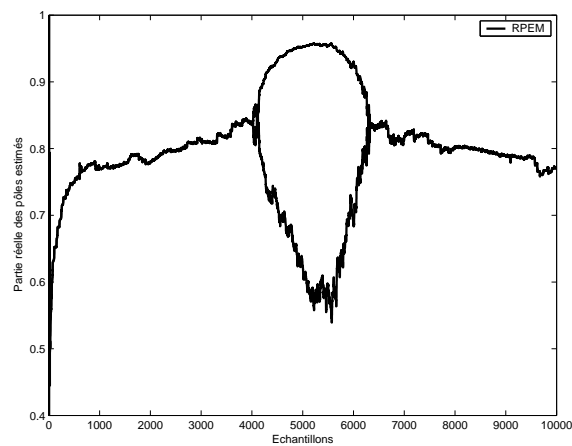


FIG. 6.51: Parties réelles des pôles estimés par l'algorithme RPEM. $F_{sin} = 0.5Hz$.

6.3 Conclusion

Ce chapitre avait pour objectif de mettre en évidence les capacités ainsi que les limitations des algorithmes PM lorsque ces derniers sont appliqués à des données réelles et simulées. Bien que non exhaustive, cette étude a permis de faire ressortir certaines de leurs propriétés, à savoir qu'elles :

- fournissent des modèles aussi fiables que les techniques employant PAST ;
- présentent une vitesse de convergence légèrement plus rapide que leurs homologues utilisant des versions approximées du critère de Yang pour un coût numérique plus faible ;
- sont capables de produire des modèles stables ayant un comportement proche de celui du système à identifier lors d'une faible sous-estimation ou surestimation de l'ordre ;
- ne demandent pas de choisir des indices de construction f et p trop élevés ;
- qu'elles sont capables, la plupart du temps, de suivre des évolutions lentes de procédés lorsque le facteur d'oubli est correctement choisi.

L'analyse réalisée sur des données réelles a également montré que ces méthodes conduisaient à des modèles dont les performances sont similaires à celles d'algorithmes hors ligne tel que PI MOESP. Elle a également dévoilé certaines limites communes aux méthodes d'identification récursive des sous-espaces.

Chapitre 7

Conclusion et perspectives

Le développement de nouvelles méthodes d'identification récursive des systèmes exploitant l'approche des sous-espaces constitue la principale motivation de cette thèse. La démarche suivie pour justifier ce travail a débuté par une synthèse bibliographique des principaux algorithmes récursifs classiques fondés sur les moindres carrés. Cette analyse a permis de faire ressortir les limites de cette classe d'identification. Elle a plus précisément montré que celles-ci sont liées à la structure même des modèles employés par cette approche classique : la fonction de transfert. La forme d'état s'est avérée, dans de nombreux cas, être une excellente alternative à la représentation polynomiale. L'éclosion, au début des années 90, des techniques d'identification des sous-espaces a permis d'exploiter au mieux l'approche d'état en identification hors ligne. Notre étude s'est plus particulièrement intéressée aux algorithmes regroupés sous l'acronyme MOESP puisque ces derniers sont à la base des méthodes récursives des sous-espaces. Ces techniques ont la particularité d'extraire les matrices d'état du système à partir de la matrice d'observabilité estimée, sans rechercher le vecteur d'état. Comme tout algorithme hors ligne des sous-espaces, elles visent à identifier un modèle d'état du système, sans calcul intermédiaire d'une représentation externe, en ajustant un ou plusieurs sous-espaces aux mesures acquises sur le procédé. Elles utilisent principalement des outils d'algèbre linéaire tels que la factorisation QR et la décomposition en valeurs propres ou singulières. Ces méthodes souffrent malencontreusement d'un inconvénient majeur pour leur implémentation en temps réel : leur coût calculatoire. La décomposition en valeurs propres ou singulières constitue l'étape la plus onéreuse numériquement. Cette constatation a conduit les chercheurs à développer des algorithmes alternatifs à cette technique de séparation. Notre étude bibliographique a mis en évidence que des solutions intéressantes étaient proposées en traitement du signal pour la détection des directions d'arrivée. Le recensement des méthodes d'identification récursive utilisant l'approche sous-espace et proposées jusqu'en 2003 a montré que ces dernières étaient toutes basées sur un critère particulier employé en traitement d'antennes : le critère de Yang. Plusieurs techniques ont ainsi été développées :

- les algorithmes de T. Gustafsson qui utilisent une version approximée du critère de Yang nommée PAST associée à une variable instrumentale afin d'être applicable en présence de bruits colorés ;
- ceux proposés par M. Lovera qui permettent :
 - * d'estimer récursivement le vecteur de sortie modifié en mettant à jour les décomposition QR des schémas d'identification Ordinary et PI/PO MOESP,
 - * d'introduire le concept du projecteur en identification en proposant une version qua-

- dratique avec contraintes du critère de Yang ;
- les travaux de H. Oku qui reposent sur la mise à jour, à l'aide du lemme d'inversion matricielle, de matrices symétriques dérivées des schémas d'identification Ordinary et PI/PO MOESP et l'application d'un algorithme de descente itératif de type gradient pour estimer une base de la matrice d'observabilité du système.

Leur analyse a permis de cerner un certain nombre d'inconvénients :

- Ces méthodes reposent toutes sur un seul et même critère d'ordre 4 ;
- De par la complexité de cette fonction coût, il est nécessaire :
 - * d'utiliser des techniques de descente itératives de type gradient pour le minimiser,
 - * d'introduire des approximations afin d'obtenir des critères quadratiques et faciliter sa minimisation,
 - * de construire des fonctions coûts d'ordre 2 avec contraintes demandant une très bonne initialisation.

L'exploitation de méthodes analysées lors de notre étude des techniques de poursuite des directions d'arrivée nous a alors permis de proposer une nouvelle approche n'utilisant plus le critère d'ordre 4 employé jusqu'ici. Celle-ci est plus précisément basée sur l'adaptation d'une technique particulière de traitement d'antennes jusqu'à présent inappliquée en identification : la méthode du propagateur.

Les principales contributions de cette thèse, hormis l'état de l'art des méthodes récursives des sous-espaces, sont :

- *Une présentation de l'analogie entre l'identification récursive des sous-espaces et la poursuite des directions d'arrivée en traitement d'antennes*

Cette description permet à la fois de clarifier la formulation du problème d'identification récursive par l'approche des sous-espaces et de mettre en évidence les deux étapes nécessaires à toute procédure adaptative des sous-espaces :

1. la mise à jour, à chaque nouvelle acquisition, du vecteur d'observation ;
2. l'estimation récursive d'une base de l'espace d'observabilité à partir de ce vecteur en adaptant des méthodes de traitement d'antennes.

Cette décomposition a également permis de prouver qu'il est possible de traiter le problème des perturbations dans l'une ou l'autre de ces étapes et ce, de manière totalement équivalente.

- *Le développement de deux nouveaux algorithmes d'estimation du vecteur d'observation*

Ces méthodes reposent sur le lemme d'inversion matricielle et sont alternatives aux méthodes de mise à jour des décompositions QR des schémas Ordinary et PI/PO MOESP proposées par M. Lovera. Elles s'inspirent de la procédure développée par H. Oku mais n'utilisent aucune matrice symétrique. Elles confondent a posteriori les affirmations de H. Oku selon lequel l'utilisation de matrices non symétriques ne serait pas appropriée à la récursion. Une analyse de complexité des algorithmes de mise à jour du vecteur d'observation a été réalisée. Cette dernière montre que les algorithmes d'estimation fondés sur le lemme d'inversion matricielle sont moins coûteux numériquement que leurs homologues utilisant la décomposition QR. Il est cependant important de noter que l'application de techniques employant le lemme d'inversion matricielle peut, dans certaines situations, conduire à une instabilité avérée.

- *L'adaptation de la méthode du propagateur à l'identification récursive des sous-espaces*
Par définition, le propagateur est un opérateur qui permet de séparer le sous-espace signal en deux sous-espaces orthogonaux à l'aide d'opérations linéaires. Ces propriétés nous ont permis de créer de nouveaux critères d'estimation quadratique, sans approximation et sans contrainte dont la minimisation fournit le sous-espace d'observabilité dans une base particulière. Deux hypothèses sont nécessaires à leur création :

- * que le système à identifier soit observable,
- * que la dimension du vecteur d'état soit connue *a priori*.

Sept algorithmes ont été développés : RPM1, RPM2, COPM, COIVPM, IVPM, EIVPM et EIV_{sqr}tPM. Ces derniers peuvent être regroupés en trois familles :

Famille 1 : algorithme RPM1. Cette technique, dont l'objectif est purement didactique, conduit à des estimées consistantes si et seulement si les perturbations agissant sur le procédé sont nulles.

Famille 2 : algorithmes RPM2 et COPM. Ces méthodes utilisent la matrice de covariance du vecteur d'observation estimée par les versions récursives de PI/PO MOESP qui sont, par construction, asymptotiquement épurées des perturbations. Ils diffèrent par l'approche employée pour les minimiser (resp. non séquentielle et séquentielle).

Famille 3 : algorithmes COIVPM, IVPM, EIVPM et EIV_{sqr}tPM. Le vecteur d'observation utilisé est estimé quant à lui à l'aide des versions récursives d'Ordinary MOESP. Puisque ce dernier est asymptotiquement biaisé lorsque, par exemple, le bruit de sortie est coloré, une variable instrumentale lui est associée afin d'aboutir à une estimée consistante du propagateur lorsque le système est sujet aux perturbations. Les algorithmes IVPM et EIVPM s'inspirent de techniques de minimisation classique telles que les moindres carrés récursifs. Ils diffèrent selon les hypothèses fixées sur le nombre d'instruments composant la variable instrumentale. IVPM n'est applicable qu'en construisant une variable instrumentale de dimension égale au nombre de variables internes. EIVPM permet quant à lui de gérer un nombre d'instruments plus conséquent. L'algorithme COIVPM peut être employé quelle que soit la taille de la variable instrumentale utilisée. Il diffère par l'approche adoptée. Le propagateur est en effet estimé à l'aide d'une minimisation séquentielle de plusieurs critères. L'algorithme EIV_{sqr}tPM est une version robuste d'EIVPM. Contrairement à ce dernier, EIV_{sqr}tPM permet de garantir la positivité de certaines matrices composant la méthode la variable instrumentale étendue.

Plusieurs résultats intéressants ont été mis en évidence lors de l'étude théorique et pratique des techniques d'identification récursive basées sur le propagateur. Une première analyse des propriétés asymptotiques des algorithmes PM a montré que :

- la méthode du propagateur sans variable instrumentale utilisant un vecteur d'observation bruité produit des estimées asymptotiquement biaisées même si les perturbations sont temporellement et spatialement blanches ;
- les algorithmes du propagateur avec variable instrumentale conduisent à des estimées non biaisées de l'espace d'observabilité si et seulement si :
 - * le système à identifier est causal, asymptotiquement stable et observable,
 - * l'entrée appliquée au procédé vérifie la condition d'excitation persistante,

- * la variable instrumentale et le bruit sont décorrélés ;
- les algorithmes IVPM et IVPT ont, tout au moins sur un exemple de simulation, des propriétés statistiques équivalentes au premier et second ordre. Bien que nous n'ayons pas encore réussi à le démontrer théoriquement, cette constatation indique qu'il est fort probable que nous puissions obtenir des résultats équivalents avec un critère quadratique sans approximation et sans contrainte et des fonctions coût d'ordre 4 ou d'ordre 2 avec contraintes.

Une étude de complexité des méthodes récursives des sous-espaces a également permis de prouver que les techniques fondées sur le propagateur sont numériquement moins onéreuses que leurs homologues utilisant l'approximation PAST.

L'application des algorithmes PM sur des données de simulation et des données réelles stationnaires et non stationnaires a mis en évidence :

- qu'ils présentent une tolérance intéressante à une faible sur-estimation et sous-estimation de l'ordre du système ;
- que le facteur d'oubli fait l'objet d'un compromis capacité de poursuite des évolutions du système / sensibilité aux perturbations ;
- que les paramètres de fenêtrage f et p ont une influence négligeable sur la fiabilité des estimées et qu'il est préférable de les choisir relativement faibles pour éviter d'augmenter le coût calculatoire ;
- qu'ils sont capables de suivre des évolutions lentes de paramètres ;
- qu'ils sont aussi performants en terme de consistance que les techniques utilisant le critère de Yang et ses dérivées.

Ces différents résultats indiquent donc que les méthodes récursives des sous-espaces utilisant le propagateur sont relativement efficaces et particulièrement prometteuses.

Perspectives de recherche

Comme nous l'avons précisé au sein des paragraphes précédents, un certain nombre de développements doivent encore faire l'objet d'une recherche approfondie. Deux axes nous semblent plus particulièrement intéressants à étudier :

- les propriétés asymptotiques des méthodes PM avec variable instrumentale ;
- le problème de l'estimation de l'ordre du système en ligne.

Ce dernier point est certainement l'inconvénient majeur des algorithmes récursifs des sous-espaces. Les pistes de recherche actuelles s'orientent plus précisément vers l'utilisation des méthodes COPASTd et COIVPASTd associées à des critères d'information tels que AIC ou SVC. Ces techniques d'identification pourraient en effet fournir des estimées des valeurs propres de la matrice de covariance du vecteur d'observation. Les études réalisées jusqu'ici n'ont cependant pas permis d'aboutir à des résultats publiables.

L'application de ces méthodes d'identification sur des procédés industriels est également une orientation qu'il nous semble nécessaire de considérer. Dans le contexte du suivi et du pilotage de systèmes industriels, il est souvent très utile de posséder des outils récursifs capables de modéliser, en temps réel, des processus non stationnaires. Les algorithmes des sous-espaces développés dans ce manuscrit pourraient, par exemple, être utilisés pour obtenir des modèles locaux de procédés évolutifs utilisés ensuite par des techniques de classification de données pour

surveiller et diagnostiquer le système étudié. L'identification est également une étape préalable nécessaire à de nombreux problèmes de commande. Les méthodes d'estimation développées au sein de ce rapport présentent le double avantage d'être récursives et d'utiliser une représentation très employée en commande de procédé : la forme d'état. Une étude de l'apport de ces algorithmes en commande adaptative ou optimale pourrait donc également être envisagée.

Annexe A

Preuves

Sommaire

| | | |
|-----|--------------------------------|-----|
| A.1 | Preuve de la formule (3.133) | 187 |
| A.2 | Preuve de la formule (3.154) | 188 |
| A.3 | Preuve de la formule (4.96) | 189 |
| A.4 | Preuve de la proposition 5.1 | 191 |
| A.5 | Preuve de l'algorithme (5.81) | 198 |
| A.6 | Preuve de l'algorithme (5.122) | 200 |
| A.7 | Preuve de l'algorithme (5.138) | 202 |

A.1 Preuve de la formule (3.133)

Nous avons montré dans le paragraphe 3.4 (cf. equ. (3.173)) que :

$$\mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \mathbf{\Xi}^T = \mathbf{\Gamma}_f \mathbf{X}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \mathbf{\Xi}^T. \quad (\text{A.1})$$

Par définition de la projection orthogonale sur le noyau de \mathbf{U}_f^+ , cette relation se réécrit alors comme suit :

$$\mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \mathbf{\Xi}^T = \mathbf{\Gamma}_f \left(\mathbf{X}_f^+ \mathbf{\Xi}^T - \mathbf{U}_f^{+T} \left(\mathbf{U}_f^+ \mathbf{U}_f^{+T} \right)^{-1} \mathbf{U}_f^+ \mathbf{\Xi}^T \right). \quad (\text{A.2})$$

Or :

$$\begin{aligned} \lim_{M \rightarrow \infty} \left(\mathbf{X}_f^+ \mathbf{\Xi}^T - \mathbf{U}_f^{+T} \left(\mathbf{U}_f^+ \mathbf{U}_f^{+T} \right)^{-1} \mathbf{U}_f^+ \mathbf{\Xi}^T \right) = \\ \mathbb{E} \left\{ \mathbf{x}_f^+ \boldsymbol{\xi}^T \right\} - \mathbb{E} \left\{ \mathbf{x}_f^+ \mathbf{u}_f^{+T} \right\} \mathbb{E} \left\{ \left(\mathbf{u}_f^+ \mathbf{u}_f^{+T} \right)^{-1} \right\} \mathbb{E} \left\{ \mathbf{u}_f^+ \boldsymbol{\xi}^T \right\} = \mathbf{R}_{\mathbf{x}\boldsymbol{\xi}} - \mathbf{R}_{\mathbf{x}\mathbf{u}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{\mathbf{u}_f^+ \boldsymbol{\xi}}. \end{aligned} \quad (\text{A.3})$$

Ainsi :

$$\lim_{M \rightarrow \infty} \left(\mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \mathbf{\Xi}^T \right) = \mathbf{\Gamma}_f \left(\mathbf{R}_{\mathbf{x}\boldsymbol{\xi}} - \mathbf{R}_{\mathbf{x}\mathbf{u}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{\mathbf{u}_f^+ \boldsymbol{\xi}} \right). \quad (\text{A.4})$$

La matrice des résidus définie par :

$$\begin{aligned} \mathbf{R}_\varepsilon &= \mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \mathbf{\Xi}^T \left(\mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \mathbf{\Xi}^T \right)^{-1} \mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \left(\mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \mathbf{\Xi}^T \left(\mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \mathbf{\Xi}^T \right)^{-1} \mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \right)^T \\ &= \mathbf{Y}_f^+ \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \mathbf{\Xi}^T \left(\mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \mathbf{\Xi}^T \right)^{-1} \mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \mathbf{Y}_f^+ \end{aligned} \quad (\text{A.5})$$

vérifie alors asymptotiquement :

$$\mathbf{R}_\varepsilon = \mathbf{\Gamma}_f \left(\mathbf{R}_{\mathbf{x}\xi} - \mathbf{R}_{\mathbf{x}\mathbf{u}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{\mathbf{u}_f^+ \xi} \right) \left(\mathbf{\Xi} \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}} \mathbf{\Xi}^T \right)^{-1} \left(\mathbf{R}_{\mathbf{x}\xi} - \mathbf{R}_{\mathbf{x}\mathbf{u}_f^+} \mathbf{R}_{\mathbf{u}_f^+}^{-1} \mathbf{R}_{\mathbf{u}_f^+ \xi} \right)^T \mathbf{\Gamma}_f^T. \quad (\text{A.6})$$

A.2 Preuve de la formule (3.154)

Par définition [GV96] :

$$\mathbf{\Pi}_\Psi = \mathbf{\Psi}^T (\mathbf{\Psi} \mathbf{\Psi}^T)^{-1} \mathbf{\Psi}. \quad (\text{A.7})$$

Puisque :

$$\mathbf{\Psi} = \begin{bmatrix} \mathbf{U}_p^- \\ \mathbf{\Xi} \end{bmatrix}, \quad (\text{A.8})$$

la projection sur $\mathbf{\Psi}$ s'écrit :

$$\mathbf{\Pi}_\Psi = \begin{bmatrix} \mathbf{U}_p^{-T} & \mathbf{\Xi}^T \end{bmatrix} \begin{bmatrix} \mathbf{U}_p^- \mathbf{U}_p^{-T} & \mathbf{U}_p^- \mathbf{\Xi}^T \\ \mathbf{\Xi} \mathbf{U}_p^{-T} & \mathbf{\Xi} \mathbf{\Xi}^T \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{U}_p^- \\ \mathbf{\Xi} \end{bmatrix}. \quad (\text{A.9})$$

En utilisant le lemme d'inversion matricielle par bloc [Bro91] :

Lemme A.1. Soit \mathbf{M} une matrice inversible décomposable de la façon suivante :

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_1 & \mathbf{M}_2 \\ \mathbf{M}_3 & \mathbf{M}_4 \end{bmatrix}. \quad (\text{A.10})$$

L'inverse de la matrice \mathbf{M} vérifie alors :

$$\mathbf{M}^{-1} = \begin{bmatrix} \mathbf{M}_1^{-1} + \mathbf{M}_1^{-1} \mathbf{M}_2 \mathbf{T}^{-1} \mathbf{M}_3 \mathbf{M}_1^{-1} & -\mathbf{M}_1^{-1} \mathbf{M}_2 \mathbf{T}^{-1} \\ -\mathbf{T}^{-1} \mathbf{M}_3 \mathbf{M}_1^{-1} & \mathbf{T}^{-1} \end{bmatrix} \quad (\text{A.11})$$

où $\mathbf{T} = \mathbf{M}_4 - \mathbf{M}_3 \mathbf{M}_1^{-1} \mathbf{M}_2$ est le complément de Schur de \mathbf{M}_1 .

la relation (A.9) devient :

$$\begin{aligned} \mathbf{\Pi}_\Psi &= \\ \begin{bmatrix} \mathbf{U}_p^- \\ \mathbf{\Xi} \end{bmatrix}^T & \begin{bmatrix} (\mathbf{U}_p^- \mathbf{U}_p^{-T})^{-1} - (\mathbf{U}_p^- \mathbf{U}_p^{-T})^{-1} \mathbf{U}_p^- \mathbf{\Xi}^T \mathbf{T}^{-1} \mathbf{\Xi} \mathbf{U}_p^- (\mathbf{U}_p^- \mathbf{U}_p^{-T})^{-1} & -(\mathbf{U}_p^- \mathbf{U}_p^{-T})^{-1} \mathbf{U}_p^- \mathbf{\Xi}^T \mathbf{T}^{-1} \\ -\mathbf{T}^{-1} \mathbf{\Xi} \mathbf{U}_p^- (\mathbf{U}_p^- \mathbf{U}_p^{-T})^{-1} & \mathbf{T}^{-1} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{U}_p^- \\ \mathbf{\Xi} \end{bmatrix} \end{aligned} \quad (\text{A.12})$$

avec :

$$\mathbf{T}^{-1} = \boldsymbol{\Xi}\boldsymbol{\Xi}^T - \boldsymbol{\Xi}\mathbf{U}_p^{-T} \left(\mathbf{U}_p^- \mathbf{U}_p^{-T} \right)^{-1} \mathbf{U}_p^- \boldsymbol{\Xi}^T = \boldsymbol{\Xi}\boldsymbol{\Pi}_{\mathbf{U}_p^-} \boldsymbol{\Xi}^T. \quad (\text{A.13})$$

Ainsi :

$$\begin{aligned} \boldsymbol{\Pi}_\Psi &= \mathbf{U}_p^{-T} \left(\left(\mathbf{U}_p^- \mathbf{U}_p^{-T} \right)^{-1} - \left(\mathbf{U}_p^- \mathbf{U}_p^{-T} \right)^{-1} \mathbf{U}_p^- \boldsymbol{\Xi}^T \mathbf{T}^{-1} \boldsymbol{\Xi} \mathbf{U}_p^- \left(\mathbf{U}_p^- \mathbf{U}_p^{-T} \right)^{-1} \right) \mathbf{U}_p^- \\ &\quad - \mathbf{U}_p^{-T} \left(\mathbf{U}_p^- \mathbf{U}_p^{-T} \right)^{-1} \mathbf{U}_p^- \boldsymbol{\Xi}^T \mathbf{T}^{-1} \boldsymbol{\Xi} - \boldsymbol{\Xi}^T \left(\mathbf{U}_p^- \mathbf{U}_p^{-T} \right)^{-1} \mathbf{U}_p^- \boldsymbol{\Xi}^T \mathbf{T}^{-1} \mathbf{U}_p^- + \boldsymbol{\Xi}^T \mathbf{T}^{-1} \boldsymbol{\Xi} \\ &= \boldsymbol{\Pi}_{\mathbf{U}_p^-} - \boldsymbol{\Pi}_{\mathbf{U}_p^-} \boldsymbol{\Xi}^T \mathbf{T}^{-1} \boldsymbol{\Xi} \boldsymbol{\Pi}_{\mathbf{U}_p^-} - \boldsymbol{\Pi}_{\mathbf{U}_p^-} \boldsymbol{\Xi}^T \mathbf{T}^{-1} \boldsymbol{\Xi} - \boldsymbol{\Xi}^T \mathbf{T}^{-1} \boldsymbol{\Xi} \boldsymbol{\Pi}_{\mathbf{U}_p^-} + \boldsymbol{\Xi}^T \mathbf{T}^{-1} \boldsymbol{\Xi} \\ &= \boldsymbol{\Pi}_{\mathbf{U}_p^-} + \underbrace{\left(\mathbf{I}_{n_{uf}} - \boldsymbol{\Pi}_{\mathbf{U}_p^-} \right)}_{\boldsymbol{\Pi}_{\mathbf{U}_p^-}^\perp} \boldsymbol{\Xi}^T \mathbf{T}^{-1} \boldsymbol{\Xi} \underbrace{\left(\mathbf{I}_{n_{uf}} - \boldsymbol{\Pi}_{\mathbf{U}_p^-} \right)}_{\boldsymbol{\Pi}_{\mathbf{U}_p^-}^\perp} \\ &= \boldsymbol{\Pi}_{\mathbf{U}_p^-} + \boldsymbol{\Pi}_{\mathbf{U}_p^-}^\perp \boldsymbol{\Xi}^T \left(\boldsymbol{\Xi} \boldsymbol{\Pi}_{\mathbf{U}_p^-} \boldsymbol{\Xi}^T \right)^{-1} \boldsymbol{\Xi} \boldsymbol{\Pi}_{\mathbf{U}_p^-}^\perp. \end{aligned} \quad (\text{A.14})$$

A.3 Preuve de la formule (4.96)

Considérons la décomposition QR suivante accessible à l'instant $\bar{t} - 1 = t + M - 2$:

$$\begin{bmatrix} \mathbf{U}_f^+(\bar{t} - 1) \\ \mathbf{Y}_f^+(\bar{t} - 1) \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}}_{11}(\bar{t} - 1) & \mathbf{0} \\ \bar{\mathbf{R}}_{21}(\bar{t} - 1) & \bar{\mathbf{R}}_{22}(\bar{t} - 1) \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1(\bar{t} - 1) \\ \bar{\mathbf{Q}}_2(\bar{t} - 1) \end{bmatrix}. \quad (\text{A.15})$$

Lorsqu'un nouveau couple de données est mesuré, la décomposition précédente est mise à jour de la façon suivante :

$$\begin{bmatrix} \sqrt{\lambda} \begin{bmatrix} \bar{\mathbf{R}}_{11}(\bar{t} - 1) & \mathbf{0} \\ \bar{\mathbf{R}}_{21}(\bar{t} - 1) & \bar{\mathbf{R}}_{22}(\bar{t} - 1) \end{bmatrix} \begin{bmatrix} \mathbf{u}_f^+(\bar{t}) \\ \mathbf{y}_f^+(\bar{t}) \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1(\bar{t} - 1) & \mathbf{0} \\ \bar{\mathbf{Q}}_2(\bar{t} - 1) & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \end{bmatrix}. \quad (\text{A.16})$$

La séquence de rotations de Givens [GV96] permettant d'annihiler $\mathbf{u}_f^+(\bar{t})$ s'écrit explicitement comme suit :

$$\mathbf{Rot}_G(\bar{t}) = \begin{bmatrix} \mathbf{Rot}_G^{11}(\bar{t}) & \mathbf{0} & \mathbf{Rot}_G^{12}(\bar{t}) \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{Rot}_G^{21}(\bar{t}) & \mathbf{0} & \mathbf{Rot}_G^{22}(\bar{t}) \end{bmatrix}. \quad (\text{A.17})$$

La mise à jour de la factorisation QR (A.16) s'écrit alors :

$$\begin{aligned} &\begin{bmatrix} \sqrt{\lambda} \begin{bmatrix} \bar{\mathbf{R}}_{11}(\bar{t} - 1) & \mathbf{0} \\ \bar{\mathbf{R}}_{21}(\bar{t} - 1) & \bar{\mathbf{R}}_{22}(\bar{t} - 1) \end{bmatrix} \begin{bmatrix} \mathbf{u}_f^+(\bar{t}) \\ \mathbf{y}_f^+(\bar{t}) \end{bmatrix} \begin{bmatrix} \bar{\mathbf{Q}}_1(\bar{t} - 1) & \mathbf{0} \\ \bar{\mathbf{Q}}_2(\bar{t} - 1) & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \end{bmatrix} \mathbf{Rot}_G(\bar{t}) = \\ &\begin{bmatrix} \sqrt{\lambda} \bar{\mathbf{R}}_{11}(\bar{t} - 1) \mathbf{Rot}_G^{11}(\bar{t}) + \mathbf{u}_f^+(\bar{t}) \mathbf{Rot}_G^{21}(\bar{t}) & \mathbf{0} & \sqrt{\lambda} \bar{\mathbf{R}}_{11}(\bar{t} - 1) \mathbf{Rot}_G^{12}(\bar{t}) + \mathbf{u}_f^+(\bar{t}) \mathbf{Rot}_G^{22}(\bar{t}) \\ \sqrt{\lambda} \bar{\mathbf{R}}_{21}(\bar{t} - 1) \mathbf{Rot}_G^{11}(\bar{t}) + \mathbf{y}_f^+(\bar{t}) \mathbf{Rot}_G^{21}(\bar{t}) & \bar{\mathbf{R}}_{22}(\bar{t} - 1) & \sqrt{\lambda} \bar{\mathbf{R}}_{21}(\bar{t} - 1) \mathbf{Rot}_G^{12}(\bar{t}) + \mathbf{y}_f^+(\bar{t}) \mathbf{Rot}_G^{22}(\bar{t}) \end{bmatrix}. \end{aligned} \quad (\text{A.18})$$

Les relations suivantes sont alors vérifiées :

$$\mathbf{0} = \sqrt{\lambda} \bar{\mathbf{R}}_{11}(\bar{t} - 1) \mathbf{Rot}_G^{12}(\bar{t}) + \mathbf{u}_f^+(\bar{t}) \mathbf{Rot}_G^{22}(\bar{t}) \quad (\text{A.19})$$

$$\bar{\mathbf{z}}_f^+(\bar{t}) = \sqrt{\lambda} \bar{\mathbf{R}}_{21}(\bar{t} - 1) \mathbf{Rot}_G^{12}(\bar{t}) + \mathbf{y}_f^+(\bar{t}) \mathbf{Rot}_G^{22}(\bar{t}). \quad (\text{A.20})$$

Nous savons de plus que :

$$\mathbf{y}_f^+(\bar{t}) = \mathbf{\Gamma}_f \mathbf{x}(\bar{t}) + \mathbf{H}_f \mathbf{u}_f^+(\bar{t}) + \mathbf{n}_f^+(\bar{t}). \quad (\text{A.21})$$

Ainsi :

$$\begin{aligned} \bar{\mathbf{z}}_f^+(\bar{t}) &= \sqrt{\lambda} \bar{\mathbf{R}}_{21}(\bar{t}-1) \mathbf{Rot}_G^{12}(\bar{t}) + \mathbf{y}_f^+(\bar{t}) \mathbf{Rot}_G^{12}(\bar{t}) \\ &= \sqrt{\lambda} \bar{\mathbf{R}}_{21}(\bar{t}-1) \mathbf{Rot}_G^{12}(\bar{t}) + \mathbf{\Gamma}_f \mathbf{x}(\bar{t}) \mathbf{Rot}_G^{22}(\bar{t}) + \mathbf{H}_f \mathbf{u}_f^+(\bar{t}) \mathbf{Rot}_G^{22}(\bar{t}) + \mathbf{n}_f^+(\bar{t}) \mathbf{Rot}_G^{22}(\bar{t}). \end{aligned} \quad (\text{A.22})$$

Or, d'après l'équation (A.19) :

$$\mathbf{u}_f^+(\bar{t}) \mathbf{Rot}_G^{22}(\bar{t}) = -\sqrt{\lambda} \bar{\mathbf{R}}_{11}(\bar{t}-1) \mathbf{Rot}_G^{12}(\bar{t}). \quad (\text{A.23})$$

Le vecteur de sortie modifié $\bar{\mathbf{z}}_f^+$ vérifie donc la relation suivante :

$$\begin{aligned} \bar{\mathbf{z}}_f^+(\bar{t}) &= \sqrt{\lambda} \bar{\mathbf{R}}_{21}(\bar{t}-1) \mathbf{Rot}_G^{12}(\bar{t}) + \mathbf{\Gamma}_f \mathbf{x}(\bar{t}) \mathbf{Rot}_G^{22}(\bar{t}) \\ &\quad - \sqrt{\lambda} \mathbf{H}_f \bar{\mathbf{R}}_{11}(\bar{t}-1) \mathbf{Rot}_G^{12}(\bar{t}) + \mathbf{n}_f^+(\bar{t}) \mathbf{Rot}_G^{22}(\bar{t}). \end{aligned} \quad (\text{A.24})$$

La suite de cette démonstration repose sur la définition de la factorisation QR :

$$\mathbf{U}_f^+(\bar{t}-1) = \bar{\mathbf{R}}_{11}(\bar{t}-1) \bar{\mathbf{Q}}_1^T(\bar{t}-1) \quad (\text{A.25})$$

$$\mathbf{Y}_f^+(\bar{t}-1) = \bar{\mathbf{R}}_{21}(\bar{t}-1) \bar{\mathbf{Q}}_1^T(\bar{t}-1) + \bar{\mathbf{R}}_{22}(\bar{t}-1) \bar{\mathbf{Q}}_2^T(\bar{t}-1). \quad (\text{A.26})$$

Ces deux expressions nous permettent en effet d'écrire que :

$$\begin{aligned} \bar{\mathbf{R}}_{21}(\bar{t}-1) &= \mathbf{Y}_f^+(\bar{t}-1) \bar{\mathbf{Q}}_1^T(\bar{t}-1) = \left(\mathbf{\Gamma}_f \mathbf{X}^+(\bar{t}-1) + \mathbf{H}_f \mathbf{U}_f^+(\bar{t}-1) + \mathbf{N}_f^+(\bar{t}-1) \right) \bar{\mathbf{Q}}_1^T(\bar{t}-1) \\ &= \mathbf{\Gamma}_f \mathbf{X}^+(\bar{t}-1) \bar{\mathbf{Q}}_1^T(\bar{t}-1) + \underbrace{\mathbf{H}_f \mathbf{U}_f^+(\bar{t}-1) \bar{\mathbf{Q}}_1^T(\bar{t}-1)}_{\bar{\mathbf{R}}_{11}(\bar{t}-1)} + \mathbf{N}_f^+(\bar{t}-1) \bar{\mathbf{Q}}_1^T(\bar{t}-1) \\ &= \mathbf{\Gamma}_f \mathbf{X}^+(\bar{t}-1) \bar{\mathbf{Q}}_1^T(\bar{t}-1) + \mathbf{H}_f \bar{\mathbf{R}}_{11}(\bar{t}-1) + \mathbf{N}_f^+(\bar{t}-1) \bar{\mathbf{Q}}_1^T(\bar{t}-1). \end{aligned} \quad (\text{A.27})$$

L'insertion de cette relation au sein de l'équation (A.24) aboutit à la formule suivante :

$$\begin{aligned} \bar{\mathbf{z}}_f^+(\bar{t}) &= \sqrt{\lambda} \mathbf{\Gamma}_f \mathbf{X}^+(\bar{t}-1) \bar{\mathbf{Q}}_1^T(\bar{t}-1) \mathbf{Rot}_G^{12}(\bar{t}) + \sqrt{\lambda} \mathbf{H}_f \bar{\mathbf{R}}_{11}(\bar{t}-1) \mathbf{Rot}_G^{12}(\bar{t}) \\ &\quad + \sqrt{\lambda} \mathbf{N}_f^+(\bar{t}-1) \bar{\mathbf{Q}}_1^T(\bar{t}-1) \mathbf{Rot}_G^{12}(\bar{t}) + \mathbf{\Gamma}_f \mathbf{x}(\bar{t}) \mathbf{Rot}_G^{22}(\bar{t}) \\ &\quad - \sqrt{\lambda} \mathbf{H}_f \bar{\mathbf{R}}_{11}(\bar{t}-1) \mathbf{Rot}_G^{12}(\bar{t}) + \mathbf{n}_f^+(\bar{t}) \mathbf{Rot}_G^{22}(\bar{t}) \\ &= \mathbf{\Gamma}_f \left(\sqrt{\lambda} \mathbf{X}^+(\bar{t}-1) \bar{\mathbf{Q}}_1^T(\bar{t}-1) \mathbf{Rot}_G^{12}(\bar{t}) + \mathbf{x}(\bar{t}-1) \mathbf{Rot}_G^{22}(\bar{t}) \right) \\ &\quad + \sqrt{\lambda} \mathbf{N}_f^+(\bar{t}-1) \bar{\mathbf{Q}}_1^T(\bar{t}-1) \mathbf{Rot}_G^{12}(\bar{t}) + \mathbf{n}_f^+(\bar{t}) \mathbf{Rot}_G^{22}(\bar{t}). \end{aligned} \quad (\text{A.28})$$

A.4 Preuve de la proposition 5.1

L'objectif est de mettre à jour terme à terme la projection :

$$\mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1)\left(\mathbf{\Xi}(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1)\right)^{-1}\mathbf{\Xi}(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} \quad (\text{A.29})$$

lorsque de nouvelles données d'entrée-sortie sont accessibles à la mesure afin d'estimer le vecteur d'observation à l'instant \bar{t} . Pour cela, calculons successivement chaque terme de la relation :

$$\mathbf{Y}_f^+(\bar{t})\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})}\mathbf{\Xi}^T(\bar{t})\left(\mathbf{\Xi}(\bar{t})\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})}\mathbf{\Xi}^T(\bar{t})\right)^{-1}\mathbf{\Xi}(\bar{t})\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})} \quad (\text{A.30})$$

sachant que :

$$\mathbf{Y}_f^+(\bar{t}) = \begin{bmatrix} \sqrt{\lambda}\mathbf{Y}_f^+(\bar{t}-1) & \mathbf{y}_f^+(\bar{t}) \end{bmatrix} \quad (\text{A.31})$$

$$\mathbf{U}_f^+(\bar{t}) = \begin{bmatrix} \sqrt{\lambda}\mathbf{U}_f^+(\bar{t}-1) & \mathbf{u}_f^+(\bar{t}) \end{bmatrix} \quad (\text{A.32})$$

$$\mathbf{\Xi}(\bar{t}) = \begin{bmatrix} \sqrt{\lambda}\mathbf{\Xi}(\bar{t}-1) & \boldsymbol{\xi}(\bar{t}) \end{bmatrix}. \quad (\text{A.33})$$

Nous savons que (cf. equ. (5.35)) :

$$\begin{aligned} \mathbf{Y}_f^+(\bar{t})\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})} &= \\ & \left[\sqrt{\lambda}\mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} - \frac{1}{\sqrt{\lambda+\boldsymbol{\delta}_f(\bar{t})}}\check{\mathbf{z}}_f^+(\bar{t})\boldsymbol{\gamma}_f^{+T}(\bar{t})\mathbf{U}_f^+(\bar{t}-1) \quad \sqrt{\frac{\lambda}{\lambda+\boldsymbol{\delta}_f(\bar{t})}}\check{\mathbf{z}}_f^+(\bar{t}) \right] \end{aligned} \quad (\text{A.34})$$

avec :

$$\boldsymbol{\gamma}_f^+(\bar{t}) = \left(\mathbf{U}_f^+(\bar{t}-1)\mathbf{U}_f^{+T}(\bar{t}-1)\right)^{-1}\mathbf{u}_f^+(\bar{t}) \quad (\text{A.35})$$

$$\boldsymbol{\delta}_f(\bar{t}) = \mathbf{u}_f^{+T}(\bar{t})\boldsymbol{\gamma}_f^+(\bar{t}) \quad (\text{A.36})$$

$$\check{\mathbf{z}}_f^+(\bar{t}) = \sqrt{\frac{\lambda}{\lambda+\boldsymbol{\delta}_f(\bar{t})}}\left(\mathbf{y}_f^+(\bar{t}) - \mathbf{Y}_f^+(\bar{t}-1)\mathbf{U}_f^{+T}(\bar{t}-1)\boldsymbol{\gamma}_f^+(\bar{t})\right) \quad (\text{A.37})$$

En postmultipliant cette expression par $\mathbf{\Xi}^T(\bar{t})$:

$$\begin{aligned} & \mathbf{Y}_f^+(\bar{t})\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})}\mathbf{\Xi}^T(\bar{t}) \\ &= \left[\sqrt{\lambda}\mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} - \frac{1}{\sqrt{\lambda+\boldsymbol{\delta}_f(\bar{t})}}\check{\mathbf{z}}_f^+(\bar{t})\boldsymbol{\gamma}_f^{+T}(\bar{t})\mathbf{U}_f^+(\bar{t}-1) \quad \sqrt{\frac{\lambda}{\lambda+\boldsymbol{\delta}_f(\bar{t})}}\check{\mathbf{z}}_f^+(\bar{t}) \right] \begin{bmatrix} \sqrt{\lambda}\mathbf{\Xi}^T(\bar{t}-1) \\ \boldsymbol{\xi}^T(\bar{t}) \end{bmatrix} \\ &= \lambda\mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1) + \check{\mathbf{z}}_f^+(\bar{t})\check{\boldsymbol{\xi}}^{-T}(\bar{t}) \end{aligned} \quad (\text{A.38})$$

avec :

$$\check{\boldsymbol{\xi}}^{-}(\bar{t}) = \sqrt{\frac{\lambda}{\lambda+\boldsymbol{\delta}_f(\bar{t})}}\left(\boldsymbol{\xi}(\bar{t}) - \mathbf{\Xi}(\bar{t}-1)\mathbf{U}_f^{+T}(\bar{t}-1)\boldsymbol{\gamma}_f^+(\bar{t})\right). \quad (\text{A.39})$$

Intéressons nous maintenant au produit matriciel $\Xi(\bar{t})\Pi_{\mathbf{U}_f^{+\perp}(\bar{t})}$. Par analogie avec l'équation (5.35), en remplaçant $\mathbf{Y}_f^+(\bar{t})$ par $\Xi(\bar{t})$ et $\mathbf{y}_f^+(\bar{t})$ par $\check{\xi}^-(\bar{t})$, toute chose égale par ailleurs, il est facile de montrer que :

$$\begin{aligned} \Xi(\bar{t})\Pi_{\mathbf{U}_f^{+\perp}(\bar{t})} = & \\ & \left[\sqrt{\lambda}\Xi(\bar{t}-1)\Pi_{\mathbf{U}_f^{+\perp}(\bar{t}-1)} - \frac{1}{\sqrt{\lambda+\delta_f(\bar{t})}}\check{\xi}^-(\bar{t})\boldsymbol{\gamma}_f^{+T}(\bar{t})\mathbf{U}_f^+(\bar{t}-1) \quad \sqrt{\frac{\lambda}{\lambda+\delta_f(\bar{t})}}\check{\xi}^-(\bar{t}) \right]. \end{aligned} \quad (\text{A.40})$$

En postmultipliant cette expression par $\Xi^T(\bar{t})$, nous avons :

$$\begin{aligned} & \Xi(\bar{t})\Pi_{\mathbf{U}_f^{+\perp}(\bar{t})}\Xi^T(\bar{t}) \\ &= \left[\sqrt{\lambda}\Xi(\bar{t}-1)\Pi_{\mathbf{U}_f^{+\perp}(\bar{t}-1)} - \frac{1}{\sqrt{\lambda+\delta_f(\bar{t})}}\check{\xi}^-(\bar{t})\boldsymbol{\gamma}_f^{+T}(\bar{t})\mathbf{U}_f^+(\bar{t}-1) \quad \sqrt{\frac{\lambda}{\lambda+\delta_f(\bar{t})}}\check{\xi}^-(\bar{t}) \right] \begin{bmatrix} \sqrt{\lambda}\Xi^T(\bar{t}-1) \\ \check{\xi}^T(\bar{t}) \end{bmatrix} \\ &= \lambda\Xi(\bar{t}-1)\Pi_{\mathbf{U}_f^{+\perp}(\bar{t}-1)}\Xi^T(\bar{t}-1) + \check{\xi}^-(\bar{t})\check{\xi}^{-T}(\bar{t}). \end{aligned} \quad (\text{A.41})$$

En appliquant le lemme d'inversion matricielle à cette dernière équation, nous obtenons :

$$\Phi_f(\bar{t}) = \left(\Xi(\bar{t})\Pi_{\mathbf{U}_f^{+\perp}(\bar{t})}\Xi^T(\bar{t}) \right)^{-1} = \left(\Phi_f(\bar{t}-1) - \frac{\boldsymbol{\chi}_f^+(\bar{t})\boldsymbol{\chi}_f^{+T}(\bar{t})}{\lambda + \zeta_f(\bar{t})} \right) / \lambda \quad (\text{A.42})$$

avec :

$$\zeta_f(\bar{t}) = \check{\xi}^{-T}(\bar{t})\Phi_f(\bar{t}-1)\check{\xi}^-(\bar{t}) \quad (\text{A.43})$$

$$\boldsymbol{\chi}_f^+(\bar{t}) = \Phi_f(\bar{t}-1)\check{\xi}^-(\bar{t}). \quad (\text{A.44})$$

En postmultipliant $\mathbf{Y}_f^+(\bar{t})\Pi_{\mathbf{U}_f^{+\perp}(\bar{t})}\Xi^T(\bar{t})$ par $\Phi_f(\bar{t})$, nous obtenons :

$$\begin{aligned} & \mathbf{Y}_f^+(\bar{t})\Pi_{\mathbf{U}_f^{+\perp}(\bar{t})}\Xi^T(\bar{t})\Phi_f(\bar{t}) \\ &= \left(\lambda\mathbf{Y}_f^+(\bar{t}-1)\Pi_{\mathbf{U}_f^{+\perp}(\bar{t}-1)}\Xi^T(\bar{t}-1) + \check{\mathbf{z}}_f^+(\bar{t})\check{\xi}^{-T}(\bar{t}) \right) \left(\Phi_f(\bar{t}-1) - \frac{\boldsymbol{\chi}_f^+(\bar{t})\boldsymbol{\chi}_f^{+T}(\bar{t})}{\lambda + \zeta_f(\bar{t})} \right) / \lambda \\ &= \mathbf{Y}_f^+(\bar{t}-1)\Pi_{\mathbf{U}_f^{+\perp}(\bar{t}-1)}\Xi^T(\bar{t}-1)\Phi_f(\bar{t}-1) - \mathbf{Y}_f^+(\bar{t}-1)\Pi_{\mathbf{U}_f^{+\perp}(\bar{t}-1)}\Xi^T(\bar{t}-1) \frac{\boldsymbol{\chi}_f^+(\bar{t})\boldsymbol{\chi}_f^{+T}(\bar{t})}{\lambda + \zeta_f(\bar{t})} \\ &+ \frac{1}{\lambda} \left(\check{\mathbf{z}}_f^+(\bar{t})\boldsymbol{\chi}_f^{+T}(\bar{t}) - \check{\mathbf{z}}_f^+(\bar{t}) \frac{\zeta_f(\bar{t})\boldsymbol{\chi}_f^{+T}(\bar{t})}{\lambda + \zeta_f(\bar{t})} \right). \end{aligned} \quad (\text{A.45})$$

En introduisant quelques simplifications, nous obtenons :

$$\begin{aligned}
& \mathbf{Y}_f^+(\bar{t})\mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t})}\mathbf{\Xi}^T(\bar{t})\mathbf{\Phi}_f(\bar{t}) = \mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1)\mathbf{\Phi}_f(\bar{t}-1) \\
& - \mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1)\frac{\boldsymbol{\chi}_f^+(\bar{t})\boldsymbol{\chi}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\zeta}_f(\bar{t})} + \frac{1}{\lambda} \left(\check{\mathbf{z}}_f^+(\bar{t}) \left(1 - \frac{\boldsymbol{\zeta}_f(\bar{t})}{\lambda + \boldsymbol{\zeta}_f(\bar{t})} \right) \boldsymbol{\chi}_f^{+T}(\bar{t}) \right) \\
& = \mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1)\mathbf{\Phi}_f(\bar{t}-1) - \mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1)\frac{\boldsymbol{\chi}_f^+(\bar{t})\boldsymbol{\chi}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\zeta}_f(\bar{t})} \\
& + \frac{\check{\mathbf{z}}_f^+(\bar{t})\boldsymbol{\chi}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\zeta}_f(\bar{t})} = \mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1)\mathbf{\Phi}_f(\bar{t}-1) + \check{\mathbf{z}}_f^+(\bar{t})\frac{\boldsymbol{\chi}_f^{+T}(\bar{t})}{\sqrt{\lambda^2 + \lambda\boldsymbol{\zeta}_f(\bar{t}+1)}}
\end{aligned} \tag{A.46}$$

avec :

$$\check{\mathbf{z}}_f^+(\bar{t}) = \sqrt{\frac{\lambda}{\lambda + \boldsymbol{\zeta}_f(\bar{t})}} \left(\check{\mathbf{z}}_f^+(\bar{t}) - \mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1)\boldsymbol{\chi}_f^+(\bar{t}) \right). \tag{A.47}$$

Enfin :

$$\begin{aligned}
& \mathbf{Y}_f^+(\bar{t})\mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t})}\mathbf{\Xi}^T(\bar{t})\mathbf{\Phi}_f(\bar{t})\mathbf{\Xi}(\bar{t})\mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t})} \\
& = \left(\mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1)\mathbf{\Phi}_f(\bar{t}-1) + \check{\mathbf{z}}_f^+(\bar{t})\frac{\boldsymbol{\chi}_f^{+T}(\bar{t})}{\sqrt{\lambda^2 + \lambda\boldsymbol{\zeta}_f(\bar{t}+1)}} \right) \\
& \left[\sqrt{\lambda}\mathbf{\Xi}(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)} - \frac{1}{\sqrt{\lambda + \boldsymbol{\delta}_f(\bar{t})}}\check{\boldsymbol{\xi}}^-(\bar{t})\boldsymbol{\gamma}_f^{+T}(\bar{t})\mathbf{U}_f^+(\bar{t}-1) \quad \sqrt{\frac{\lambda}{\lambda + \boldsymbol{\delta}_f(\bar{t})}}\check{\boldsymbol{\xi}}^-(\bar{t}) \right].
\end{aligned} \tag{A.48}$$

Afin de faciliter l'écriture finale de ce produit matriciel, calculons séparément l'expression $\left(\mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1)\mathbf{\Phi}_f(\bar{t}-1) + \check{\mathbf{z}}_f^+(\bar{t})\frac{\boldsymbol{\chi}_f^{+T}(\bar{t})}{\sqrt{\lambda^2 + \lambda\boldsymbol{\zeta}_f(\bar{t})}} \right) \check{\boldsymbol{\xi}}^-(\bar{t})$ commune aux deux composantes de la matrice (A.48) :

$$\begin{aligned}
& \left(\mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1)\mathbf{\Phi}_f(\bar{t}-1) + \check{\mathbf{z}}_f^+(\bar{t})\frac{\boldsymbol{\chi}_f^{+T}(\bar{t})}{\sqrt{\lambda^2 + \lambda\boldsymbol{\zeta}_f(\bar{t})}} \right) \check{\boldsymbol{\xi}}^-(\bar{t}) \\
& = \mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1)\boldsymbol{\chi}_f^+(\bar{t}) + \check{\mathbf{z}}_f^+(\bar{t})\frac{\boldsymbol{\zeta}_f(\bar{t})}{\sqrt{\lambda^2 + \lambda\boldsymbol{\zeta}_f(\bar{t})}} + \check{\mathbf{z}}_f^+(\bar{t}) - \check{\mathbf{z}}_f^+(\bar{t}) \\
& = \check{\mathbf{z}}_f^+(\bar{t}) + \check{\mathbf{z}}_f^+(\bar{t})\frac{\boldsymbol{\zeta}_f(\bar{t})}{\sqrt{\lambda^2 + \lambda\boldsymbol{\zeta}_f(\bar{t})}} - \sqrt{\frac{\lambda + \boldsymbol{\zeta}_f(\bar{t})}{\lambda}}\check{\mathbf{z}}_f^+(\bar{t}) = \check{\mathbf{z}}_f^+(\bar{t}) - \sqrt{\frac{\lambda}{\lambda + \boldsymbol{\zeta}_f(\bar{t})}}\check{\mathbf{z}}_f^+(\bar{t}),
\end{aligned} \tag{A.49}$$

expression qu'on nommera désormais $\check{\mathbf{z}}_f^+(\bar{t})$. Les mf premières colonnes de l'équation (A.48) valent alors :

$$\begin{aligned} & \left[\mathbf{Y}_f^+(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t})} \mathbf{\Xi}^T(\bar{t}) \mathbf{\Phi}_f(\bar{t}) \mathbf{\Xi}(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t})} \right]_{(:,1:n_{uf})} \\ &= \sqrt{\lambda} \mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) \mathbf{\Phi}_f(\bar{t}-1) \mathbf{\Xi}(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)} \\ &+ \check{\mathbf{z}}_f^+(\bar{t}) \frac{\boldsymbol{\chi}_f^{+T}(\bar{t})}{\sqrt{\lambda + \boldsymbol{\zeta}_f(\bar{t})}} \mathbf{\Xi}(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)} - \frac{1}{\sqrt{\lambda + \boldsymbol{\delta}_f(\bar{t})}} \check{\mathbf{z}}_f^+(\bar{t}) \boldsymbol{\gamma}_f^{+T}(\bar{t}) \mathbf{U}_f^+(\bar{t}-1), \end{aligned} \quad (\text{A.50})$$

la dernière colonne ayant pour expression :

$$\left[\mathbf{Y}_f^+(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t})} \mathbf{\Xi}^T(\bar{t}) \mathbf{\Phi}_f(\bar{t}) \mathbf{\Xi}(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t})} \right]_{(:,n_{uf}+1)} = \sqrt{\frac{\lambda}{\lambda + \boldsymbol{\delta}_f(\bar{t})}} \check{\mathbf{z}}_f^+(\bar{t}). \quad (\text{A.51})$$

Nous aboutissons ainsi à l'algorithme de mise à jour suivant :

$$\boldsymbol{\gamma}_f^+(\bar{t}) = \left(\mathbf{U}_f^+(\bar{t}-1) \mathbf{U}_f^{+T}(\bar{t}-1) \right)^{-1} \mathbf{u}_f^+(\bar{t}) = \quad (\text{A.52a})$$

$$\boldsymbol{\delta}_f(\bar{t}) = \mathbf{u}_f^{+T}(\bar{t}) \boldsymbol{\gamma}_f^+(\bar{t}) \quad (\text{A.52b})$$

$$\check{\mathbf{z}}_f^+(\bar{t}) = \sqrt{\frac{\lambda}{\lambda + \boldsymbol{\delta}_f(\bar{t})}} \left(\mathbf{y}_f^+(\bar{t}) - \mathbf{Y}_f^+(\bar{t}-1) \mathbf{U}_f^{+T}(\bar{t}-1) \boldsymbol{\gamma}_f^+(\bar{t}) \right) \quad (\text{A.52c})$$

$$\mathbf{Y}_f^+(\bar{t}) \mathbf{U}_f^{+T}(\bar{t}) = \lambda \mathbf{Y}_f^+(\bar{t}-1) \mathbf{U}_f^{+T}(\bar{t}-1) + \mathbf{y}_f^+(\bar{t}) \mathbf{u}_f^{+T}(\bar{t}) \quad (\text{A.52d})$$

$$\left(\mathbf{U}_f^+(\bar{t}) \mathbf{U}_f^{+T}(\bar{t}) \right)^{-1} = \left(\left(\mathbf{U}_f^+(\bar{t}-1) \mathbf{U}_f^{+T}(\bar{t}-1) \right)^{-1} - \frac{\boldsymbol{\gamma}_f^+(\bar{t}) \boldsymbol{\gamma}_f^{+T}(\bar{t}+1)}{\lambda + \boldsymbol{\delta}_f(\bar{t})} \right) / \lambda \quad (\text{A.52e})$$

$$\check{\boldsymbol{\xi}}^-(\bar{t}) = \sqrt{\frac{\lambda}{\lambda + \boldsymbol{\delta}_f(\bar{t})}} \left(\boldsymbol{\xi}(\bar{t}) - \mathbf{\Xi}(\bar{t}-1) \mathbf{U}_f^{+T}(\bar{t}-1) \boldsymbol{\gamma}_f^+(\bar{t}) \right) \quad (\text{A.52f})$$

$$\boldsymbol{\chi}_f^+(\bar{t}) = \mathbf{\Phi}_f(\bar{t}-1) \check{\boldsymbol{\xi}}^-(\bar{t}) \quad (\text{A.52g})$$

$$\boldsymbol{\zeta}_f(\bar{t}) = \boldsymbol{\chi}_f^{+T}(\bar{t}) \check{\boldsymbol{\xi}}^-(\bar{t}) \quad (\text{A.52h})$$

$$\check{\mathbf{z}}_f^+(\bar{t}) = \sqrt{\frac{\lambda}{\lambda + \boldsymbol{\zeta}_f(\bar{t})}} \left(\check{\mathbf{z}}_f^+(\bar{t}) - \mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) \boldsymbol{\chi}_f^+(\bar{t}) \right) \quad (\text{A.52i})$$

$$\check{\mathbf{z}}_f^+(\bar{t}) = \check{\mathbf{z}}_f^+(\bar{t}) - \sqrt{\frac{\lambda}{\lambda + \boldsymbol{\zeta}_f(\bar{t})}} \check{\mathbf{z}}_f^+(\bar{t}) \quad (\text{A.52j})$$

$$\mathbf{\Xi}(\bar{t}) \mathbf{U}_f^{+T}(\bar{t}) = \lambda \mathbf{\Xi}(\bar{t}-1) \mathbf{U}_f^{+T}(\bar{t}-1) + \boldsymbol{\xi}(\bar{t}) \mathbf{u}_f^{+T}(\bar{t}) \quad (\text{A.52k})$$

$$\mathbf{\Phi}_f(\bar{t}) = \left(\mathbf{\Phi}_f(\bar{t}-1) - \frac{\boldsymbol{\chi}_f^+(\bar{t}) \boldsymbol{\chi}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\zeta}_f(\bar{t})} \right) / \lambda \quad (\text{A.52l})$$

$$\mathbf{Y}_f^+(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t})} \mathbf{\Xi}^T(\bar{t}) = \lambda \mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) + \check{\mathbf{z}}_f^+(\bar{t}) \check{\boldsymbol{\xi}}^{-T}(\bar{t}). \quad (\text{A.52m})$$

La dernière étape de cette démonstration consiste à expliciter la relation liant $\check{\mathbf{z}}_f^+$ et $\check{\mathbf{z}}_f^+$ au vecteur d'observation \mathbf{z}_f^+ . Pour cela, il est nécessaire de calculer :

$$\begin{aligned} \mathbf{Z}_f^+(t)\mathbf{Z}_f^{+T}(t) &= \left[\sqrt{\lambda}\mathbf{Z}_f^+(t-1) \quad \mathbf{z}_f^+(t) \right] \left[\sqrt{\lambda}\mathbf{Z}_f^+(t-1) \quad \mathbf{z}_f^+(t) \right]^T \\ &= \lambda\mathbf{Z}_f^+(t-1)\mathbf{Z}_f^{+T}(t-1) + \mathbf{z}_f^+(t)\mathbf{z}_f^{+T}(t) \quad (\text{A.53}) \end{aligned}$$

que nous savons vérifier que :

$$\begin{aligned} &\mathbf{Z}_f^+(t)\mathbf{Z}_f^{+T}(t) \\ &= \mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1) \left(\mathbf{\Xi}(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1) \right)^{-1} \mathbf{\Xi}(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} \\ &\quad \left(\mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1) \left(\mathbf{\Xi}(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1) \right)^{-1} \mathbf{\Xi}(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} \right)^T \\ &= \mathbf{Y}_f^+(\bar{t})\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})}\mathbf{\Xi}^T(\bar{t})\mathbf{\Phi}_f(\bar{t})\mathbf{\Xi}(\bar{t})\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})}\mathbf{Y}_f^{+T}(\bar{t}). \end{aligned} \quad (\text{A.54})$$

Intéressons-nous dans un premier temps au produit de $\mathbf{Y}_f^+(\bar{t})\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})}\mathbf{\Xi}^T(\bar{t})$ par $\mathbf{\Phi}_f(\bar{t})$:

$$\begin{aligned} &\mathbf{Y}_f^+(\bar{t})\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})}\mathbf{\Xi}^T(\bar{t})\mathbf{\Phi}_f(\bar{t}) \\ &= \left(\lambda\mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1) + \check{\mathbf{z}}_f^+(\bar{t})\check{\mathbf{\xi}}^{-T}(\bar{t}) \right) \left(\mathbf{\Phi}_f(\bar{t}-1) - \frac{\boldsymbol{\chi}_f^+(\bar{t})\boldsymbol{\chi}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\zeta}_f(\bar{t})} \right) / \lambda \\ &= \mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1)\mathbf{\Phi}_f(\bar{t}-1) - \mathbf{Y}_f^+(\bar{t}-1)\mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)}\mathbf{\Xi}^T(\bar{t}-1) \frac{\boldsymbol{\chi}_f^+(\bar{t})\boldsymbol{\chi}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\zeta}_f(\bar{t})} \\ &\quad + \frac{1}{\lambda}\check{\mathbf{z}}_f^+(\bar{t})\check{\mathbf{\xi}}^{-T}(\bar{t})\mathbf{\Phi}_f(\bar{t}-1) - \frac{1}{\lambda}\check{\mathbf{z}}_f^+(\bar{t})\check{\mathbf{\xi}}^{-T}(\bar{t}) \frac{\boldsymbol{\chi}_f^+(\bar{t})\boldsymbol{\chi}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\zeta}_f(\bar{t})}. \end{aligned} \quad (\text{A.55})$$

Sachant que :

$$\boldsymbol{\chi}_f^+(\bar{t}) = \mathbf{\Phi}_f(\bar{t}-1)\check{\mathbf{\xi}}^-(\bar{t}) \quad (\text{A.56})$$

$$\boldsymbol{\zeta}_f(\bar{t}) = \boldsymbol{\chi}_f^{+T}(\bar{t})\check{\mathbf{\xi}}^-(\bar{t}) \quad (\text{A.57})$$

L'équation (A.55) se simplifie comme suit :

$$\begin{aligned}
& \mathbf{Y}_f^+(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t})} \mathbf{\Xi}^T(\bar{t}) \mathbf{\Phi}_f(\bar{t}) \\
&= \mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) \mathbf{\Phi}_f(\bar{t}-1) - \mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) \frac{\boldsymbol{\chi}_f^+(\bar{t}) \boldsymbol{\chi}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\zeta}_f(\bar{t})} \\
&+ \frac{1}{\lambda} \check{\mathbf{z}}_f^+(\bar{t}) \boldsymbol{\chi}_f^{+T}(\bar{t}) - \frac{1}{\lambda} \check{\mathbf{z}}_f^+(\bar{t}) \frac{\boldsymbol{\zeta}_f(\bar{t}) \boldsymbol{\chi}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\zeta}_f(\bar{t})} \\
&= \mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) \mathbf{\Phi}_f(\bar{t}-1) - \mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) \frac{\boldsymbol{\chi}_f^+(\bar{t}) \boldsymbol{\chi}_f^{+T}(\bar{t})}{\lambda + \boldsymbol{\zeta}_f(\bar{t})} \\
&+ \frac{\lambda}{\lambda + \boldsymbol{\zeta}_f(\bar{t})} \check{\mathbf{z}}_f^+(\bar{t}) \boldsymbol{\chi}_f^{+T}(\bar{t}).
\end{aligned} \tag{A.58}$$

Il est alors possible de mettre en évidence le vecteur $\check{\mathbf{z}}_f^+$:

$$\begin{aligned}
\mathbf{Y}_f^+(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t})} \mathbf{\Xi}^T(\bar{t}) \mathbf{\Phi}_f(\bar{t}) &= \mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) \mathbf{\Phi}_f(\bar{t}-1) \\
&+ \frac{1}{\lambda + \boldsymbol{\zeta}_f(\bar{t})} \underbrace{\left(\check{\mathbf{z}}_f^+(\bar{t}) - \mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) \boldsymbol{\chi}_f^+(\bar{t}) \right)}_{\sqrt{\frac{\lambda + \boldsymbol{\zeta}_f(\bar{t})}{\lambda}} \check{\mathbf{z}}_f^+(\bar{t})} \boldsymbol{\chi}_f^{+T}(\bar{t}) \tag{A.59}
\end{aligned}$$

ce qui, après simplification, donne :

$$\begin{aligned}
\mathbf{Y}_f^+(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t})} \mathbf{\Xi}^T(\bar{t}) \mathbf{\Phi}_f(\bar{t}) &= \mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{+\perp}(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) \mathbf{\Phi}_f(\bar{t}-1) \\
&+ \frac{1}{\sqrt{\lambda^2 + \lambda \boldsymbol{\zeta}_f(\bar{t})}} \check{\mathbf{z}}_f^+(\bar{t}) \boldsymbol{\chi}_f^{+T}(\bar{t}). \tag{A.60}
\end{aligned}$$

La dernière phase de la démonstration consiste à multiplier l'équation précédente à droite par

$$\left(\mathbf{Y}_f^+(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^+(\bar{t})} \mathbf{\Xi}^T(\bar{t}) \right)^T :$$

$$\begin{aligned} & \mathbf{Y}_f^+(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^+(\bar{t})} \mathbf{\Xi}^T(\bar{t}) \mathbf{\Phi}_f(\bar{t}) \mathbf{\Xi}(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^+(\bar{t})} \mathbf{Y}_f^{+T}(\bar{t}) \\ &= \left(\mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^+(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) \mathbf{\Phi}_f(\bar{t}-1) + \frac{1}{\sqrt{\lambda^2 + \lambda \zeta_f(\bar{t})}} \check{\mathbf{z}}_f^+(\bar{t}) \boldsymbol{\chi}_f^{+T}(\bar{t}) \right) \\ & \quad \left(\lambda \mathbf{\Xi}(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^+(\bar{t}-1)} \mathbf{Y}_f^{+T}(\bar{t}-1) + \check{\boldsymbol{\xi}}^-(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t}) \right) \\ &= \lambda \mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^+(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) \mathbf{\Phi}_f(\bar{t}-1) \mathbf{\Xi}(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^+(\bar{t}-1)} \mathbf{Y}_f^{+T}(\bar{t}-1) \quad (\text{A.61}) \\ &+ \sqrt{\frac{\lambda}{\lambda + \zeta_f(\bar{t})}} \check{\mathbf{z}}_f^+(\bar{t}) \boldsymbol{\chi}_f^{+T}(\bar{t}) \mathbf{\Xi}(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^+(\bar{t}-1)} \mathbf{Y}_f^{+T}(\bar{t}-1) \\ &+ \mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^+(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) \mathbf{\Phi}_f(\bar{t}-1) \check{\boldsymbol{\xi}}^-(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t}) \\ &+ \frac{1}{\sqrt{\lambda^2 + \lambda \zeta_f(\bar{t})}} \check{\mathbf{z}}_f^+(\bar{t}) \boldsymbol{\chi}_f^{+T}(\bar{t}) \check{\boldsymbol{\xi}}^-(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t}). \end{aligned}$$

Pour simplifier quelque peu les calculs, il convient alors d'ajouter et de soustraire $\check{\mathbf{z}}_f^+(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t})$:

$$\begin{aligned} & \mathbf{Y}_f^+(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^+(\bar{t})} \mathbf{\Xi}^T(\bar{t}) \mathbf{\Phi}_f(\bar{t}) \mathbf{\Xi}(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^+(\bar{t})} \mathbf{Y}_f^{+T}(\bar{t}) \\ &= \lambda \mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^+(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) \mathbf{\Phi}_f(\bar{t}-1) \mathbf{\Xi}(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^+(\bar{t}-1)} \mathbf{Y}_f^{+T}(\bar{t}-1) \\ & \quad - \check{\mathbf{z}}_f^+(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t}) + \check{\mathbf{z}}_f^+(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t}) + \sqrt{\frac{\lambda}{\lambda + \zeta_f(\bar{t})}} \check{\mathbf{z}}_f^+(\bar{t}) \boldsymbol{\chi}_f^{+T}(\bar{t}) \mathbf{\Xi}(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^+(\bar{t}-1)} \mathbf{Y}_f^{+T}(\bar{t}-1) \\ &+ \mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^+(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) \mathbf{\Phi}_f(\bar{t}-1) \check{\boldsymbol{\xi}}^-(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t}) \\ &+ \frac{1}{\sqrt{\lambda^2 + \lambda \zeta_f(\bar{t})}} \check{\mathbf{z}}_f^+(\bar{t}) \boldsymbol{\chi}_f^{+T}(\bar{t}) \check{\boldsymbol{\xi}}^-(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t}). \end{aligned} \quad (\text{A.62})$$

Il est ainsi possible de procéder à quelques simplifications et regroupements :

$$\begin{aligned}
& \mathbf{Y}_f^+(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})} \mathbf{\Xi}^T(\bar{t}) \mathbf{\Phi}_f(\bar{t}) \mathbf{\Xi}(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})} \mathbf{Y}_f^{+T}(\bar{t}) \\
&= \lambda \mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) \mathbf{\Phi}_f(\bar{t}-1) \mathbf{\Xi}(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} \mathbf{Y}_f^{+T}(\bar{t}-1) \\
&\quad - \check{\mathbf{z}}_f^+(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t}) + \sqrt{\frac{\lambda}{\lambda + \zeta_f(\bar{t})}} \check{\mathbf{z}}_f^+(\bar{t}) \boldsymbol{\chi}_f^{+T}(\bar{t}) \mathbf{\Xi}^T(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} \mathbf{Y}_f^{+T}(\bar{t}-1) \\
&\quad + \frac{\zeta_f(\bar{t})}{\sqrt{\lambda^2 + \lambda \zeta_f(\bar{t})}} \check{\mathbf{z}}_f^+(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t}) + \sqrt{\frac{\lambda}{\lambda + \zeta_f(\bar{t})}} \check{\mathbf{z}}_f^+(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t}) \\
&\quad - \sqrt{\frac{\lambda}{\lambda + \zeta_f(\bar{t})}} \check{\mathbf{z}}_f^+(\bar{t}) \boldsymbol{\chi}_f^{+T}(\bar{t}) \mathbf{\Xi}^T(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} \mathbf{Y}_f^{+T}(\bar{t}-1) \\
&= \lambda \mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) \mathbf{\Phi}_f(\bar{t}-1) \mathbf{\Xi}(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} \mathbf{Y}_f^{+T}(\bar{t}-1) - \check{\mathbf{z}}_f^+(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t}) \\
&\quad + \mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) \boldsymbol{\chi}_f^+(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t}) + \sqrt{\frac{\lambda + \zeta_f(\bar{t})}{\lambda}} \check{\mathbf{z}}_f^+(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t}).
\end{aligned} \tag{A.63}$$

On reconnaît alors $\check{\mathbf{z}}_f^+(\bar{t})$ et cette dernière équation devient :

$$\begin{aligned}
& \mathbf{Y}_f^+(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})} \mathbf{\Xi}^T(\bar{t}) \mathbf{\Phi}_f(\bar{t}) \mathbf{\Xi}(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})} \mathbf{Y}_f^{+T}(\bar{t}) \\
&= \lambda \mathbf{Y}_f^+(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} \mathbf{\Xi}^T(\bar{t}-1) \mathbf{\Phi}_f(\bar{t}-1) \mathbf{\Xi}(\bar{t}-1) \mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t}-1)} \mathbf{Y}_f^{+T}(\bar{t}-1) \\
&\quad - \check{\mathbf{z}}_f^+(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t}) + \check{\mathbf{z}}_f^+(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t}).
\end{aligned} \tag{A.64}$$

En associant les équations (A.53) et (A.64) et sachant que :

$$\mathbf{Z}_f^+(t) \mathbf{Z}_f^{+T}(t) = \mathbf{Y}_f^+(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})} \mathbf{\Xi}^T(\bar{t}) \mathbf{\Phi}_f(\bar{t}) \mathbf{\Xi}(\bar{t}) \mathbf{\Pi}_{\mathbf{U}_f^{\perp}(\bar{t})} \mathbf{Y}_f^{+T}(\bar{t}), \tag{A.65}$$

la relation suivante est vérifiée :

$$\mathbf{z}_f^+(\bar{t}) \mathbf{z}_f^{+T}(\bar{t}) = \check{\mathbf{z}}_f^+(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t}) - \check{\mathbf{z}}_f^+(\bar{t}) \check{\mathbf{z}}_f^{+T}(\bar{t}). \tag{A.66}$$

A.5 Preuve de l'algorithme (5.81)

Considérons le critère suivant :

$$J(\mathbf{P}_f) = \mathbb{E} \left\| \mathbf{z}_{f_2}^+ - \mathbf{P}_f^T \mathbf{z}_{f_1}^+ \right\|^2. \tag{A.67}$$

Son minimum est atteint lorsque sa dérivée s'annule i.e. pour :

$$\mathbf{P}_f^T(t) = \mathbf{R}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+}(t) \mathbf{R}_{\mathbf{z}_{f_1}^+}^{-1}(t). \tag{A.68}$$

A chaque nouvelle acquisition, ces matrices de covariance sont mises à jour de la façon suivante :

$$\mathbf{R}_{\mathbf{z}_{f_1}^+}(t) = \lambda \mathbf{R}_{\mathbf{z}_{f_1}^+}(t-1) + \mathbf{z}_{f_1}^+(t) \mathbf{z}_{f_1}^{+T}(t) \quad (\text{A.69})$$

$$\mathbf{R}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+}(t) = \lambda \mathbf{R}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+}(t-1) + \mathbf{z}_{f_2}^+(t) \mathbf{z}_{f_1}^{+T}(t). \quad (\text{A.70})$$

En appliquant le lemme d'inversion matricielle à la matrice $\mathbf{L}_f(t) = \mathbf{R}_{\mathbf{z}_{f_1}^+}^{-1}(t)$, nous obtenons :

$$\begin{aligned} \mathbf{L}_f(t) &= \left(\lambda \mathbf{R}_{\mathbf{z}_{f_1}^+}(t-1) + \mathbf{z}_{f_1}^+(t) \mathbf{z}_{f_1}^{+T}(t) \right)^{-1} = \\ &= \frac{1}{\lambda} \left(\mathbf{L}_f(t-1) - \frac{\mathbf{L}_f(t-1) \mathbf{z}_{f_1}^+(t) \mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t-1)}{\lambda + \mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t-1) \mathbf{z}_{f_1}^+(t)} \right). \end{aligned} \quad (\text{A.71})$$

Le projecteur estimé vaut alors :

$$\mathbf{P}_f^T(t) = \left(\lambda \mathbf{R}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+}(t-1) + \mathbf{z}_{f_2}^+(t) \mathbf{z}_{f_1}^{+T}(t) \right) \mathbf{L}_f(t) \quad (\text{A.72})$$

avec :

$$\mathbf{L}_f(t) = \frac{1}{\lambda} \left(\mathbf{L}_f(t-1) - \frac{\mathbf{L}_f(t-1) \mathbf{z}_{f_1}^+(t) \mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t-1)}{\lambda + \mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t-1) \mathbf{z}_{f_1}^+(t)} \right). \quad (\text{A.73})$$

Il est possible de simplifier cette écriture en remarquant que :

$$\lambda \mathbf{R}_{\mathbf{z}_{f_2}^+ \mathbf{z}_{f_1}^+}(t-1) = \lambda \mathbf{P}_f^T(t-1) \mathbf{R}_{\mathbf{z}_{f_1}^+}(t-1) = \mathbf{P}_f^T(t-1) \left(\mathbf{L}_f^{-1}(t-1) - \mathbf{z}_{f_1}^+(t-1) \mathbf{z}_{f_1}^{+T}(t-1) \right). \quad (\text{A.74})$$

Ainsi :

$$\begin{aligned} \mathbf{P}_f^T(t) &= \left(\mathbf{P}_f^T(t-1) \left(\mathbf{L}_f^{-1}(t-1) - \mathbf{z}_{f_1}^+(t-1) \mathbf{z}_{f_1}^{+T}(t-1) \right) + \mathbf{z}_{f_2}^+(t) \mathbf{z}_{f_1}^{+T}(t) \right) \mathbf{L}_f(t) \\ &= \mathbf{P}_f^T(t-1) + \left(\mathbf{z}_{f_2}^+(t) - \mathbf{P}_f^T(t-1) \mathbf{z}_{f_1}^+(t) \right) \mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t). \end{aligned} \quad (\text{A.75})$$

De plus :

$$\begin{aligned} \mathbf{K}_f(t) &= \mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t) \\ &= \frac{1}{\lambda} \left(\mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t-1) - \frac{\mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t-1) \mathbf{z}_{f_1}^+(t) \mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t-1)}{\lambda + \mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t-1) \mathbf{z}_{f_1}^+(t)} \right) \\ &= \frac{\mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t-1)}{\lambda + \mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t-1) \mathbf{z}_{f_1}^+(t)} \end{aligned} \quad (\text{A.76})$$

puisque $\mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t-1) \mathbf{z}_{f_1}^+(t) \in \mathbb{R}$. L'algorithme récursif complet s'écrit alors :

$$\mathbf{K}_f(t) = \frac{\mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t-1)}{\lambda + \mathbf{z}_{f_1}^{+T}(t) \mathbf{L}_f(t-1) \mathbf{z}_{f_1}^+(t)} \quad (\text{A.77a})$$

$$\mathbf{L}_f(t) = \frac{1}{\lambda} \left(\mathbf{L}_f(t-1) - \mathbf{L}_f(t-1) \mathbf{z}_{f_1}^+(t) \mathbf{K}_f(t) \right) \quad (\text{A.77b})$$

$$\mathbf{P}_f^T(t) = \mathbf{P}_f^T(t-1) + \left(\mathbf{z}_{f_2}^+(t) - \mathbf{P}_f^T(t-1) \mathbf{z}_{f_1}^+(t) \right) \mathbf{K}_f(t). \quad (\text{A.77c})$$

A.6 Preuve de l'algorithme (5.122)

Nous savons que :

$$\mathbf{P}_f^T(t) = \left(\mathbf{R}_{\underline{z}_{f_2}^+ \underline{z}_{f_1}^+}(t) - \mathbf{R}_{\underline{z}_{f_2}^+ \underline{z}_{f_1}^+}(t) \right) \left(\mathbf{R}_{\underline{z}_{f_1}^+}(t) - \mathbf{R}_{\underline{z}_{f_1}^+}(t) \right)^{-1} \quad (\text{A.78})$$

avec :

$$\mathbf{R}_{\underline{z}_{f_1}^+}(t) = \lambda \mathbf{R}_{\underline{z}_{f_1}^+}(t-1) + \underline{z}_{f_1}^+(t) \underline{z}_{f_1}^{+T}(t) \quad (\text{A.79})$$

$$\mathbf{R}_{\underline{z}_{f_1}^+}(t) = \lambda \mathbf{R}_{\underline{z}_{f_1}^+}(t-1) + \underline{z}_{f_1}^+(t) \underline{z}_{f_1}^{+T}(t) \quad (\text{A.80})$$

$$\mathbf{R}_{\underline{z}_{f_2}^+ \underline{z}_{f_1}^+}(t) = \lambda \mathbf{R}_{\underline{z}_{f_2}^+ \underline{z}_{f_1}^+}(t-1) + \underline{z}_{f_2}^+(t) \underline{z}_{f_1}^{+T}(t) \quad (\text{A.81})$$

$$\mathbf{R}_{\underline{z}_{f_2}^+ \underline{z}_{f_1}^+}(t) = \lambda \mathbf{R}_{\underline{z}_{f_2}^+ \underline{z}_{f_1}^+}(t-1) + \underline{z}_{f_2}^+(t) \underline{z}_{f_1}^{+T}(t). \quad (\text{A.82})$$

Ainsi :

$$\mathbf{R}_{\underline{z}_{f_1}^+}(t) - \mathbf{R}_{\underline{z}_{f_1}^+}(t) = \lambda \left(\mathbf{R}_{\underline{z}_{f_1}^+}(t-1) - \mathbf{R}_{\underline{z}_{f_1}^+}(t-1) \right) + \underline{z}_{f_1}^+(t) \underline{z}_{f_1}^{+T}(t) - \underline{z}_{f_1}^+(t) \underline{z}_{f_1}^{+T}(t). \quad (\text{A.83})$$

La principale étape de cette démonstration est de proposer une mise à jour récursive de cette équation. Pour cela, nous allons lui appliquer par deux fois le lemme d'inversion matricielle. Rappelons à ce stade ce lemme [Bro91] :

Lemme A.2. Soient \mathbf{A} , \mathbf{B} , \mathbf{C} et \mathbf{D} quatre matrices telles que \mathbf{A} , \mathbf{C} et $\mathbf{A} + \mathbf{BCD}$ soient inversibles. Alors :

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{C} + \mathbf{DA}^{-1} \mathbf{B})^{-1} \mathbf{DA}^{-1}. \quad (\text{A.84})$$

Ainsi, en posant dans un premier temps :

$$\mathbf{A} = \lambda \left(\mathbf{R}_{\underline{z}_{f_1}^+}(t-1) - \mathbf{R}_{\underline{z}_{f_1}^+}(t-1) \right) + \underline{z}_{f_1}^+(t) \underline{z}_{f_1}^{+T}(t) = \underline{\mathbf{L}}_f^{-1}(t) \quad (\text{A.85})$$

$$\mathbf{B} = \underline{z}_{f_1}^+(t) \quad (\text{A.86})$$

$$\mathbf{C} = 1 \quad (\text{A.87})$$

$$\mathbf{D} = \underline{z}_{f_1}^{+T}(t) \quad (\text{A.88})$$

et en appliquant le lemme d'inversion matricielle à partir de cette décomposition, nous obtenons :

$$\underline{\mathbf{L}}_f(t) = \left(\mathbf{R}_{\underline{z}_{f_1}^+}(t) - \mathbf{R}_{\underline{z}_{f_1}^+}(t) \right)^{-1} = \underline{\mathbf{L}}_f(t) - \frac{\underline{\mathbf{L}}_f(t) \underline{z}_{f_1}^+(t) \underline{z}_{f_1}^{+T}(t) \underline{\mathbf{L}}_f(t)}{1 + \underline{z}_{f_1}^{+T}(t) \underline{\mathbf{L}}_f(t) \underline{z}_{f_1}^+(t)} \quad (\text{A.89})$$

avec :

$$\underline{\mathbf{L}}_f(t) = \left(\lambda \left(\mathbf{R}_{\underline{z}_{f_1}^+}(t-1) - \mathbf{R}_{\underline{z}_{f_1}^+}(t-1) \right) + \underline{z}_{f_1}^+(t) \underline{z}_{f_1}^{+T}(t) \right)^{-1}. \quad (\text{A.90})$$

En appliquant ensuite le lemme d'inversion matricielle à l'équation (A.90), la relation suivante est vérifiée :

$$\underline{\mathbf{L}}_f(t) = \frac{1}{\lambda} \left(\underline{\mathbf{L}}_f(t-1) - \frac{\underline{\mathbf{L}}_f(t-1) \underline{z}_{f_1}^+(t) \underline{z}_{f_1}^{+T}(t) \underline{\mathbf{L}}_f(t-1)}{\lambda + \underline{z}_{f_1}^{+T}(t) \underline{\mathbf{L}}_f(t-1) \underline{z}_{f_1}^+(t)} \right) \quad (\text{A.91})$$

en posant :

$$\underline{\mathbf{L}}_f(t-1) = \left(\mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+}(t-1) - \mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+}(t-1) \right)^{-1}. \quad (\text{A.92})$$

La dernière phase de cette démonstration consiste à exprimer $\mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_1}^+}(t) - \mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_1}^+}(t)$ en fonction des matrices de covariance de l'instant précédent et des nouveaux vecteurs acquis. Or, nous savons que :

$$\begin{aligned} \mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_1}^+}(t) - \mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_1}^+}(t) &= \lambda \left(\mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_1}^+}(t-1) - \mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_1}^+}(t-1) \right) \\ &\quad + \mathbf{z}_{f_2}^+(t) \mathbf{z}_{f_1}^{+T}(t) - \mathbf{z}_{f_2}^+(t) \mathbf{z}_{f_1}^{+T}(t) \end{aligned} \quad (\text{A.93})$$

à l'aide des équations (A.81) et (A.82). Puisque :

$$\mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_1}^+}(t-1) - \mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_1}^+}(t-1) = \mathbf{P}_f^T(t-1) \left(\mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+}(t-1) - \mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+}(t-1) \right) \quad (\text{A.94})$$

et que :

$$\lambda \left(\mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+}(t-1) - \mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+}(t-1) \right) = \mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+}(t) - \mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+}(t) - \mathbf{z}_{f_1}^+(t) \mathbf{z}_{f_1}^{+T}(t) - \mathbf{z}_{f_1}^+(t) \mathbf{z}_{f_1}^{+T}(t), \quad (\text{A.95})$$

nous pouvons écrire que :

$$\begin{aligned} \mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_1}^+}(t) - \mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_1}^+}(t) &= \mathbf{P}_f^T(t-1) \left(\mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+}(t) - \mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+}(t) \right) \\ &\quad + \left(\mathbf{z}_{f_2}^+(t) - \hat{\mathbf{P}}_f^T(t-1) \mathbf{z}_{f_1}^+(t) \right) \mathbf{z}_{f_1}^{+T}(t) - \left(\mathbf{z}_{f_2}^+(t) - \mathbf{P}_f^T(t-1) \mathbf{z}_{f_1}^+(t) \right) \mathbf{z}_{f_1}^{+T}(t). \end{aligned} \quad (\text{A.96})$$

En postmultipliant cette équation par $\underline{\mathbf{L}}_f(t) = \left(\mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+}(t) - \mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+}(t) \right)^{-1}$, nous obtenons :

$$\mathbf{P}_f^T(t) = \left(\mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_1}^+}(t) - \mathbf{R}_{\underline{\mathbf{z}}_{f_2}^+ \underline{\mathbf{z}}_{f_1}^+}(t) \right) \left(\mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+}(t) - \mathbf{R}_{\underline{\mathbf{z}}_{f_1}^+}(t) \right)^{-1} \quad (\text{A.97})$$

$$= \mathbf{P}_f^T(t-1) + \left(\mathbf{z}_{f_2}^+(t) - \hat{\mathbf{P}}_f^T(t-1) \mathbf{z}_{f_1}^+(t) \right) \mathbf{z}_{f_1}^{+T}(t) \underline{\mathbf{L}}_f(t) \quad (\text{A.98})$$

$$- \left(\mathbf{z}_{f_2}^+(t) - \mathbf{P}_f^T(t-1) \mathbf{z}_{f_1}^+(t) \right) \mathbf{z}_{f_1}^{+T}(t) \underline{\mathbf{L}}_f(t). \quad (\text{A.99})$$

L'algorithme final s'écrit :

$$\underline{\mathbf{L}}_f(t) = \frac{1}{\lambda} \left(\underline{\mathbf{L}}_f(t-1) - \frac{\underline{\mathbf{L}}_f(t-1) \mathbf{z}_{f_1}^+(t) \mathbf{z}_{f_1}^{+T}(t) \underline{\mathbf{L}}_f(t-1)}{\lambda + \mathbf{z}_{f_1}^{+T}(t) \underline{\mathbf{L}}_f(t-1) \mathbf{z}_{f_1}^+(t)} \right) \quad (\text{A.100a})$$

$$\underline{\mathbf{L}}_f(t) = \underline{\mathbf{L}}_f(t) + \frac{\underline{\mathbf{L}}_f(t) \mathbf{z}_{f_1}^+(t) \mathbf{z}_{f_1}^{+T}(t) \underline{\mathbf{L}}_f(t)}{\lambda + \mathbf{z}_{f_1}^{+T}(t) \underline{\mathbf{L}}_f(t) \mathbf{z}_{f_1}^+(t)} \quad (\text{A.100b})$$

$$\begin{aligned} \mathbf{P}_f^T(t) &= \mathbf{P}_f^T(t-1) + \left(\mathbf{z}_{f_2}^+(t) - \hat{\mathbf{P}}_f^T(t-1) \mathbf{z}_{f_1}^+(t) \right) \mathbf{z}_{f_1}^{+T}(t) \underline{\mathbf{L}}_f(t) \\ &\quad - \left(\mathbf{z}_{f_2}^+(t) - \mathbf{P}_f^T(t-1) \mathbf{z}_{f_1}^+(t) \right) \mathbf{z}_{f_1}^{+T}(t) \underline{\mathbf{L}}_f(t). \end{aligned} \quad (\text{A.100c})$$

A.7 Preuve de l'algorithme (5.138)

Lorsque le nombre d'instruments composant la variable instrumentale ξ est strictement supérieur à n , le minimum du critère J_{IV} défini par :

$$J_{IV}(\mathbf{P}_f(t)) = \text{Tr} \left(\left(\mathbf{R}_{z_{f_2}} \xi(t) - \mathbf{P}_f^T(t) \mathbf{R}_{z_{f_1}} \xi(t) \right) \left(\mathbf{R}_{z_{f_2}}^T \xi(t) - \mathbf{R}_{z_{f_1}}^T \xi(t) \mathbf{P}_f(t) \right) \right). \quad (\text{A.101})$$

est atteint pour :

$$\mathbf{P}_f^T(t) = \mathbf{R}_{z_{f_2}} \xi(t) \mathbf{R}_{z_{f_1}}^\dagger \xi(t) = \mathbf{R}_{z_{f_2}} \xi(t) \mathbf{R}_{z_{f_1}}^T \xi(t) \left(\mathbf{R}_{z_{f_1}} \xi(t) \mathbf{R}_{z_{f_1}}^T \xi(t) \right)^{-1} \quad (\text{A.102})$$

puisque la matrice $\mathbf{R}_{z_{f_1}} \xi$ n'est pas carrée. L'application du lemme d'inversion matricielle n'est donc plus directement applicable. Il est cependant possible de s'inspirer des travaux de B. Friedlander sur la méthode de la variable instrumentale étendue [Fri84] pour développer un algorithme récursif capable de gérer ce type de matrice. Considérons que nous ayons accès à de nouvelles mesures d'entrée et de sortie à l'instant t . Les matrices de covariance $\mathbf{R}_{z_{f_1}} \xi$ et $\mathbf{R}_{z_{f_2}} \xi$ sont alors mises à jour comme suit :

$$\mathbf{R}_{z_{f_1}} \xi(t) = \lambda \mathbf{R}_{z_{f_1}} \xi(t-1) + \mathbf{z}_{f_1}(t) \xi^T(t) \quad (\text{A.103})$$

$$\mathbf{R}_{z_{f_2}} \xi(t) = \lambda \mathbf{R}_{z_{f_2}} \xi(t-1) + \mathbf{z}_{f_2}(t) \xi^T(t). \quad (\text{A.104})$$

En posant :

$$\mathbf{L}_f(t) = \left(\mathbf{R}_{z_{f_1}} \xi(t) \mathbf{R}_{z_{f_1}}^T \xi(t) \right)^{-1}, \quad (\text{A.105})$$

il est aisé de vérifier que :

$$\begin{aligned} \mathbf{P}_f^T(t) &= \mathbf{R}_{z_{f_2}} \xi(t) \mathbf{R}_{z_{f_1}}^T \xi(t) \mathbf{L}_f(t) + \mathbf{P}_f^T(t-1) - \mathbf{P}_f^T(t-1) \\ &= \mathbf{R}_{z_{f_2}} \xi(t) \mathbf{R}_{z_{f_1}}^T \xi(t) \mathbf{L}_f(t) + \mathbf{P}_f^T(t-1) - \mathbf{P}_f^T(t-1) \underbrace{\mathbf{R}_{z_{f_1}} \xi(t) \mathbf{R}_{z_{f_1}}^T \xi(t) \mathbf{L}_f(t)}_{\mathbf{I}} \\ &= \mathbf{P}_f^T(t-1) + \left(\mathbf{R}_{z_{f_2}} \xi(t) - \mathbf{P}_f^T(t-1) \mathbf{R}_{z_{f_1}} \xi(t) \right) \mathbf{R}_{z_{f_1}}^T \xi(t) \mathbf{L}_f(t) \\ &= \mathbf{P}_f^T(t-1) + \boldsymbol{\varepsilon}(t) \mathbf{R}_{z_{f_1}}^T \xi(t) \mathbf{L}_f(t) \end{aligned} \quad (\text{A.106})$$

avec :

$$\boldsymbol{\varepsilon}(t) = \mathbf{R}_{z_{f_2}} \xi(t) - \mathbf{P}_f^T(t-1) \mathbf{R}_{z_{f_1}} \xi(t). \quad (\text{A.107})$$

Calculons dans un premier temps le terme $\boldsymbol{\varepsilon}(t) \mathbf{R}_{z_{f_1}}^T \xi(t)$ lorsque de nouvelles données sont mesurées. Pour cela, introduisons les relations (A.103) et (A.104) au sein même de $\boldsymbol{\varepsilon}(t) \mathbf{R}_{z_{f_1}}^T \xi(t)$:

$$\begin{aligned} \boldsymbol{\varepsilon}(t) \mathbf{R}_{z_{f_1}}^T \xi(t) &= \left(\mathbf{R}_{z_{f_2}} \xi(t) - \mathbf{P}_f^T(t-1) \mathbf{R}_{z_{f_1}} \xi(t) \right) \mathbf{R}_{z_{f_1}}^T \xi(t) \\ &= \left(\lambda \mathbf{R}_{z_{f_2}} \xi(t-1) + \mathbf{z}_{f_2}(t) \xi^T(t) - \mathbf{P}_f^T(t-1) \left(\lambda \mathbf{R}_{z_{f_1}} \xi(t-1) + \mathbf{z}_{f_1}(t) \xi^T(t) \right) \right) \\ &\quad \left(\lambda \mathbf{R}_{z_{f_1}} \xi(t-1) + \mathbf{z}_{f_1}(t) \xi^T(t) \right)^T. \end{aligned} \quad (\text{A.108})$$

En développant les différentes composantes de cette expression, nous obtenons :

$$\begin{aligned}
\boldsymbol{\varepsilon}(t)\mathbf{R}_{\mathbf{z}_{f_1}}^T\xi(t) &= \left(\lambda \left(\mathbf{R}_{\mathbf{z}_{f_2}}\xi(t-1) - \mathbf{P}_f^T(t-1)\mathbf{R}_{\mathbf{z}_{f_1}}\xi(t-1) \right) \right. \\
&\quad \left. + \left(\mathbf{z}_{f_2}(t) - \mathbf{P}_f^T(t-1)\mathbf{z}_{f_1}(t) \right) \xi^T(t) \right) \left(\lambda \mathbf{R}_{\mathbf{z}_{f_1}}\xi(t-1) + \mathbf{z}_{f_1}(t)\xi^T(t) \right)^T \\
&= \lambda^2 \underbrace{\left(\mathbf{R}_{\mathbf{z}_{f_2}}\xi(t-1) - \mathbf{P}_f^T(t-1)\mathbf{R}_{\mathbf{z}_{f_1}}\xi(t-1) \right) \mathbf{R}_{\mathbf{z}_{f_1}}^T\xi(t-1)}_0 \\
&\quad + \lambda \left(\mathbf{R}_{\mathbf{z}_{f_2}}\xi(t-1) - \mathbf{P}_f^T(t-1)\mathbf{R}_{\mathbf{z}_{f_1}}\xi(t-1) \right) \xi(t)\mathbf{z}_{f_1}^T(t) \\
&\quad + \lambda \left(\mathbf{z}_{f_2}(t) - \mathbf{P}_f^T(t-1)\mathbf{z}_{f_1}(t) \right) \xi^T(t)\mathbf{R}_{\mathbf{z}_{f_1}}^T\xi(t-1) \\
&\quad + \left(\mathbf{z}_{f_2}(t) - \mathbf{P}_f^T(t-1)\mathbf{z}_{f_1}(t) \right) \xi^T(t)\xi(t)\mathbf{z}_{f_1}^T(t) \\
&= \lambda \left(\mathbf{R}_{\mathbf{z}_{f_2}}\xi(t-1) - \mathbf{P}_f^T(t-1)\mathbf{R}_{\mathbf{z}_{f_1}}\xi(t-1) \right) \xi(t)\mathbf{z}_{f_1}^T(t) \\
&\quad + \left(\mathbf{z}_{f_2}(t) - \mathbf{P}_f^T(t-1)\mathbf{z}_{f_1}(t) \right) \left(\xi^T(t)\xi(t)\mathbf{z}_{f_1}^T(t) + \lambda \xi^T(t)\mathbf{R}_{\mathbf{z}_{f_1}}^T\xi(t-1) \right) \\
&= \begin{bmatrix} \xi^T(t)\mathbf{R}_{\mathbf{z}_{f_2}}^T\xi(t-1) - \xi^T(t)\mathbf{R}_{\mathbf{z}_{f_1}}^T\xi(t-1)\mathbf{P}_f(t-1) \\ \mathbf{z}_{f_2}^T(t) - \mathbf{z}_{f_1}^T(t)\mathbf{P}_f(t-1) \end{bmatrix}^T \begin{bmatrix} \lambda \mathbf{z}_{f_1}^T(t) \\ \xi^T(t)\xi(t)\mathbf{z}_{f_1}^T(t) + \lambda \xi^T(t)\mathbf{R}_{\mathbf{z}_{f_1}}^T\xi(t-1) \end{bmatrix} \\
&= \begin{bmatrix} \left[\xi^T(t)\mathbf{R}_{\mathbf{z}_{f_2}}^T\xi(t-1) \right] - \left[\xi^T(t)\mathbf{R}_{\mathbf{z}_{f_1}}^T\xi(t-1) \right] \mathbf{P}_f(t-1) \\ \mathbf{z}_{f_2}^T(t) \end{bmatrix}^T \begin{bmatrix} 0 & \lambda \\ \lambda & \xi^T(t)\xi(t) \end{bmatrix} \begin{bmatrix} \xi^T(t)\mathbf{R}_{\mathbf{z}_{f_1}}^T\xi(t-1) \\ \mathbf{z}_{f_1}^T(t) \end{bmatrix}.
\end{aligned} \tag{A.109}$$

En posant :

$$\mathbf{g}_f(t) = \begin{bmatrix} \mathbf{R}_{\mathbf{z}_{f_2}^+}\xi(t)\xi(t) & \mathbf{z}_{f_2}^+(t) \end{bmatrix} \tag{A.110}$$

$$\boldsymbol{\Lambda}^{-1}(t) = \begin{bmatrix} 0 & \lambda \\ \lambda & \xi^T(t)\xi(t) \end{bmatrix} \tag{A.111}$$

$$\boldsymbol{\Psi}_f(t) = \begin{bmatrix} \mathbf{R}_{\mathbf{z}_{f_1}^+}\xi(t-1)\xi(t) & \mathbf{z}_{f_1}^+(t) \end{bmatrix}, \tag{A.112}$$

nous montrons que :

$$\boldsymbol{\varepsilon}(t)\mathbf{R}_{\mathbf{z}_{f_1}}^T\xi(t) = (\mathbf{g}_f(t) - \mathbf{P}_f^T(t-1)\boldsymbol{\Psi}_f(t))\boldsymbol{\Lambda}^{-1}(t)\boldsymbol{\Psi}_f^T(t). \tag{A.113}$$

Intéressons-nous maintenant à la mise à jour de $\mathbf{L}_f(t) = \left(\mathbf{R}_{\mathbf{z}_{f_1}} \boldsymbol{\xi}(t) \mathbf{R}_{\mathbf{z}_{f_1}}^T \boldsymbol{\xi}(t) \right)^{-1}$. A chaque nouvelle acquisition :

$$\begin{aligned}
\mathbf{L}_f^{-1}(t) &= \left(\mathbf{R}_{\mathbf{z}_{f_1}} \boldsymbol{\xi}(t) \mathbf{R}_{\mathbf{z}_{f_1}}^T \boldsymbol{\xi}(t) \right)^{-1} \\
&= \left(\lambda \mathbf{R}_{\mathbf{z}_{f_1}} \boldsymbol{\xi}(t-1) + \mathbf{z}_{f_1}(t) \boldsymbol{\xi}^T(t) \right) \left(\lambda \mathbf{R}_{\mathbf{z}_{f_1}} \boldsymbol{\xi}(t-1) + \mathbf{z}_{f_1}(t) \boldsymbol{\xi}^T(t) \right)^T \\
&= \lambda^2 \mathbf{L}_f^{-1}(t-1) + \lambda \mathbf{R}_{\mathbf{z}_{f_1}} \boldsymbol{\xi}(t-1) \boldsymbol{\xi}^T(t) \mathbf{z}_{f_1}^T(t) \\
&\quad + \lambda \mathbf{z}_{f_1}(t) \boldsymbol{\xi}^T(t) \mathbf{R}_{\mathbf{z}_{f_1}}^T \boldsymbol{\xi}(t-1) + \mathbf{z}_{f_1}(t) \boldsymbol{\xi}^T(t) \boldsymbol{\xi}(t) \mathbf{z}_{f_1}^T(t) \\
&= \lambda^2 \mathbf{L}_f^{-1}(t-1) + \begin{bmatrix} \mathbf{R}_{\mathbf{z}_{f_1}^+} \boldsymbol{\xi}(t-1) \boldsymbol{\xi}^T(t) & \mathbf{z}_{f_1}^+(t) \end{bmatrix} \begin{bmatrix} 0 & \lambda \\ \lambda & \boldsymbol{\xi}^T(t) \boldsymbol{\xi}(t) \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi}^T(t) \mathbf{R}_{\mathbf{z}_{f_1}^+}^T \boldsymbol{\xi}(t-1) \\ \mathbf{z}_{f_1}^{+T}(t) \end{bmatrix} \\
&= \lambda^2 \mathbf{L}_f^{-1}(t-1) + \boldsymbol{\Psi}_f(t) \boldsymbol{\Lambda}^{-1}(t) \boldsymbol{\Psi}_f^T(t).
\end{aligned} \tag{A.114}$$

En appliquant le lemme d'inversion matricielle, nous montrons que :

$$\begin{aligned}
\mathbf{L}_f(t) &= \frac{1}{\lambda^2} \left(\mathbf{L}_f(t-1) - \mathbf{L}_f(t-1) \boldsymbol{\Psi}_f(t) \right. \\
&\quad \left. \left(\boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1) \boldsymbol{\Psi}_f(t) + \lambda^2 \boldsymbol{\Lambda}(t) \right)^{-1} \boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1) \right) \tag{A.115}
\end{aligned}$$

avec :

$$\lambda^2 \boldsymbol{\Lambda}(t) = \begin{bmatrix} -\boldsymbol{\xi}^T(t) \boldsymbol{\xi}(t) & \lambda \\ \lambda & 0 \end{bmatrix}. \tag{A.116}$$

Finalement :

$$\begin{aligned}
&\boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t) \\
&= \frac{1}{\lambda^2} \left(\mathbf{I} - \boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1) \boldsymbol{\Psi}_f(t) \left(\boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1) \boldsymbol{\Psi}_f(t) + \lambda^2 \boldsymbol{\Lambda}(t) \right)^{-1} \right) \boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1) \\
&= \frac{1}{\lambda^2} \left(\boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1) \boldsymbol{\Psi}_f(t) + \lambda^2 \boldsymbol{\Lambda}(t) - \boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1) \boldsymbol{\Psi}_f(t) \right. \\
&\quad \left. \left(\boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1) \boldsymbol{\Psi}_f(t) + \lambda^2 \boldsymbol{\Lambda}(t) \right)^{-1} \boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1) \right) \\
&= \boldsymbol{\Lambda}(t) \left(\boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1) \boldsymbol{\Psi}_f(t) + \lambda^2 \boldsymbol{\Lambda}(t) \right)^{-1} \boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1).
\end{aligned} \tag{A.117}$$

Ainsi :

$$\begin{aligned}
\boldsymbol{\varepsilon}(t) \mathbf{R}_{\mathbf{z}_{f_1}}^T \boldsymbol{\xi}(t) \mathbf{L}_f(t) &= \left(\mathbf{g}_f(t) - \mathbf{P}_f^T(t-1) \boldsymbol{\Psi}_f(t) \right) \boldsymbol{\Lambda}^{-1}(t) \boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t) \\
&= \left(\mathbf{g}_f(t) - \mathbf{P}_f^T(t-1) \boldsymbol{\Psi}_f(t) \right) \left(\boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1) \boldsymbol{\Psi}_f(t) + \lambda^2 \boldsymbol{\Lambda}(t) \right)^{-1} \boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1). \tag{A.118}
\end{aligned}$$

L'algorithme récursif complet s'écrit alors :

$$\mathbf{g}_f(t) = \begin{bmatrix} \mathbf{R}_{\mathbf{z}_{f_2}^+} \boldsymbol{\xi}(t) & \mathbf{z}_{f_2}^+(t) \end{bmatrix} \quad (\text{A.119a})$$

$$\boldsymbol{\Lambda}(t) = \begin{bmatrix} -\boldsymbol{\xi}^T(t) \boldsymbol{\xi}(t) & \lambda \\ \lambda & 0 \end{bmatrix} \quad (\text{A.119b})$$

$$\boldsymbol{\Psi}_f(t) = \begin{bmatrix} \mathbf{R}_{\mathbf{z}_{f_1}^+} \boldsymbol{\xi}(t-1) \boldsymbol{\xi}(t) & \mathbf{z}_{f_1}^+(t) \end{bmatrix} \quad (\text{A.119c})$$

$$\mathbf{K}_f(t) = (\boldsymbol{\Lambda}(t) + \boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1) \boldsymbol{\Psi}_f(t))^{-1} \boldsymbol{\Psi}_f^T(t) \mathbf{L}_f(t-1) \quad (\text{A.119d})$$

$$\mathbf{P}_f^T(t) = \mathbf{P}_f^T(t-1) + (\mathbf{g}_f(t) - \mathbf{P}_f^T(t-1) \boldsymbol{\Psi}_f(t)) \mathbf{K}_f(t) \quad (\text{A.119e})$$

$$\mathbf{R}_{\mathbf{z}_{f_1}^+} \boldsymbol{\xi}(t) = \lambda \mathbf{R}_{\mathbf{z}_{f_1}^+} \boldsymbol{\xi}(t-1) + \mathbf{z}_{f_1}^+(t) \boldsymbol{\xi}^T(t) \quad (\text{A.119f})$$

$$\mathbf{R}_{\mathbf{z}_{f_2}^+} \boldsymbol{\xi}(t) = \lambda \mathbf{R}_{\mathbf{z}_{f_2}^+} \boldsymbol{\xi}(t-1) + \mathbf{z}_{f_2}^+(t) \boldsymbol{\xi}^T(t) \quad (\text{A.119g})$$

$$\mathbf{L}_f(t) = \frac{1}{\lambda^2} (\mathbf{L}_f(t-1) - \mathbf{L}_f(t-1) \boldsymbol{\Psi}(t) \mathbf{K}_f(t)) \quad (\text{A.119h})$$

Annexe B

Étude du système de transmission flexible

Sommaire

| | |
|---|-----|
| B.1 Étude du moteur à courant continu | 207 |
| B.2 Étude mécanique du système de transmission flexible | 209 |

Le système de transmission flexible étudié est constitué de trois poulies reliées par deux courroies métalliques comportant chacune deux ressorts (cf. fig. B.1). La première poulie est commandée par un moteur à courant continu asservi en position et en vitesse par un régulateur PD. La position de la poulie 3 est mesurée à l'aide d'un potentiomètre.

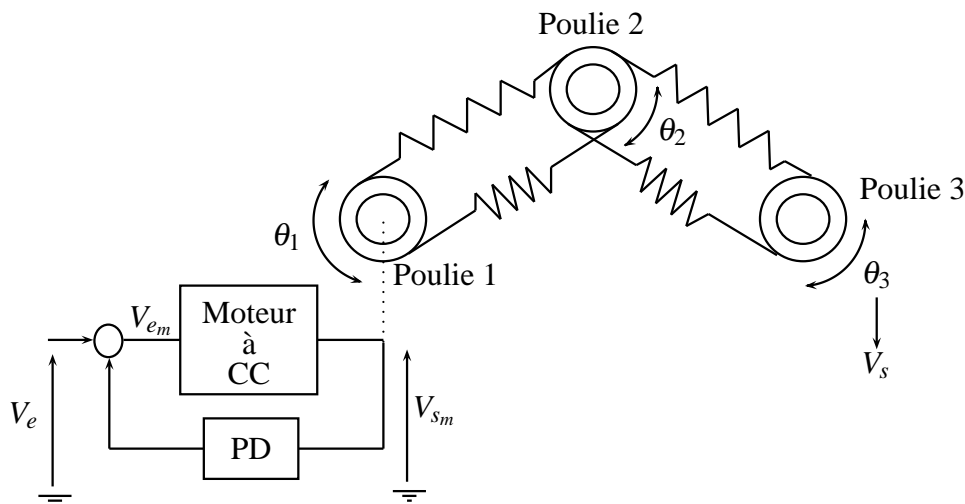


FIG. B.1: Schéma de principe du système de transmission flexible.

L'objectif de cet annexe consiste à justifier certains choix et certaines caractéristiques mises en évidence lors de l'identification de ce procédé au sein du paragraphe 6.2.1.

B.1 Étude du moteur à courant continu

Puisque le système considéré est commandé par un moteur à courant continu, il est nécessaire de s'intéresser, dans un premier temps, au modèle physique de ce dernier. Cet exercice est,

somme toute, très classique en automatique. Nous n'en rappellerons donc que les principales étapes.

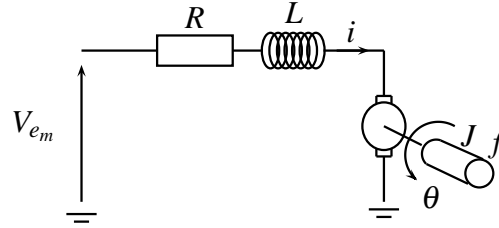


FIG. B.2: Moteur à courant continu.

Les équations fondamentales de l'électricité et de la mécanique nous permettent d'écrire que (cf. fig. B.2) :

$$V_{em}(t) = K_e \omega(t) + Ri(t) + L \frac{di(t)}{dt} \quad (\text{B.1})$$

$$C_m(t) = J \frac{d\omega(t)}{dt} + f\omega(t) \quad (\text{B.2})$$

$$C_m(t) = K_c i(t) \quad (\text{B.3})$$

$$\omega(t) = \frac{d\theta(t)}{dt} \quad (\text{B.4})$$

où R et L sont respectivement la résistance et l'inductance de l'induit, K_e la constante de force contre électromotrice, K_c la constante de couple, C_m le couple moteur, J la matrice d'inertie du l'ensemble axe moteur et f le coefficient de frottement. Le signal d'entrée est la tension de commande V_{em} . Le signal de sortie est la position angulaire de l'axe moteur θ . En passant dans l'espace de Laplace, en supposant que les conditions initiales soient nulles, nous montrons facilement que :

$$V_{em}(p) = K_e \Omega(p) + (R + Lp)I(p) \quad (\text{B.5})$$

$$C_m(p) = (Jp + f)\Omega(p) \quad (\text{B.6})$$

$$C_m(p) = K_c I(p) \quad (\text{B.7})$$

$$\Omega(p) = p\Theta(p) \quad (\text{B.8})$$

où $V_{em}(p)$, $\Omega(p)$, $I(p)$, $C_m(p)$ et $\Theta(p)$ sont les transformées de Laplace de V_{em} , ω , i , C_m et θ . La relation entre la tension de commande V_{em} et la position angulaire Θ s'écrit alors :

$$\frac{\Theta(p)}{V_{em}(p)} = \frac{K_c}{p(K_e K_c + (R + Lp)(f + Jp))}. \quad (\text{B.9})$$

En associant à cette fonction de transfert le gain d'amplificateur, le terme de réduction et le gain du convertisseur $rad/s \longleftrightarrow V$, l'expression finale devient :

$$\frac{V_{sm}(p)}{V_{em}(p)} = \frac{K}{p(K_e K_c + (R + Lp)(f + Jp))}. \quad (\text{B.10})$$

B.2 Étude mécanique du système de transmission flexible

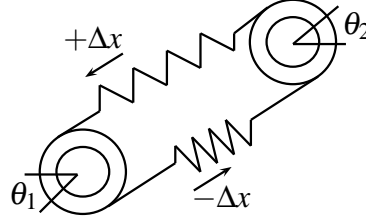


FIG. B.3: Système composé de deux poulies reliées par une transmission flexible.

Afin de calculer la fonction de transfert reliant la position angulaire de la poulie 1 à celle de la poulie 3 (cf. fig. B.1), il est nécessaire d'étudier séparément chaque poulie et de lui appliquer le principe fondamental de la dynamique. Considérons donc le système de transmission de la figure B.3. Il est alors aisé de vérifier que :

$$J_1 \frac{d^2 \theta_1}{dt^2} = C_m(t) - f_1 \frac{d\theta_1}{dt} - 2kr\Delta x \quad (\text{B.11})$$

où J_1 est la matrice d'inertie de l'ensemble « moteur-réducteur-poulie 1 », C_m le couple moteur, f_1 le coefficient de frottement visqueux de l'axe moteur, k la constante de raideur des ressorts et r le rayon des poulies. La relation entre l'allongement Δx et les angles θ_1 et θ_2 vaut quant à elle :

$$\Delta x = r(\theta_1 - \theta_2). \quad (\text{B.12})$$

Ainsi :

$$J_1 \frac{d^2 \theta_1}{dt^2} = C_m(t) - f_1 \frac{d\theta_1}{dt} - 2k(\theta_1 - \theta_2)r^2. \quad (\text{B.13})$$

De même, pour la poulie 2 :

$$J_2 \frac{d^2 \theta_2}{dt^2} = -f_2 \frac{d\theta_2}{dt} - 2k(\theta_2 - \theta_1)r^2 - 2k(\theta_2 - \theta_3)r^2 \quad (\text{B.14})$$

où J_2 et f_2 sont respectivement la matrice d'inertie de la poulie 2 et le coefficient de frottement visqueux de l'axe. Enfin, pour la dernière poulie :

$$J_3 \frac{d^2 \theta_3}{dt^2} = -f_3 \frac{d\theta_3}{dt} - 2k(\theta_3 - \theta_2)r^2, \quad (\text{B.15})$$

J_3 et f_3 étant respectivement la matrice d'inertie de la poulie 3 et le coefficient de frottement visqueux de l'axe. En passant dans l'espace de Laplace, en supposant, comme précédemment, que les conditions initiales soient nulles, nous avons :

$$(J_1 p^2 + F_1 p + 2kr^2) \Theta_1(p) = C_m(p) + 2kr^2 \Theta_2(p) \quad (\text{B.16})$$

$$(J_2 p^2 + F_2 p + 4kr^2) \Theta_2(p) = 2kr^2 \Theta_1(p) + 2kr^2 \Theta_3(p) \quad (\text{B.17})$$

$$(J_3 p^2 + F_3 p + 2kr^2) \Theta_3(p) = 2kr^2 \Theta_2(p). \quad (\text{B.18})$$

Après simplification, la relation entre les positions angulaires extrêmes Θ_1 et Θ_3 s'écrit :

$$\frac{\Theta_3(p)}{\Theta_1(p)} = \frac{1}{1 + \frac{F_2 + 2F_3}{2kr^2} p + \left(\frac{J_2 + 2J_3}{2kr^2} + \frac{F_2 F_3}{(2kr^2)^2} \right) p^2 + \frac{F_2 J_3 + F_3 J_2}{(2kr^2)^2} p^3 + \frac{J_2 J_3}{(2kr^2)^2} p^4}. \quad (\text{B.19})$$

Le processus de transmission flexible considéré est donc un système d'ordre 4.

A partir des équations (B.10) et (B.19), sachant que le régulateur utilisé pour asservir le moteur en position et en vitesse est de type PD, il est ensuite facile de montrer que l'ensemble « moteur-transmission flexible » possède un gain statique égale 1. Cette dernière remarque explique la valeur obtenue sur le tracé fréquentiel 6.35.

Annexe C

Abréviations, acronymes et notations

Sommaire

| | | |
|------------|----------------------------------|------------|
| C.1 | Abréviations et acronymes | 211 |
| C.2 | Symboles et notations | 212 |

C.1 Abréviations et acronymes

| | |
|------------------------|--|
| 4SID | Subspace State Space System IDentification |
| ARMAX | AutoRegressive Moving Average with eXogenous inputs |
| ARX | AutoRegressive model with eXogenous inputs |
| BB | Bruit Blanc |
| BC | Bruit Coloré |
| CEP | Condition d'Excitation Persistante |
| CCA | Canonical Correlation Analysis |
| COIVPAST | COrelated based Instrumental Variable version of Projection Approximation Subspace Tracking |
| COIVPASTd | COrelated based Instrumental Variable version of Projection Approximation Subspace Tracking with deflation |
| COIVPM | COrelated based Instrumental Variable version of the Propagator Method |
| COPAST | COrelated based Projection Approximation Subspace Tracking |
| COPASTd | COrelated based Projection Approximation Subspace Tracking with deflation |
| COPM | COrelated based Propagator Method |
| CVA | Canonical Variate Analysis |
| DDA | Détection des Directions d'Arrivée |
| DVS | Décomposition en Valeurs Singulières |
| EIVPAST | Extended Instrumental Variable Projection Approximation Subspace Tracking |
| EIVPM | Extended Instrumental Variable Propagator Method |
| EIV _{sqrt} PM | Extended Instrumental Variable square root Propagator Method |

| | |
|----------------------|---|
| ESPRIT | Estimation of Signal Parameters by Rotational Invariance Techniques |
| IVPT | Instrumental Variable Projector Tracking |
| MIMO | Multi-Input Multi-Output |
| MOESP | Mimo Output Error State sPace |
| MUSIC | MUltiple SIgnal Classification |
| N4SID | Numerical algorithm for Subspace State Space System IDentification |
| PAST | Projection Approximation Subspace Tracking |
| PI | Past Input |
| PM | Propagator Method |
| PO | Past Output |
| PO_EIV | Past Output Error In Variables |
| PT | Projector Tracking |
| ROM _{im} | Recursive Ordinary Moesp with inversion matrix lemma update |
| ROM _{qr} | Recursive Ordinary Moesp with qr factorization update |
| RPIPO _{im} | Recursive Past Input Past Output with inversion matrix lemma update |
| RPIPO _{qr} | Recursive Past Input Past Output with qr factorization update |
| RPM1 | Recursive propagator method 1 |
| RPM2 | Recursive propagator method 2 |
| SWEDE | Subspace method WithoutEigenDEcomposition |
| UDPAST | Update DwnDate Projection Approximation Subspace Tracking |
| Yang _{grad} | Yang criterion with gradient minimisation |

C.2 Symboles et notations

| | |
|---------------------------|--|
| n_u | nombre d'entrées du systèmes |
| n_y | nombre de sorties du systèmes |
| n_x | dimension du vecteur d'état |
| n_ξ | dimension de la variable instrumentale |
| n_s | nombre de sources émettrices |
| n_z | nombre de capteurs de l'antenne |
| f | indice futur |
| p | indice passé |
| \mathbb{R} | corps de réels |
| \mathbb{C} | corps des complexes |
| \mathbb{R}^l | espace vectoriel des vecteurs réels de dimension l |
| $\mathbb{R}^{l \times m}$ | espace vectoriel des matrices réelles de dimension $l \times m$ |
| $\mathbb{E}\{\cdot\}$ | espérance mathématique |
| \mathbf{R}_a | matrice de covariance du vecteur \mathbf{a} |
| \mathbf{R}_{ab} | matrice d'intercorrélation des vecteurs \mathbf{a} et \mathbf{b} |
| \mathbf{I}_n | matrice identité de dimension n |
| $\mathbf{M}(i, :)$ | i^{e} ligne de la matrice \mathbf{M} |
| $\mathbf{M}(:, j)$ | j^{e} colonne de la matrice \mathbf{M} |
| $\mathbf{M}(i : j, :)$ | lignes i à j de la matrice \mathbf{M} |
| $\mathbf{M}(:, i : j)$ | colonnes i à j de la matrice \mathbf{M} |
| \mathbf{M}^T | transposée de la matrice \mathbf{M} |

| | |
|--|--|
| \mathbf{M}^H | transposée hermitienne de la matrice \mathbf{M} |
| \mathbf{M}^{-1} | inverse de la matrice \mathbf{M} |
| $\mathbf{M}^{1/2}$ | racine carrée de la \mathbf{M} telle que $\mathbf{M}^{1/2}\mathbf{M}^{T/2} = \mathbf{M}$ |
| \mathbf{M}^\dagger | pseudo inverse de Moore Penrose de la matrice \mathbf{M} |
| \mathbf{M}^\perp | orthogonale de la matrice \mathbf{M} telle que $\mathbf{M}\mathbf{M}^\perp = \mathbf{0}$ |
| $\Pi_{\mathbf{M}}^\perp$ | projection orthogonale sur le noyau de \mathbf{M} |
| $\Pi_{\mathbf{M}}^{\parallel \mathbf{N}}$ | projection oblique sur l'espace ligne de la matrice \mathbf{M} parallèlement à l'espace ligne de la matrice \mathbf{N} |
| $\text{Tr}(\mathbf{M})$ | trace de la matrice \mathbf{M} |
| $\text{Com}(\mathbf{M})$ | comatrice de la matrice \mathbf{M} |
| $\text{dim}(\mathbf{M})$ | dimension de la matrice \mathbf{M} |
| $\text{rang}(\mathbf{M})$ | rang de la matrice \mathbf{M} |
| $\text{Im}_{\text{col}}(\mathbf{M})$ | espace vectoriel engendré par les colonnes de la matrice \mathbf{M} |
| $\text{Im}_{\text{lig}}(\mathbf{M})$ | espace vectoriel engendré par les lignes de la matrice \mathbf{M} |
| $\text{Ker}(\mathbf{M})$ | noyau de la matrice \mathbf{M} |
| $\mathbf{M} \otimes \mathbf{N}$ | produit de Kronecker des matrices \mathbf{M} et \mathbf{N} |
| $\text{vec}(\mathbf{M})$ | mise en colonnes de la matrice \mathbf{M} |
| $\ \mathbf{x}\ $ | norme du vecteur \mathbf{x} |
| $\ \mathbf{M}\ _F$ | norme de Frobenius de la matrice \mathbf{M} |
| $\ \mathbf{M}\ _2$ | norme 2 de la matrice \mathbf{M} |
| $\arg \min V$ | argument minimum du critère V |
| $\hat{\mathbf{x}}$ | estimée du vecteur \mathbf{x} |
| $[\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}]$ | matrices d'état du système |
| \mathbf{u} | vecteur d'entrée |
| $\tilde{\mathbf{u}}$ | vecteur d'entrée non bruitée |
| \mathbf{y} | vecteur de sortie |
| $\tilde{\mathbf{y}}$ | vecteur de sortie non bruitée |
| \mathbf{x} | vecteur d'état |
| \mathbf{v} | vecteur de bruit de mesure en entrée |
| \mathbf{v} | vecteur de bruit de mesure en sortie |
| \mathbf{w} | vecteur de bruit d'état |
| \mathbf{n} | vecteur de bruit |
| ξ | vecteur de variable instrumentale |
| \mathbf{z} | vecteur de sortie des capteur de l'antenne |
| \mathbf{s} | vecteur des sources émettrices |
| $\bar{\mathbf{h}}$ | réponse impulsionnelle |
| \mathcal{U}_j | matrice de commandabilité d'ordre j |
| Γ_i | matrice d'observabilité d'ordre i |
| Γ_i^\uparrow | matrice réduite obtenue en supprimant les n_y dernières lignes de Γ_i |
| Γ_i^\downarrow | matrice réduite obtenue en supprimant les n_y premières lignes de Γ_i |
| \mathbf{G}_i | matrice de Toeplitz d'ordre i liée au bruit d'état |
| \mathbf{H}_i | matrice de Toeplitz d'ordre i contenant les paramètres de Markov du système |
| Ξ | matrice de variable instrumentale |

| | |
|--|---|
| $\bar{\mathbf{R}}, \bar{\mathbf{Q}}$ | matrices résultantes de la décomposition QR |
| $\bar{\mathbf{U}}, \bar{\mathbf{\Sigma}}, \bar{\mathbf{V}}$ | matrices résultantes de la décomposition en valeurs singulières |
| $\underline{\mathbf{U}}, \underline{\mathbf{\Lambda}}, \underline{\mathbf{V}}$ | matrices résultantes de la décomposition en valeurs propres |
| $\bar{\sigma}$ | valeur singulière |
| $\underline{\lambda}$ | valeur propre |

Pour $\bar{t} = t + M - 1$

$$\mathbf{y}_f^+(t) = [\mathbf{y}^T(t) \ \cdots \ \mathbf{y}^T(t+f-1)]^T \in \mathbb{R}^{n_y f \times 1}$$

$$\mathbf{Y}_f^+(\bar{t}) = [\mathbf{y}_f^+(t) \ \cdots \ \mathbf{y}_f^+(\bar{t})] \in \mathbb{R}^{n_y f \times M}$$

$$\mathbf{y}_p^-(t) = [\mathbf{y}^T(t-p) \ \cdots \ \mathbf{y}^T(t-1)]^T \in \mathbb{R}^{n_y p \times 1}$$

$$\mathbf{Y}_p^-(\bar{t}) = [\mathbf{y}_p^-(t) \ \cdots \ \mathbf{y}_p^-(\bar{t})] \in \mathbb{R}^{n_y p \times M}$$

$$\mathbf{u}_f^+(t) = [\mathbf{u}^T(t) \ \cdots \ \mathbf{u}^T(t+f-1)]^T \in \mathbb{R}^{n_u f \times 1}$$

$$\mathbf{U}_f^+(\bar{t}) = [\mathbf{u}_f^+(t) \ \cdots \ \mathbf{u}_f^+(\bar{t})] \in \mathbb{R}^{n_u f \times M}$$

$$\mathbf{u}_p^-(t) = [\mathbf{u}^T(t-p) \ \cdots \ \mathbf{u}^T(t-1)]^T \in \mathbb{R}^{n_u p \times 1}$$

$$\mathbf{U}_p^-(\bar{t}) = [\mathbf{u}_p^-(t) \ \cdots \ \mathbf{u}_p^-(\bar{t})] \in \mathbb{R}^{n_u p \times M}$$

$$\mathbf{v}_f^+(t) = [\mathbf{v}^T(t) \ \cdots \ \mathbf{v}^T(t+f-1)]^T \in \mathbb{R}^{n_v f \times 1}$$

$$\mathbf{V}_f^+(\bar{t}) = [\mathbf{v}_f^+(t) \ \cdots \ \mathbf{v}_f^+(\bar{t})] \in \mathbb{R}^{n_v f \times M}$$

$$\mathbf{v}_p^-(t) = [\mathbf{v}^T(t-p) \ \cdots \ \mathbf{v}^T(t-1)]^T \in \mathbb{R}^{n_v p \times 1}$$

$$\mathbf{V}_p^-(\bar{t}) = [\mathbf{v}_p^-(t) \ \cdots \ \mathbf{v}_p^-(\bar{t})] \in \mathbb{R}^{n_v p \times M}$$

$$\mathbf{v}_f^+(t) = [\mathbf{v}^T(t) \ \cdots \ \mathbf{v}^T(t+f-1)]^T \in \mathbb{R}^{n_y f \times 1}$$

$$\mathbf{V}_f^+(\bar{t}) = [\mathbf{v}_f^+(t) \ \cdots \ \mathbf{v}_f^+(\bar{t})] \in \mathbb{R}^{n_y f \times M}$$

$$\mathbf{v}_p^-(t) = [\mathbf{v}^T(t-p) \ \cdots \ \mathbf{v}^T(t-1)]^T \in \mathbb{R}^{n_y p \times 1}$$

$$\mathbf{V}_p^-(\bar{t}) = [\mathbf{v}_p^-(t) \ \cdots \ \mathbf{v}_p^-(\bar{t})] \in \mathbb{R}^{n_y p \times M}$$

$$\mathbf{w}_f^+(t) = [\mathbf{w}^T(t) \ \cdots \ \mathbf{w}^T(t+f-1)]^T \in \mathbb{R}^{n_x f \times 1}$$

$$\mathbf{W}_f^+(\bar{t}) = [\mathbf{w}_f^+(t) \ \cdots \ \mathbf{w}_f^+(\bar{t})] \in \mathbb{R}^{n_x f \times M}$$

$$\mathbf{w}_p^-(t) = [\mathbf{w}^T(t-p) \ \cdots \ \mathbf{w}^T(t-1)]^T \in \mathbb{R}^{n_x p \times 1}$$

$$\mathbf{W}_p^-(\bar{t}) = [\mathbf{w}_p^-(t) \ \cdots \ \mathbf{w}_p^-(\bar{t})] \in \mathbb{R}^{n_x p \times M}$$

$$\mathbf{n}_f^+(t) = [\mathbf{n}^T(t) \ \cdots \ \mathbf{n}^T(t+f-1)]^T \in \mathbb{R}^{n_y f \times 1}$$

$$\mathbf{N}_f^+(\bar{t}) = [\mathbf{n}_f^+(t) \ \cdots \ \mathbf{n}_f^+(\bar{t})] \in \mathbb{R}^{n_y f \times M}$$

$$\mathbf{n}_p^-(t) = [\mathbf{n}^T(t-p) \ \cdots \ \mathbf{n}^T(t-1)]^T \in \mathbb{R}^{n_y p \times 1}$$

$$\mathbf{N}_p^-(\bar{t}) = [\mathbf{n}_p^-(t) \ \cdots \ \mathbf{n}_p^-(\bar{t})] \in \mathbb{R}^{n_y p \times M}$$

$$\mathbf{X}^-(\bar{t}) = [\mathbf{x}(t-p) \ \cdots \ \mathbf{x}(t-p)] \in \mathbb{R}^{n_x \times M}$$

$$\mathbf{X}^+(\bar{t}) = [\mathbf{x}(t) \ \cdots \ \mathbf{x}(\bar{t})] \in \mathbb{R}^{n_x \times M}$$

Annexe D

Bibliographie personnelle

G. Mercère, S. Lecœuche and C. Vasseur. Output noise insensitive identification of systems : a new recursive state-space method. *The 3rd workshop on Physics in Signal and Image Processing*. Grenoble, France. January 2003.

G. Mercère and S. Lecœuche. Identification en ligne de systèmes bruités : une nouvelle méthode récursive des sous-espaces. *Les Journées Doctorales d'Automatique. Valenciennes*, France. June 2003.

G. Mercère, S. Lecœuche and C. Vasseur. A new recursive method for subspace identification of noisy systems : EIVPM. *The 13th IFAC Symposium on System Identification*. Rotterdam, The Netherlands. August 2003.

S. Lecœuche, C. Lurette, G. Mercère, S. Lalot and C. Vasseur. Surveillance et diagnostic d'évolutions temporelles par classifieur autoadaptatif. *Le 14^e Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle*. Toulouse, France. January 2004.

G. Mercère, S. Lecœuche and M. Lovera. Recursive subspace identification based on instrumental variable unconstrained quadratic optimization. *International Journal of Adaptive Control and Signal Processing*. Vol. 18, pp 771–797. 2004.

G. Mercère, S. Lecœuche and C. Vasseur. Adaptation robuste de la méthode du propagateur à l'identification récursive des sous-espaces. *La Conférence Internationale Francophone d'Automatique*. Douz, Tunisia. November 2004.

S. Lecœuche, G. Mercère and S. Lalot. Evaluating time-dependent heat fluxes using artificial neural networks. *Inverse Problems in Science and Engineering*. Accepted, to appear.

Bibliographie

- [Abi88] H. Abidin-Tjokronegoro. *Estimation adaptative des paramètres de systèmes variant dans le temps et détection de ruptures*. PhD thesis, Institut National Polytechnique de Grenoble, Grenoble, France, 1988.
- [ÅE71] K. J. Åström and P. Eykhoff. System identification : a survey. *Automatica*, 7 :123–167, 1971.
- [ÅHHH93] K. J. Åström, T. Hägglund, C. C. Hang, and W. K. Ho. Automatic tuning and adaptation for PID controllers : a survey. *Control Engineering Practice*, 1 :699–714, 1993.
- [Aka74] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19 :716–723, 1974.
- [AL97] R. Amirthalingam and J. H. Lee. Subspace identification based inferential control of a continuous pulp digester. *Computers Chemical Engineering*, 21 :1143–1148, 1997.
- [AM90] B. Anderson and J. Moore. *Optimal control : linear quadratic methods*. Prentice Hall, Englewood Cliffs, 1990.
- [And85] B. D. O. Andersson. Identification of scalar errors in variables models with dynamics. *Automatica*, 21 :709–716, 1985.
- [AVVD98] M. Abdelghani, M. Verhaegen, P. Van Overschee, and B. De Moor. Comparison study of subspace identification methods applied to flexible structures. *Mechanical Systems and Signal Processing*, 12 :679–692, 1998.
- [ÅW95] K. J. Åström and B. Wittenmark. *Adaptive control*. Addison Wesley, 2nd edition, 1995.
- [BA92] S Bannour and M. R. Azimi-Sadjadi. An adaptive approach for optimal data reduction using recursive least squares learning methods. In *The IEEE International Conference on Acoustics, Speech and Signal Processing*, San Fransisco, California, USA, March 1992.
- [BAB00] M. Basseville, M. Abdelghani, and A. Benveniste. Subspace based fault detection algorithms for vibration monitoring. *Automatica*, 36 :101–109, 2000.
- [Bas97] T. Bastogne. *Identification des systèmes multivariables par les méthodes des sous espaces. Application à un système d’entraînement de bande*. PhD thesis, Université Henri Poincaré, Nancy, France, 1997.
- [Bau98] D. Bauer. *Some asymptotic theory for the estimation of linear systems using maximum likelihood methods or subspace algorithms*. PhD thesis, Vienna University of Technology, Vienna, Austria, 1998.

- [Bau99] D. Bauer. Order estimation in the context of MOESP subspace identification methods. In *The 5th European Control Conference*, Karlsruhe, Germany, August 1999.
- [Bau01] D. Bauer. Order estimation for subspace methods. *Automatica*, 37 :1561–1573, 2001.
- [Bau03a] D. Bauer. Choosing integer parameters in subspace methods : a survey on asymptotic results. In *The 13th IFAC Symposium on System Identification*, Rotterdam, The Netherlands, August 2003.
- [Bau03b] D. Bauer. Subspace algorithms. In *The 13th IFAC Symposium on System Identification*, Rotterdam, The Netherlands, August 2003.
- [Bau04] D. Bauer. Asymptotic properties of subspace estimators. Submitted to *Automatica*, 2004.
- [BDR⁺92] P. Borne, G. Dauphin Tanguy, J. P. Richard, F. Rotella, and I. Zambettakis. *Modélisation et identification des processus*. Technip, 1992.
- [BDS99] D. Bauer, M. Deistler, and W. Scherrer. Consistency and asymptotic normality of some subspace algorithms for systems without observed inputs. *Automatica*, 35 :1243–1254, 1999.
- [BGRS97] S. Bittanti, E. Gatti, G. Ripamonti, and S. Savaresi. An application of subspace based techniques to nuclear spectroscopy. In *The 11th IFAC Symposium on System Identification*, Fukuoka, Japan, September 1997.
- [Bil97] P. Bildstein. Synthèse et réalisation de filtres actifs. In *Traité d'électronique. Techniques de l'Ingénieur*, 1997.
- [BJ00] D. Bauer and M. Jansson. Analysis of the asymptotic properties of the MOESP type of subspace algorithms. *Automatica*, 36 :497–509, 2000.
- [BK79] G. Bienvenu and L. Kopp. Adaptivity to background noise spatial coherence for high resolution passive methods. In *International Conference on Acoustics, Speech and Signal Processing*, Washington, disctrict of Columbia, USA, April 1979.
- [BK83] G. Bienvenu and L. Kopp. Optimality of high resolution array processing using the eigensystem approach. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 31 :1235–1247, 1983.
- [BNSR98] T. Bastogne, H. Noura, P. Sibille, and A. Richard. Multivariable identification of a winding process by subspace methods for tension control. *Control Engineering Practice*, 6 :1077–1088, 1998.
- [Böh95] J. F. Böhme. Statistical array signal processing of measured sonar and seismic data. In *The SPIE proceedings : 2563 Advanced Signal Processing Algorithms*, San Diego, California, USA, July 1995.
- [BPP96] I. Branica, N. Peric, and I. Petrovic. Comparison of several recursive identification methods. *Automatika*, 3 :99–104, 1996.
- [BR02] T. Bastogne and A. Richard. Modélisation et simulation d'un système d'entraînement de bande en présence de glissement. In *Conférence Internationale Francophone d'Automatique*, Nantes, France, July 2002.

- [Bro91] W. L. Brogan. *Modern control theory*. Prentice Hall, Englewood Cliffs, 3rd edition, 1991.
- [BRS98] T. Bastogne, A. Richard, and P. Sibille. Identification des systèmes multivariables : méthodes des sous-espaces. partie 2 : applicabilité et intérêt. *Journal Européen des Systèmes Automatisés*, 32 :235–365, 1998.
- [BSM03] B. Bhikkaji, T. Söderström, and K. Mahata. A recursive algorithm for estimating parameters in a one dimensional diffusion system. In *The 13th IFAC Symposium on System Identification*, Rotterdam, The Netherlands, August 2003.
- [BSR98] T. Bastogne, P. Sibille, and A. Richard. Identification des systèmes multivariables : méthodes des sous-espaces. partie 1 : état de l’art. *Journal Européen des Systèmes Automatisés*, 32 :207–234, 1998.
- [BYR98] K.V.S. Babu, Y. Yoganandam, and V.U. Reddy. Adaptive estimation of eigensubspace and tracking the directions of arrival. *Signal Processing*, 68 :317–339, 1998.
- [Cad90] J. A. Cadzow. Multiple source location : the signal subspace approach. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38 :1110–1125, 1990.
- [Cai76] P. E. Caines. On the asymptotic normality of instrumental variable and least squares estimators. *IEEE Transactions on Automatic Control*, 21 :598–600, 1976.
- [CG90] P. Comon and G. H. Golub. Tracking a few extreme singular values and vectors in signal processing. *Proceedings of the IEEE*, 78 :1327–1343, 1990.
- [Che02] T. Y. Chen. Application of adaptive predictive control to a floor heating system with a large thermal lag. *Energy and Buildings*, 34 :45–51, 2002.
- [Chi97] N. Chiu. *Subspace methods and informative experiments for system identification*. PhD thesis, University of Cambridge, Cambridge, United Kingdom, 1997.
- [CK95] Y. Cho and T. Kailath. Fast subspace based system identification : an instrumental variable approach. *Automatica*, 31 :903–905, 1995.
- [Cla67] D. W. Clarke. Generalized least squares estimation of parameters of a dynamic model. In *The first IFAC Symposium on Identification in Automatic Control Systems*, Prague, Czech Republic, June 1967.
- [CP04] A. Chiuso and G. Picci. The asymptotic variance of subspace estimates. *Journal of Econometrics*, 118 :257–291, 2004.
- [CR01] F. J. Carrillo and F. Rotella. Identification en temps réel des paramètres d’un système continu pour l’estimation d’usure d’outils de coupe. In *Journées Identification et Modélisation Expérimentale*, Vandoeuvre-les-Nancy, France, March 2001.
- [Cra46] H. Cramér. *Mathematical methods of statistics*. Princeton University Press, 1946.
- [CV97] C. T. Chou and M. Verhaegen. Subspace algorithms for the identification of multivariable dynamic errors-in-variables models. *Automatica*, 33 :1857–1869, 1997.
- [CXK94] Y. M. Cho, G. Xu, and T. Kailath. Fast recursive identification of state space models via exploitation displacement structure. *Automatica*, 30 :45–60, 1994.
- [CYH02] J. C. Chen, K. Yao, and R. E. Hudson. Source localization and beamforming. *IEEE Signal Processing Magazine*, 19 :30–40, 2002.

- [DAC95] P. Dorato, C. Abdallah, and V. Cerone. *Linear quadratic control : an introduction*. Prentice Hall, 1995.
- [DAD94] E.M. Dowling, L.P. Ammann, and R.D. DeGroat. A TQR iteration based adaptive SVD for real time angle and frequency tracking. *IEEE Transactions on Signal Processing*, 42 :914–926, 1994.
- [DBL⁺01] B. Dubuisson, B. O. Bouamama, R. Litwak, D. Maquin, I. Nikiforov, J. Ragot, M. Staroswiecki, and G. Zwingelstein. *Automatique et statistiques pour le diagnostic*, volume Systèmes automatisés, Information Commande Communication. Hermès, 2001.
- [DD02] K. De Cock and B. De Moor. Subspace angles for fault detection. In *The International Conference on Noise and Vibration Engineering*, Leuven, Belgium, September 2002.
- [DDL95] E.M. Dowling, R.D. DeGroat, D.A. Linebarger, and H. Ye. *SVD and Signal Processing III*, chapter Sphericalized SVD updating for subspace tracking, pages 227–234. Elsevier Publishing Company, 1995.
- [De 88] B. De Moor. *Mathematical concepts and techniques for modelling of static and dynamic systems*. PhD thesis, Katholieke Universiteit Leuven, Faculteit der Toegepaste Wetenschappen, Leuven, Belgium, 1988.
- [De 02] K. De Cock. *Principal angles in system theory, information theory and signal processing*. PhD thesis, Katholieke Universiteit Leuven, Faculteit der Toegepaste Wetenschappen, Leuven, Belgium, 2002.
- [DeG92] R. D. DeGroat. Noniterative subspace tracking. *IEEE Transactions on Signal Processing*, 40 :571–577, 1992.
- [DPS95] M. Deistler, K. Peternell, and W. Scherrer. Consistency and relative efficiency of subspace methods. *Automatica*, 31 :1865–1875, 1995.
- [DPV⁺02] K. De Cock, B. Peters, A. Vecchio, H. Van der Auweraer, and B. De Moor. Subspace system identification for mechanical engineering. In *The International Conference on Noise and Vibration Engineering*, Leuven, Belgium, September 2002.
- [DR90] R. D. DeGroat and R. A. Roberts. Efficient, numerically stabilized rank one eigenstructure updating. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38 :301–316, 1990.
- [ESS94] A. Eriksson, P. Stoica, and T. Söderström. On line subspace algorithms for tracking moving sources. *IEEE Transactions on Signal Processing*, 42 :2319–2330, 1994.
- [Far92] A. Farina. *Antenna based signal processing techniques for radar systems*. Artech House, Norwood, 1992.
- [FDV⁺98] W. Favoreel, B. De Moor, S. Van Huffel, V. Sima, and M. Verhaegen. Benchmark for subspace system identification algorithms. Technical report, Katholieke Universiteit Leuven, Leuven, Belgium, 1998.
- [FDV00] W. Favoreel, B. De Moor, and P. Van Overschee. Subspace state space system identification for industrial processes. *Journal of Process Control*, 10 :149–155, 2000.

- [FK94] A. J. Fenn and G. A. King. Adaptive radio frequency hyperthermia based array system for improved cancer therapy : phantom target measurements. *International Journal of Hyperthermia*, 10 :189–208, 1994.
- [Fos86] L. V. Foster. Rank and null space calculations using matrix decomposition without column interchanges. *Linear Algebra and its Applications*, 74 :47–71, 1986.
- [Fri84] B. Friedlander. The overdetermined recursive instrumental variable method. *IEEE Transactions on Automatic Control*, 4 :353–356, 1984.
- [Fuc92] J. J. Fuchs. Estimation of the number of signals in the presence of unknown correlated sensor noise. *IEEE Transactions on Signal Processing*, 40 :1053–1061, 1992.
- [FV98] T. W. Flint and R. J. Vaccaro. Performance analysis of N4SID state space system identification. In *The American Control Conference*, Philadelphia, Pennsylvania, USA, June 1998.
- [GBT99] G. O. Glentis, K. Berberidi, and S. Theodoridis. Efficient least squares adaptive algorithms for FIR transversal filtering. *IEEE Signal Processing Magazine*, 16 :13–41, 1999.
- [Ger98] J. J. Gertler. *Fault detection and diagnosis in engineering systems*. Marcel Dekker, 1998.
- [Giv58] W. Givens. Computation of plane unitary rotations transforming a general matrix to triangular form. *SIAM*, 6 :26–50, 1958.
- [GL95] L. Guo and L. Ljung. Performance analysis of general tracking algorithms. *IEEE Transactions on Automatic Control*, 40 :1388–1402, 1995.
- [GLV98] T. Gustafsson, M. Lovera, and M. Verhaegen. A novel algorithm for recursive instrumental variable based subspace identification. In *The 37th IEEE Conference on Decision and Control*, Tampa, Florida, USA, December 1998.
- [GS84] G. Goodwin and K. S. Sin. *Adaptive filtering prediction and control*. Prentice Hall, 1984.
- [GSMR96] H. Garnier, P. Sibille, M. Mensler, and A. Richard. Pilot crane identification and control in presence of friction. In *The 13th IFAC World Congress*, San Fransisco, California, USA, July 1996.
- [Gus96] T. Gustafsson. Instrumental variable subspace tracking using projection approximation. <http://citeseer.ist.psu.edu/gustafsson96instrumental.html>, 1996.
- [Gus97a] T. Gustafsson. Recursive system identification using instrumental variable subspace tracking. In *The 11th IFAC Symposium on System Identification*, Fukuoka, Japan, September 1997.
- [Gus97b] T. Gustafsson. System identification using subspace based instrumental variable methods. In *The 11th IFAC Symposium on System Identification*, Fukuoka, Japan, September 1997.
- [Gus98] T. Gustafsson. Instrumental variable subspace tracking using projection approximation. *IEEE Transactions on Signal Processing*, 46 :669–681, 1998.

- [Gus99] T. Gustafsson. *Subspace methods for system identification and signal processing*. PhD thesis, Department of Signals and Systems, Chalmers University of Technology, Göteborg, Sweden, 1999.
- [Gus00] F. Gustafsson. *Adaptive filtering and change detection*. John Wiley & Sons, Ltd, 2000.
- [Gus01] T. Gustafsson. Subspace identification using instrumental variable techniques. *Automatica*, 37 :2005–2010, 2001.
- [GV96] G. H. Golub and C. F. Van Loan. *Matrix computations*. John Hopkins University Press, Baltimore MD, 3rd edition, 1996.
- [HK66] B. L. Ho and R. E. Kalman. Effective construction of linear state variable models from input output data. *Regelungstechnik*, 14 :545–592, 1966.
- [HK86] I. W. Hunter and M. J. Korenberg. The identification of nonlinear biological systems : Wiener and Hammerstein cascade models. *Biological Cybernetics*, 55 :135–144, 1986.
- [HV97] B. Havekamp and M. Verhaegen. SMI Toolbox : state space model identification software for multivariable dynamical systems v. 1.0. Delft, the netherlands, Delft University of Technology, 1997.
- [Ise80] R. Isermann. Practical aspects of process identification. *Automatica*, 16 :575–587, 1980.
- [Jan97] M. Jansson. *On subspace methods in system identification and sensor array signal processing*. PhD thesis, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, 1997.
- [JD93] D. H. Johnson and D. E. Dudgeon. *Array signal processing : concepts and techniques*. Prentice Hall, Engelwood Cliffs, 1993.
- [JKDW01] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter. *Applied interval analysis with examples in parameter and state estimation, robust control and robotics*. Springer Verlag, 2001.
- [JSL01] B. C. Juricek, D. E. Seborg, and W. E. Larimore. Identification of the tennessee eastman challenge process with subspace methods. *Control Engineering Practice*, 9 :1337–1351, 2001.
- [JW98] M. Jansson and B. Wahlberg. On consistency of subspace methods for system identification. *Automatica*, 34 :1507–1519, 1998.
- [Kai80] T. Kailath. *Linear Systems*. Prentice Hall, Engelwood Cliffs, 1980.
- [Kar84] J. Karhunen. Adaptive algorithms for estimating eigenvectors of correlation type matrices. In *The IEEE International Conference on Acoustics, Speech and Signal Processing*, San Diego, California, USA, March 1984.
- [Kar86] I. Karasalo. Estimating the covariance matrix by signal subspace averaging. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 34 :8–12, 1986.
- [KB61] R. E. Kalman and R. S. Bucy. New results in linear filtering and prediction theory. *Journal of Basic Engineering*, 83 :95–108, 1961.

- [KB86] M. Kaveh and A. J. Barabell. The statistical performance of the MUSIC and the Minimum Norm algorithms in resolving plane waves in noise. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 34 :331–341, 1986.
- [KKKN01] S. Kim, Y. Kim, H. Kim, and C. Nam. Adaptive reconfigurable flight control system based on recursive system identification. In *The 15th International Sessions in the 39th Aircraft Symposium*, Gifu, Japan, October 2001.
- [KS61] M. G. Kendall and A. Stuart. *The advanced theory of statistics*, volume 2. Griffin, London, 1961.
- [KT83] R. Kumaresan and D. W. Tufts. Estimating the angles of arrival of multiple plane waves. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 19 :134–139, 1983.
- [Kun78] S. Y. Kung. A new identification and model reduction algorithm via singular value decomposition. In *The 12th Asilomar Conference on Circuits, Systems and Computer*, Pacific Grove, California, USA, November 1978.
- [KV96] H. Krim and M. Viberg. Two decades of array signal processing research : the parametric approach. *IEEE Signal Processing Magazine*, 13 :67–94, 1996.
- [Lan50] C. Lanczos. An iterative method for the solution of the eigenvalue problem of linear differential and integral operators. *Journal of Research of the National Bureau of Standards*, 45 :225–280, 1950.
- [Lan74] I. D. Landau. A survey of model reference adaptive techniques : theory and applications. *Automatica*, 10 :353–379, 1974.
- [Lar83] W. E. Larimore. System identification, reduced order filters and modeling via canonical variate analysis. In *The American Control Conference*, San Francisco, California, USA, June 1983.
- [Lar90] W. E. Larimore. Canonical variate analysis in identification, filtering and adaptive control. In *The 29th IEEE Conference on Decision and Control*, Honolulu, Hawaii, USA, December 1990.
- [LBB⁺01] I. D. Landau, A. Besançon-Voda, G. Besançon, C. Durieu, A. Karimi, I. D. Landau, M. Namar, T. Poinot, L. Pronzato, H. F. Raynaud, J. Richalet, J. C. Trigeassou, E. Walter, and A. Zolghardi. *Identification des systèmes*, volume Systèmes automatisés, Information Commande Communication. Hermès, 2001.
- [LFB98] M. Lovera, A. Falcetti, and S. Bittanti. On the estimation of the A matrix in subspace model identification. In *The Conference on the Mathematical Theory of Networks and Systems*, Padova, Italy, July 1998.
- [LGV00] M. Lovera, T. Gustafsson, and M. Verhaegen. Recursive subspace identification of linear and non linear Wiener state space models. *Automatica*, 36 :1639–1650, 2000.
- [Liu92] K. Liu. Identification of multi input and multi output systems by observability range space extraction. In *The 31st IEEE Conference on Decision and Control*, Tucson, Arizona, USA, December 1992.
- [Lju81] L. Ljung. Analysis of a general recursive prediction error identification algorithm. *Automatica*, 17 :89–100, 1981.

- [Lju96] L. Ljung. Development of system identification. In *The 13th IFAC World Congress*, San Francisco, California, July 1996.
- [Lju99] L. Ljung. *System identification. Theory for the user*. PTR Prentice Hall Information and System Sciences Series. T. Kailath, Series Editor, Upper Saddle River, 2nd edition edition, 1999.
- [Lju02] L. Ljung. Recursive identification algorithms. *Circuits, Systems and Signal Processing*, 21 :57–68, 2002.
- [LLM97] I. Landau, R. Lozano, and M. M'Saad. *Adaptive Control*. Springer Verlag, New York, 1997.
- [LM96] L. Ljung and T. McKelvey. Subspace identification from closed loop data. *Signal Processing*, 52 :209–215, 1996.
- [Lov98] M. Lovera. *Subspace identification methods : theory and applications*. PhD thesis, Dipartimento di Elettronica e Informazione, Politecnico di Milano, Milan, Italia, 1998.
- [Lov03] M. Lovera. Recursive subspace identification based on projector tracking. In *The 13th IFAC Symposium on System Identification*, Rotterdam, The Netherlands, August 2003.
- [LPV01] M. Lovera, T. Parisini, and M. Verhaegen. Fault detection : a subspace identification approach. In *The 40th IEEE Conference on Decision and Control*, Orlando, Florida, USA, December 2001.
- [LQL04] W. Lin, S. J. Qin, and L. Ljung. A framework for closed loop subspace identification with innovation estimation. Technical report, Texas Wisconsin Modeling and Control Consortium, Wisconsin, USA, 2004.
- [LS83] L. Ljung and T. Söderström. *Theory and practice of recursive identification*. The MIT Press, Cambridge, 1983.
- [LS92] K. Liu and R. E. Skelton. Identification of linear systems from their pulse responses. In *The American Control Conference*, Chicago, Illinois, USA, June 1992.
- [LT01] R. Lozano and D. Taoutaou. *Identification et commande adaptative*, volume Systèmes automatisés, Information Commande Communication. Hermès, 2001.
- [LV98] M. Lovera and M. Verhaegen. Recursive subspace identification of linear and non linear Wiener type models. In *The IEEE Conference on Computability and Complexity in Analysis*, Trieste, Italy, September 1998.
- [LVT89] F. Li, J. Vaccaro, and D. W. Tufts. Min norm linear prediction for arbitrary sensor arrays. In *The IEEE International Conference on Acoustics, Speech and Signal Processing*, Glasgow, United Kingdom, May 1989.
- [May67] D. Q. Mayne. A method for estimating discret time transfer functions. In *The 2nd United Kingdom Automatic Control Convention*, Bristol, United Kingdom, 1967.
- [MBS96] S. Marcos, M. Benidir, and J. Sanchez-Araujo. An adaptive tracking algorithm for direction finding and array shape estimation in a nonstationary environment. *Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology*, 14 :107–118, 1996.

- [MC00] M. M'Saad and J. Chebassier. Commande adaptative des systèmes. In *Traité d'informatique industrielle*, pages S7-426-1-S7-426-25. Techniques de l'Ingénieur, 2000.
- [McK95] T. McKelvey. *Identification of state space models from time and frequency data*. PhD thesis, Department of Electrical Engineering, Linköping University, Linköping, Sweden, 1995.
- [MD91] J. Munier and G. Y. Delisle. Spatial analysis using new properties of the cross spectral matrix. *IEEE Transactions on Signal Processing*, 39 :746-749, 1991.
- [MDVV89] M. Moonen, B. De Moor, L. Vandenberghe, and J. Vandewalle. On and off line identification of linear state space models. *International Journal of Control*, 49 :219-232, 1989.
- [Mer01] G. Mercère. Méthodologies d'identification pour la commande prédictive des procédés industriels, 2001. Rapport de DEA. ENSI de Caen, Laboratoire d'Automatique et des Procédés.
- [MJD88] J. Munier, G. Jourdain, and G. Y. Delisle. A new algorithm for the identification of distorted wavefronts. In *The NATO Advanced Study Institute*, Ontario, Canada, July 1988.
- [ML03] G. Mercère and S. Lecœuche. Identification en ligne de systèmes bruités : une nouvelle méthode récursive des sous-espaces. In *Les Journées Doctorales d'Automatique*, Valenciennes, France, June 2003.
- [MLL04] G. Mercère, S. Lecœuche, and M. Lovera. Recursive subspace identification based on instrumental variable unconstrained quadratic optimization. *International Journal of Adaptive Control and Signal Processing*, 18 :771-797, 2004.
- [MLV03a] G. Mercère, S. Lecœuche, and C. Vasseur. A new recursive method for subspace identification of noisy systems : EIVPM. In *The 13th IFAC Symposium on System Identification*, Rotterdam, The Netherlands, August 2003.
- [MLV03b] G. Mercère, S. Lecœuche, and C. Vasseur. Output noise insensitive identification of systems : a new recursive state-space method. In *The 3rd workshop on Physics in Signal and Image Processing*, Grenoble, France, January 2003.
- [MLV04] G. Mercère, S. Lecœuche, and C. Vasseur. Adaptation robuste de la méthode du propagateur à l'identification récursive des sous-espaces. In *La Conférence Internationale Francophone d'Automatique*, Douz, Tunisia, November 2004.
- [MMB95] S. Marcos, A. Marsal, and M. Benidir. The propagator method for source bearing estimation. *Signal Processing*, 42 :121-138, 1995.
- [Muc84] R. A. Mucci. A comparison of efficient beamforming algorithms. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 32 :548-558, 1984.
- [MV96] C. S. MacInnes and R. J. Vaccaro. Tracking directions of arrival with invariant subspace updating. *Signal Processing*, 50 :137-150, 1996.
- [Nai00] P. S. Naidu. *Sensor array signal processing*. CRC Press, 2000.
- [ÖG02] M. Östring and S. Gunnarsson. Recursive identification of physical parameters in a flexible robot arm. Technical report, Department of Electrical Engineering, Linköping University, Linköping, Sweden, 2002.

- [OK99] H. Oku and H. Kimura. A recursive 4SID from the input output point of view. *Asian Journal of control*, 1 :258–269, 1999.
- [OK02] H. Oku and H. Kimura. Recursive 4SID algorithms using gradient type subspace tracking. *Automatica*, 38 :1035–1043, 2002.
- [Oku00] H. Oku. *Sequential subspace state space system identification and state estimation of unknown multivariable systems*. PhD thesis, Department of Mathematical Engineering, Information and Physics, University of Tokyo, Tokyo, Japan, 2000.
- [Oku03] H. Oku. Application of a recursive subspace identification algorithm to change detection. In *The 13th IFAC Symposium on System Identification*, Rotterdam, The Netherlands, August 2003.
- [ONVV01] H. Oku, G. Nijse, M. Verhaegen, and V. Verdult. Change detection in the dynamics with recursive subspace identification. In *The 40th IEEE Conference on Decision and Control*, Orlando, Florida, USA, December 2001.
- [Öst02] M. Östring. Identification, diagnosis and control of a flexible robot arm. Master's thesis, Department of Electrical Engineering, Linköping University, Linköping, Sweden, 2002.
- [OV94] B. Ottersten and M. Viberg. A subspace based instrumental variable method for state space system identification. In *The 10th IFAC Symposium on System Identification*, Copenhagen, Denmark, July 1994.
- [OVK91] B. Ottersten, M. Viberg, and T. Kailath. Performance analysis of the total least squares ESPRIT algorithm. *IEEE Transactions on Signal Processing*, 39 :1122–1135, 1991.
- [OVK92] B. Ottersten, M. Viberg, and T. Kailath. Analysis of subspace fitting and ML techniques for parameter estimation from sensor array data. *IEEE Transactions on Signal Processing*, 40 :590–600, 1992.
- [OVSN93] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai. Exact and large sample ML techniques for parameter estimation and detection in array processing. In *Radar Array Processing*, pages 99–151. Springer Verlag, haykin, litva and sherperd editors edition, 1993.
- [PC98] M. Prandini and M.C. Campi. A new recursive identification algorithm for singularity free adaptive control. *Systems and Control Letters*, 34 :177–183, 1998.
- [Pet95] K. Peternell. *Identification of linear dynamic systems by subspace and realization-based algorithms*. PhD thesis, Technical University of Vienna, Vienna, Austria, 1995.
- [PF89] B. Porat and B. Friedlander. The square root overdetermined recursive instrumental variable algorithm. *IEEE Transactions on Automatic Control*, 34 :656–658, 1989.
- [PGMR03] K. M. Pekpe, K. Gasso, G. Mourit, and J. Ragot. Subspace identification of switching model. In *The 13th IFAC Symposium on System Identification*, Rotterdam, The Netherlands, August 2003.
- [Pis73] V. F. Pisarenko. The retrieval of harmonics from a covariance function. *Geophysic Journal of the Royal Astronomy Society*, 33 :347–366, 1973.

- [PSD96] K. Peterzell, W. Sherrer, and M. Deistler. Statistical analysis of novel subspace identification methods. *Signal Processing*, 52 :161–177, 1996.
- [QL03] S.J. Qin and L. Ljung. Closed loop subspace identification with innovation estimation. In *The 13th IFAC Symposium on System Identification*, Rotterdam, The Netherlands, August 2003.
- [Rab96] D. J. Rabideau. Fast, rank adaptive subspace tracking and applications. *IEEE Transactions on Signal Processing*, 44 :2229–2244, 1996.
- [RAB⁺01] J. P. Richard, J. Assan, A. Barraud, C. Commault, M. Dambrine, J. M. Dion, V. Kharitonov and J. F. Lafay, S. Lesecq, J. J. Loiseau, S. Mondié, S. I. Niculescu, W. Perruquetti, J. Ragot, V. Rasvan, and F. Rotella. *Algèbre et analyse pour l'automatique*, volume Systèmes automatisés, Information Commande Communication. Hermès, 2001.
- [Rei41] O. Reiersøl. Confluence analysis by means of lag moments and other methods of confluence analysis. *Econometrica*, 9 :1–23, 1941.
- [RK89] R. Roy and T. Kailath. ESPRIT estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37 :984–995, 1989.
- [Rot01] F. Rotella. *Algèbre et Analyse pour l'Automatique*, volume Systèmes automatisés, Information Commande Communication, chapter Formulaire de calcul matriciel, pages 215–264. Hermès, 2001.
- [RPK86] R. Roy, A. Paulraj, and T. Kailath. ESPRIT : A subspace rotation approach to estimation of parameters of cisoids in noise. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 34 :1340–1342, 1986.
- [Sch81] R. O. Schmidt. *A signal subspace approach to multiple emitter location and spectral estimation*. PhD thesis, Stanford University, Stanford, California, 1981.
- [Sch86] R. O. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas Propagation*, 34 :276–280, 1986.
- [Sch89] R. Schreiber. Implementation of adaptive array algorithms. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 34 :1550–1556, 1989.
- [SES93] P. Stoica, A. Eriksson, and T. Söderström. Subspace based algorithms without eigendecomposition for array processing. In *The IEEE International Conference on Acoustics, Speech and Signal Processing*, Minneapolis, Minnesota, USA, April 1993.
- [Shy89] J. J. Shynk. Adaptive IIR filtering. *IEEE Signal Processing Magazine*, 6 :4–21, 1989.
- [SK94] A. H. Sayed and T. Kailath. A state space approach to adaptive RLS filtering. *IEEE Signal Processing Magazine*, 11 :18–60, 1994.
- [SN90] P. Stoica and A. Nehorai. MUSIC, maximum likelihood and Cramer Rao bound : further results and comparisons. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38 :2140–2150, 1990.

- [Söd73] T. Söderström. An online algorithm for approximate maximum likelihood identification of linear dynamic systems. Technical report, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden, 1973.
- [Söd81] T. Söderström. Identification of stochastic linear systems in presence of input noise. *Automatica*, 17 :713–725, 1981.
- [SPDR03] J. Shoukens, R. Pintelon, T. Dobrowiecki, and Y. Rolain. Identification of linear systems with nonlinear distortions. In *The 13th IFAC Symposium on System Identification*, Rotterdam, The Netherlands, August 2003.
- [SS83a] P. Stoica and T. Söderström. Optimal instrumental variable estimation and approximate implementation. *IEEE Transactions on Automatic Control*, 28 :757–772, 1983.
- [SS83b] P. Stoica and T. Söderström. Optimal instrumental variable methods for identification of multivariable linear systems. *Automatica*, 19 :425–429, 1983.
- [SS89] T. Söderström and P. Stoica. *System identification*. Prentice Hall International Series in Systems and Control Engineering, New York, 1989.
- [SS92] P. Stoica and T. Söderström. Statistical analysis of a subspace method for bearing estimation without eigendecomposition. *IEEE Proceedings on Radar and Signal Processing*, 4 :301–305, 1992.
- [SVWW98] P. Stoica, M. Viberg, M. Wong, and Q. Wu. A unified instrumental variable approach to direction finding in colored noise fields. *Digital Signal Processing Handbook*, pages 1–18, 1998.
- [TBDM91] P. C. Tan, C. S. Berger, K. P. Dabke, and R. G. Mein. Recursive identification and adaptive prediction of wastewater flows. *Automatica*, 27 :761–768, 1991.
- [Tia01] Y. Tian. Etude de méthodes de diagnostic appliquées au suivi en ligne des matériels ferroviaires roulants. Technical report, INRETS et USTL, Lille, France, 2001.
- [TLS97] A. K. Tangirala, S. Lakshminarayanan, and S. L. Shah. Canonical variate analysis (CVA) for closed loop identification. In *The Canadian Society for Chemical Engineering Annual Meeting*, Edmonton, Canada, October 1997.
- [Ton00] P. Tona. *Conception de systèmes de commande avancée par ordinateur : méthodologie et applications*. PhD thesis, Institut National Polytechnique de Grenoble, Laboratoire d'Automatique de Grenoble, Grenoble, France, 2000.
- [UB03] T. Uhl and M. Bogacz. Real time modal analysis and its application for flutter testing. In *The 13th IFAC Symposium on System Identification*, Rotterdam, The Netherlands, August 2003.
- [Uts02] W. Utschick. Tracking of signal subspace projectors. *IEEE Transactions on Signal Processing*, 50 :769–778, 2002.
- [Vac94] R. J. Vaccaro. A second order perturbation expansion for the SVD. *SIAM Journal on Matrix Analysis and Applications*, 15 :661–671, 1994.
- [VB88] B. D. Van Veen and K. M. Buckley. Beamforming : a versatile approach to spatial filtering. *Signal Processing Magazine*, pages 4–24, 1988.

- [VD91] M. Verhaegen and E. Deprettere. A fast recursive MIMO state space model identification algorithm. In *The 30th IEEE Conference on Decision and Control*, Brighton, United Kingdom, December 1991.
- [VD92a] M. Verhaegen and P. Dewilde. Subspace model identification part 1 : output error state space model identification class of algorithms. *International Journal of Control*, 56 :1187–1210, 1992.
- [VD92b] M. Verhaegen and P. Dewilde. Subspace model identification part 2 : analysis of the elementary output error state space model identification algorithm. *International Journal of Control*, 56 :1211–1241, 1992.
- [VD93] P. Van Overschee and B. De Moor. N4SID : numerical algorithms for state space subspace system identification. In *The 12th IFAC World Congress*, Sydney, Australia, July 1993.
- [VD94] P. Van Overschee and B. De Moor. N4SID : subspace algorithms for the identification of combined deterministic stochastic systems. *Automatica*, 30 :75–93, 1994.
- [VD95] P. Van Overschee and B. De Moor. A unifying theorem for three subspace system identification algorithms. *Automatica*, 31 :1853–1864, 1995.
- [VD96a] P. Van Overschee and B. De Moor. Closed loop subspace system identification. Technical report, Katholieke Universiteit Leuven, Leuven, Belgium, 1996.
- [VD96b] P. Van Overschee and B. De Moor. *Subspace identification for linear systems. theory, implementation, applications*. Kluwer Academic Publishers, 1996.
- [VDCV94] S. Van Huffel, C. Decanniere, H. Chen, and P. Van Hecke. Algorithm for time domain NMR data fitting based on total least squares. *Journal of Magnetic Resonance*, 110 :228–237, 1994.
- [VDS93] A. J. Van der Veen, E. F. Deprettere, and A. L. Swindlehurst. Subspace based signal analysis using singular value decomposition. *Proceedings of the IEEE*, 81 :1277–1308, 1993.
- [Ver93] M. Verhaegen. Subspace model identification part 3 : analysis of the ordinary output error state space model identification algorithm. *International Journal of Control*, 58 :555–586, 1993.
- [Ver94] M. Verhaegen. Identification of the deterministic part of MIMO state space models given in innovations form from input output data. *Automatica*, 30 :61–74, 1994.
- [Vib95] M. Viberg. Subspace based methods for the identification of linear time invariant systems. *Automatica*, 31 :1835–1851, 1995.
- [VO91] M. Viberg and B. Ottersten. Sensor array processing based on subspace fitting. *IEEE Transactions on Signal Processing*, 39 :1110–1121, 1991.
- [VOK91] M. Viberg, B. Ottersten, and T. Kailath. Detection and estimation in sensor arrays using weighted subspace fitting. *IEEE Transactions on Signal Processing*, 39 :2436–2449, 1991.
- [VOWL93] M. Viberg, B. Ottersten, B. Wahlberg, and L. Ljung. Performance of subspace based state space system identification methods. Technical report, Department of Signal, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, 1993.

- [VS94] M. Viberg and A. L. Swindlehurst. A bayesian approach to auto calibration for parametric array signal processing. *IEEE Transactions on Signal Processing*, 42 :3495–3506, 1994.
- [VSO95] M. Viberg, P. Stoica, and B. Ottersten. Array processing in correlated noise fields based on instrumental variables and subspace fitting. *IEEE Transactions on Signal Processing*, 43 :1187–1199, 1995.
- [VV91] S. Van Huffel and J. Vandervalle. *The total least squares problem : computational aspects and analysis, frontiers in applied mathematics series*, volume 9. SIAM, Philadelphia, 1991.
- [VVG94] M. Verhaegen, A. Varga, and G. Grübel. Some experience with the MOESP class of subspace model identification methods in identifying the BO105 helicopter. Technical report, German Aerospace Research Establishment, Laboratory for Flight Systems Dynamics, Wessling, Germany, 1994.
- [VWO97] M. Viberg, B. Wahlberg, and B. Ottersten. Analysis of state space system identification methods based on instrumental variables and subspace fitting. *Automatica*, 33 :1603–1616, 1997.
- [War77] R. K. Ward. Notes on the instrumental variable method. *IEEE Transactions on Automatic Control*, 22 :482–484, 1977.
- [Wax91] M. Wax. Detection and localization of multiple sources via the stochastic signals models. *IEEE Transactions on Signal Processing*, 39 :2450–2456, 1991.
- [Wax92] M. Wax. Detection and localization of multiple sources in noise with unknown covariance. *IEEE Transactions on Signal Processing*, 40 :245–249, 1992.
- [WH60] B. Widrow and M.E. Hoff. Adaptive switching circuits. In *IRE Western Electric Show and Convention Record*, August 1960.
- [WKL03] X. Wang, U. Kruger, and B. Lennox. Recursive partial least squares algorithms for monitoring complex industrial processes. *Control Engineering Practice*, 11 :613–632, 2003.
- [WP67] K. Y. Wong and E. Polak. Identification of linear discrete time systems using the instrumental variable method. *IEEE Transactions on Automatic Control*, 12 :707–718, 1967.
- [WS85] B. Widrow and S. Stearns. *Adaptive Signal processing*. Prentice Hall, Englewood Cliffs, 1985.
- [WV96] D. Westwick and M. Verhaegen. Identifying MIMO Wiener systems using subspace model identification methods. *Signal Processing*, 52 :235–258, 1996.
- [XK94] G. Xu and T. Kailath. Fast subspace decomposition. *IEEE Transactions on Signal Processing*, 42 :539–551, 1994.
- [XRK94] G. Xu, R. H. Roy, and T. Kailath. Detection of number of sources via exploitation of centro symmetry property. *IEEE Transactions on Signal Processing*, 42 :102–112, 1994.
- [Yan95a] B. Yang. An extension of PASTd algorithm to both rank and subspace tracking. *IEEE Signal Processing Letters*, 2 :179–182, 1995.

- [Yan95b] B. Yang. Projection approximation subspace tracking. *IEEE Transactions on Signal Processing*, 43 :95–107, 1995.
- [Yan96] B. Yang. Asymptotic convergence analysis of the projection approximation subspace tracking algorithms. *Signal Processing*, 50 :123–136, 1996.
- [Yer00] A. Yeredor. The extended least squares criterion : minimization algorithms and applications. *IEEE Transactions on Signal Processing*, 49 :74–87, 2000.
- [YG94] B. Yang and F. Gersemsky. An adaptive algorithm of linear computational complexity for both rank and subspace tracking. In *The IEEE International Conference on Acoustics, Speech and Signal Processing*, Adelaide, Australia, April 1994.
- [YK88] J. Yang and M. Kaveh. Adaptive eigensubspace algorithms for direction or frequency estimation and tracking. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 36 :241–251, 1988.
- [You68] P. C. Young. The use of linear regression and related procedures for the identification of dynamic processes. In *The 7th IEEE Symposium on Adaptive Processes*, San Antonio, Texas, USA, October 1968.
- [You70] P. C. Young. An instrumental variable method for real time identification of noisy process. *Automatica*, 6 :271–287, 1970.
- [You84] P. C. Young. *Recursive estimation and time series analysis*. Springer Verlag, Berlin, 1984.
- [YSA89] X. Yang, T. S. Sarkar, and E. Arvas. A survey of conjugate gradient algorithms for solution of extreme eigenproblem of a symmetric matrix. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37 :1550–1556, 1989.
- [Yu00] J. L. Yu. A novel subspace tracking using correlation based projection approximation. *Signal Processing*, 80 :2517–2525, 2000.
- [ZD98] K. Zhou and J. C. Doyle. *Essentials of robust control*. Prentice Hall, Saddle River, New Jersey, 1998.
- [ZF89] W. X. Zheng and C. B. Feng. Unbiased parameter estimation of linear systems in the presence of input and output noise. *International Journal of Adaptive Control and Signal Processing*, 3 :231–251, 1989.
- [ZF92] W. X. Zheng and C. B. Feng. Identification of a class of dynamic error-in-variables models. *International Journal of Adaptive Control and Signal Processing*, 6 :431–440, 1992.
- [ZM74] H. Zeiger and A. McEwen. Approximate linear realizations of given dimension via Ho’s algorithm. *IEEE Transactions on Automatic Control*, 19 :153, 1974.
- [ZVDL94] Y. Zhu, P. Van Overschee, B. De Moor, and L. Ljung. Comparison of three classes of identification methods. In *The 10th IFAC Symposium on System Identification*, Copenhagen, Denmark, July 1994.
- [ZW93] Q. T. Zhang and K. M. Wong. Information theoretic criteria for the determination of the number of signals in spatially correlated noise. *IEEE Transactions on Signal Processing*, 41 :1652–1663, 1993.

Contribution à l'identification récursive des systèmes par l'approche des sous-espaces

Le travail présenté dans ce mémoire concerne l'étude et le développement de méthodes d'identification récursive. Le chapitre 2 présente une synthèse des principaux algorithmes adaptatifs classiques fondés sur les moindres carrés. Les difficultés rencontrées lors de l'utilisation de telles approches dans certaines situations pratiques motivent la création de techniques récursives employant la forme d'état. Ce type de représentation a démontré son efficacité en identification grâce à l'émergence des méthodes hors ligne des sous-espaces. Ainsi, le chapitre 3 propose une description des algorithmes MOESP. Ces derniers sont en effet considérés comme le point de départ des techniques récursives des sous-espaces. Bien que robuste, la décomposition en valeurs singulières (DVS) employée par ces méthodes les rend inapplicables en ligne de par sa charge calculatoire. Le chapitre 4 est dédié à un état de l'art des algorithmes récursifs des sous-espaces énoncés jusqu'en 2003. Ces méthodes ont en commun d'adapter un critère particulier de traitement d'antennes conduisant à des techniques alternatives à la DVS. Une description de ces algorithmes dans un contexte unifié est plus précisément présentée. Ces techniques présentent néanmoins un certain nombre de limitations principalement liées à la fonction coût utilisée. Afin de remédier à ces inconvénients, nous développons, au sein du chapitre 5, de nouvelles méthodes récursives des sous-espaces. L'approche proposée consiste à adapter un opérateur particulier du traitement d'antennes jusqu'alors inexploité en identification : le propagateur. L'ajustement de ce dernier au problème d'identification récursive permet d'obtenir des critères quadratiques sans approximation et sans contrainte. Le problème des perturbations est traité en introduisant une variable instrumentale. L'étude expérimentale sur des données réelles et simulées est réalisée dans le chapitre 6.

Mots clef : identification des systèmes, systèmes multivariables, algorithmes récursifs, méthodes des sous-espaces, traitement d'antennes, variable instrumentale.

Contribution to recursive system identification by the subspace approach

This work deals with the study and the development of recursive identification methods. Chapter 2 presents a synthesis of the main classical adaptive algorithms based on the least squares. The difficulties met with such approaches in some practical situations justify the creation of recursive techniques employing the state space structure. This kind of representation has proved its efficiency in the identification framework thanks to the emergence of the off line subspace methods. Chapter 3 proposes a description of the MOESP algorithms. These ones are in fact considered as the starting point of the subspace recursive techniques. Despite its robustness, the singular value decomposition (SVD) employed by these methods makes them on line unusable due to its prohibitive numerical complexity. Chapter 4 is dedicated to an overview of the recursive subspace algorithms developed till 2003. These methods have in common to adapt a particular array signal processing criterion in order to provide alternative techniques to the SVD. A description of these algorithms in a unified framework is more precisely introduced. These techniques have nevertheless some limitations mostly due to the used cost function. In order to overcome these drawbacks, we develop new recursive subspace methods in chapter 5. The proposed approach consists in adapting a particular array signal processing operator unused in identification so far: the propagator. Its adjustment to the recursive identification problem leads to quadratic criteria needing neither approximation nor constraint. The use of an instrumental variable makes possible the treatment of the perturbations. Chapter 6 is dedicated to an experimental validation of these new techniques.

Key words: system identification, MIMO systems, recursive algorithms, subspace methods, array signal processing, instrumental variable.