



HAL
open science

Contributions à l'identification de modèles avec des erreurs en les variables

Stéphane Thil

► **To cite this version:**

Stéphane Thil. Contributions à l'identification de modèles avec des erreurs en les variables. Sciences de l'ingénieur [physics]. Université Henri Poincaré - Nancy I, 2007. Français. NNT: . tel-00270956

HAL Id: tel-00270956

<https://theses.hal.science/tel-00270956>

Submitted on 8 Apr 2008

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.

Contributions à l'identification de modèles avec des erreurs en les variables

Thèse présentée pour l'obtention du

Doctorat de l'Université Henri Poincaré, Nancy 1

Spécialité Automatique, Traitement du Signal et Génie Informatique

par

Stéphane Thil

Soutenance publique le 04 décembre 2007

<i>Président :</i>	Jamal Daafouz	Professeur à l'Institut National Polytechnique de Lorraine
<i>Rapporteurs :</i>	Michel de Mathelin	Professeur à l'École Nationale Sup. de Physique de Strasbourg
	Eric Walter	Directeur de Rech. CNRS au Laboratoire des Signaux et Systèmes
<i>Examineurs :</i>	Torsten Söderström	Professeur à l'Université d'Uppsala, Suède
	Marion Gilson	Maître de Conférences à l'Université Henri Poincaré, Nancy 1
	Hugues Garnier	Professeur à l'Université Henri Poincaré, Nancy 1

à Marcel,

Remerciements

Je tiens tout d'abord à remercier les membres du jury, qui m'ont fait l'honneur d'accepter de participer à l'examen de ce travail.

Mes remerciements à Michel DE MATHELIN, professeur à l'École Nationale Supérieure de Physique de Strasbourg, et à Éric WALTER, directeur de recherche CNRS au Laboratoire des Signaux et Systèmes, pour avoir accepté d'être rapporteurs. Je leur suis reconnaissant d'avoir examiné avec attention mon travail, et d'avoir contribué à significativement améliorer le présent manuscrit.

Je suis très reconnaissant envers Torsten SÖDERSTRÖM, professeur à l'université d'Uppsala en Suède, d'avoir accepté d'être examinateur de ma thèse. Ce fut un grand honneur d'avoir au sein du jury l'un des plus éminents chercheurs du domaine. J'en profite également pour le remercier, ainsi que tous les autres membres du département « Information Technology », pour l'accueil chaleureux qu'ils m'ont réservé lors de mon séjour en Suède. Agnes, Björn, Erik, Håkan, Linda, Magnus, Mei, Torsten... tack för allt !

Merci enfin à Jamal DAAFOUZ, professeur à l'Institut National Polytechnique de Lorraine, d'avoir accepté de présider ce jury.

Avant de commencer mon diplôme d'études approfondies, j'avais déjà une idée assez précise du sujet du stage de recherche que j'allais choisir, ainsi que de l'encadrant (c'est d'ailleurs après en avoir discuté avec celui-ci que je m'étais décidé à suivre cette voie). Ce ne fut que quelques semaines avant le début de ce stage que j'appris que les choses ne se passeraient pas de cette manière. Que, finalement, il était *préférable* que mon stage soit encadré par une autre personne – et qu'il porte sur un autre sujet.

Ainsi, c'est suite à une mésaventure que, dès mon diplôme d'études approfondies, j'ai été encadré par Marion. Quelques mois plus tard, je débutais une thèse co-encadrée par Marion et Hugues : je connaissais peu le domaine, la co-encadrante à peine, le directeur de thèse pas du tout... Après trois longues années passées sous leur joug dur et cruel, je mesure aujourd'hui à quel point cette mésaventure n'en a pas été une. Merci à tous les deux !

Pour leur accueil enjoué et leur bonne humeur, je tiens également à remercier tous les membres du département Réseaux et Télécommunications de l'I.U.T. Nancy-Brabois, au sein duquel j'ai effectué mes enseignements (une mention particulière à mes amis volleyeurs).

Nul doute que j'eus fini ces trois années de thèse un peu plus inquiet de l'état avancé de mon délabrement mental, si je n'avais eu à mes côtés des amis me prouvant chaque jour que c'est un phénomène normal, que tous ont subi – que d'aucuns subissent encore – à des degrés divers... Puisqu'aux doctorants se mêlent maîtres de conférences et professeurs, et afin d'éviter l'écueil de la classification, qu'il me soit permis de ne pas procéder à une longue et fastidieuse énumération. Je sais qu'ils me pardonneront.

Pour finir, mes remerciements les plus chaleureux pour ma famille, sans qui rien n'eût été possible.

Table des matières

Table des figures	v
Liste des tableaux	vii
Liste des abréviations et symboles	ix
Liste des hypothèses	xi
1 Introduction	1
1.1 L'identification des systèmes	1
1.2 Formulation du problème EIV	3
1.2.1 Hypothèses générales – Problème considéré	3
1.2.2 Modèle du système « vrai »	5
1.2.3 Adéquation entre le modèle utilisé et le système « vrai »	6
1.3 Difficultés inhérentes à l'identification de modèles EIV	6
1.3.1 Un exemple introductif : mesure d'une résistance	7
1.3.2 Nécessité de poser des hypothèses supplémentaires	9
1.3.2.1 Estimateurs convergents	9
1.3.2.2 Identifiabilité	9
1.4 Organisation de la thèse et contributions	10
2 Identification de modèles à temps discret	13
2.1 Position du problème et notations	13
2.2 Estimateurs fondés sur les statistiques de second ordre	15
2.2.1 Un bref tour d'horizon	15
2.2.1.1 Méthode de Frisch (frisch)	15
2.2.1.2 Méthodes de l'erreur de prédiction (pem) et du maximum de vraisemblance (ml)	18
2.2.2 Compensation du biais de l'estimateur des moindres carrés (bcls)	20
2.2.2.1 Premier cas : bruit blanc en entrée et en sortie	21
2.2.2.2 Deuxième cas : bruit blanc en entrée et coloré en sortie	27
2.2.3 Méthodes fondées sur une variable instrumentale (iv et xiv)	30

2.2.3.1	Cas d'un bruit blanc en entrée et coloré en sortie	32
2.2.4	Compensation du biais de l'estimateur de la variable instrumentale	34
2.2.4.1	Méthode des moindres carrés séparables (ecls)	34
2.2.4.2	Méthode de compensation du biais de la variable instrumentale (bciv)	35
2.2.5	Récapitulatif des méthodes présentées	38
2.2.6	Simulations numériques	39
2.2.6.1	Analyse comparative des différentes méthodes bcls	40
2.2.6.2	Analyse comparative des méthodes fondées sur une variable instrumentale	45
2.3	Estimateurs fondés sur les statistiques d'ordre supérieur	48
2.3.1	Discussion des hypothèses d'application	49
2.3.2	Méthodes fondées sur la minimisation de critères faisant intervenir les cumulants d'ordre supérieur	50
2.3.3	Méthodes fondées sur une variable instrumentale (tociv, fociv)	52
2.3.4	Méthodes fondées sur l'équation du modèle avec cumulants	53
2.3.4.1	Estimateur des moindres carrés (tocls)	54
2.3.4.2	Estimateur des moindres carrés itératifs (tocils)	55
2.3.5	Matrice de covariance asymptotique de l'estimateur tocls	55
2.3.6	Simulations numériques	60
2.3.6.1	Analyse comparative des méthodes tociv, tocls et tocils	60
2.3.6.2	Matrice de covariance asymptotique	61
2.4	Conclusion	63
3	Identification de modèles à temps continu	67
3.1	Introduction et position du problème	67
3.2	Estimateurs fondés sur les statistiques de second ordre	70
3.2.1	Succinct état de l'art	70
3.2.2	Méthode fondée sur les moindres carrés séparables (eivsvf)	71
3.3	Estimateurs fondés sur les statistiques d'ordre supérieur	74
3.3.1	Méthodes fondées sur l'équation du modèle avec cumulants	74
3.3.2	Méthodes fondées sur les cumulants d'ordre trois	76
3.3.2.1	Estimateur des moindres carrés (tocls)	76
3.3.2.2	Estimateur des moindres carrés itératifs (tocils)	78
3.3.3	Mise en œuvre des algorithmes tocls et tocils	79
3.3.4	Méthodes fondées sur les cumulants d'ordre quatre (focls, focils)	82
3.4	Simulations numériques	84
3.4.1	Estimateurs fondés sur les cumulants d'ordre trois	85
3.4.1.1	Analyse des performances	85

3.4.1.2	Influence de l'horizon de calcul M	87
3.4.1.3	Influence de la pulsation de coupure du filtre des variables d'état λ	89
3.4.2	Estimateurs fondés sur les cumulants d'ordre quatre	90
3.4.2.1	Analyse des performances	90
3.4.2.2	Influence de l'horizon de calcul M	93
3.5	Conclusion	93
4	Conclusion et perspectives	97
A	Statistiques d'ordre supérieur : définitions et propriétés	101
A.1	Cas des variables aléatoires réelles scalaires	101
A.2	Cas des vecteurs aléatoires	103
A.3	Propriétés des cumulants	104
A.3.1	Variables aléatoires indépendantes	105
A.4	Signaux aléatoires	106
A.5	Aspects pratiques liés à l'utilisation de cumulants	107
A.5.1	Lignes de cumulants	107
A.5.2	Estimation des cumulants	108
A.5.2.1	Cumulants d'ordre trois	108
A.5.2.2	Cumulants d'ordre quatre	108
B	Annexe des chapitres 2 et 3	111
B.1	Borne de Cramér-Rao	111
B.2	Corrélation des résidus de la méthode des moindres carrés	111
B.3	Preuve des propositions relatives à la méthode bels	112
B.3.1	Preuve de la proposition 2.2.2 (fbels)	112
B.3.2	Preuve de la proposition 2.2.5 (bels2)	113
B.4	Preuve de la proposition 2.2.4 (pbcls)	114
B.5	Corrélation des résidus de la méthode des moindres carrés filtrés	114
B.6	Preuves des résultats du paragraphe 2.3.5	116
B.6.1	Preuve du lemme 2.3.1	116
B.6.2	Preuve du théorème 2.3.1	116
B.6.3	Preuve du lemme 2.3.2	117
B.6.4	Preuve du lemme 2.3.3	119
B.7	Preuve des propositions 2.3.1 et 3.3.1 (tocls , tocils , focls , focils)	120
C	Ajouts à la boîte à outils Contsid	123
C.1	Algorithme tocls	123
C.2	Algorithme tocils	125

C.3 Algorithme focls	127
C.4 Algorithme focils	129
Bibliographie	133

Table des figures

1.1	La procédure d'identification	1
1.2	Objectifs de l'identification et modèles recherchés	2
1.3	Modèle EIV à temps continu (hybride)	3
2.1	Schéma détaillé du modèle EIV à temps discret	14
2.2	Exemple 2.1 : diagrammes de Bode de $H_o^{u_o}$ et G_o	41
2.3	Exemple 2.1 : diagrammes de Bode du vrai système et des modèles estimés	43
2.4	Exemple 2.2 : diagrammes de Bode de $H_o^{u_o}$, $H_o^{\tilde{y}}$ et G_o	44
2.5	Exemple 2.2 : diagrammes de Bode du vrai système et des modèles estimés	45
2.6	Exemple 2.3 : diagrammes de Bode de $H_o^{u_o}$, $H_o^{\tilde{y}}$ et G_o	46
2.7	Exemple 2.3 : diagrammes de Bode du vrai système et des modèles estimés	47
2.8	Exemple 2.4 : diagrammes de Bode de $H_o^{u_o}$, $H_o^{\tilde{u}}$, $H_o^{\tilde{y}}$ et G_o	61
2.9	Exemple 2.4 : diagrammes de Bode du vrai système et des modèles estimés	62
3.1	Schéma détaillé du modèle EIV à temps continu	68
3.2	Exemple 3.1 : diagrammes de Bode de $H_o^{\tilde{u}}$, $H_o^{\tilde{y}}$ et G_o	85
3.3	Exemple 3.1 : partie des signaux d'entrée/sortie bruités et non-bruités	86
3.4	Exemple 3.1 : diagrammes de Bode du vrai système et des modèles estimés	88
3.5	Exemple 3.1 : tocls – Moyenne des estimées et écart-type en fonction de M	88
3.6	Exemple 3.1 : tocils – Moyenne des estimées et écart-type en fonction de M	89
3.7	Exemple 3.1 : EQMR en fonction de l'horizon de calcul M	89
3.8	Exemple 3.2 : partie des signaux d'entrée/sortie bruités et non-bruités	91
3.9	Exemple 3.2 : diagramme de Bode de $H_o^{\tilde{y}}$	91
3.10	Exemple 3.2 : diagrammes de Bode du vrai système et des modèles estimés	93
3.11	Exemple 3.2 : EQMR en fonction de l'horizon de calcul M	94

Liste des tableaux

2.2	Récapitulatif des méthodes utilisant les statistiques de second ordre	40
2.3	Exemple 2.1 : comparaison des méthodes bcls	42
2.4	Exemple 2.2 : comparaison des méthodes bels2 et abels2	45
2.5	Exemple 2.3 : comparaison des méthodes xiv , bciv et bels2	46
2.6	Exemple 2.4 : comparaison des méthodes tociv , tocls et tocils	62
2.7	Exemple 2.5 : comparaison des résultats théoriques et des résultats de simulation pour la méthode tocls	63
3.1	Exemple 3.1 : comparaison des méthodes ivsvf , srivc , tocls et tocils	87
3.2	Exemple 3.1 : influence de la pulsation de coupure λ	90
3.3	Exemple 3.2 : comparaison de eivsvf , focls et focls (bruits blancs, $N = 1000$)	92
3.4	Exemple 3.2 : comparaison de eivsvf , focls et focls (bruits blancs, $N = 5000$)	92
3.5	Exemple 3.2 : comparaison de eivsvf , focls et focls (bruits blanc/coloré, $N = 5000$)	92
B.1	Valeurs prises par τ_a , τ_b , τ_c , τ_d , τ_e et τ_f dans (B.71)	121

Liste des abréviations et symboles

ARMA	Abréviation pour modèle auto-régressif à moyenne ajustée
ARX	Abréviation pour modèle auto-régressif avec entrée exogène
\mathcal{E}^*	Ensemble de modèles de l'entrée non bruitée
$\bar{E}\{\cdot\}$	Espérance généralisée d'un processus quasi-stationnaire (voir (2.8))
EIV	Abréviation pour erreurs en les variables (<i>errors-in-variables</i> en anglais)
G_o	Fonction de transfert du processus « vrai » entre u_o et y_o
$G(\cdot, \boldsymbol{\theta})$	Fonction de transfert du modèle du processus dans une structure donnée, correspondante au vecteur des paramètres $\boldsymbol{\theta}$
\mathcal{G}^*	Ensemble de modèles du processus à identifier
\mathcal{H}^*	Ensemble de modèles des bruits en entrée/sortie
\mathbf{I}_n	Matrice identité de dimension n
$\mathcal{I}_n(k)$	Matrice carrée de dimension n ayant la $k^{\text{ième}}$ diagonale supérieure et la $k^{\text{ième}}$ diagonale inférieure composées de 1, et des zéros partout ailleurs
λ_e	Variance du processus aléatoire e
λ	Pulsation de coupure du filtre des variables d'état
\mathcal{M}^*	Ensemble de modèles, comprenant le modèle du processus, les modèles des bruits en entrée/sortie, ainsi que le modèle de l'entrée non-bruitée
MA	Abréviation pour modèle à moyenne ajustée
N	Nombre d'échantillons du jeu de données disponible
p	Opérateur différentiel
p.s.	Abréviation pour presque sûrement
$\boldsymbol{\varphi}$	Vecteur de régression
q	Opérateur avance
q^{-1}	Opérateur retard
RSB	Abréviation pour rapport signal à bruit
s	Variable de Laplace
\mathcal{S}	Système « vrai » générant les données
SOS	Abréviation pour statistiques d'ordre supérieur
T_e	Période d'échantillonnage
t_k	Instant d'échantillonnage $t_k = kT_e$

θ	Vecteur utilisé pour la paramétrisation des modèles d'entrée/sortie
$\hat{\theta}$	Estimée du vecteur des paramètres à partir du jeu de données disponible
θ_o	Vecteur des paramètres « vrais »
$\hat{\theta}^i$	Estimée de θ obtenue à la $i^{\text{ième}}$ itération
$\hat{\theta}_j$	Estimée de θ obtenue à la $j^{\text{ième}}$ réalisation de la simulation de Monte Carlo
θ^*	Limite de l'estimée de θ lorsque $N \rightarrow \infty$
u	Signal d'entrée mesuré
u_o	Signal d'entrée du modèle (sans bruit)
\check{u}_o	Signal d'entrée du système « vrai » (sans bruit)
\tilde{u}	Bruit additif sur la mesure de l'entrée du modèle
$\check{\tilde{u}}$	Bruit additif sur la mesure de l'entrée du système « vrai »
y	Signal de sortie mesuré
y_o	Signal de sortie du modèle (sans bruit)
\check{y}_o	Signal de sortie du système « vrai » (sans bruit)
\tilde{y}	Bruit additif sur la mesure de la sortie du modèle
$\check{\tilde{y}}$	Bruit additif sur la mesure de la sortie du système « vrai »
$y^{(n)}$	$n^{\text{ième}}$ dérivée du signal y
$(\cdot)^T$	Transposée d'une matrice
$\ \cdot \ $	Norme euclidienne
$(\cdot)^\dagger$	Pseudo-inverse d'une matrice
\star	Opérateur de convolution
\xrightarrow{dist}	Convergence en distribution

Liste des hypothèses

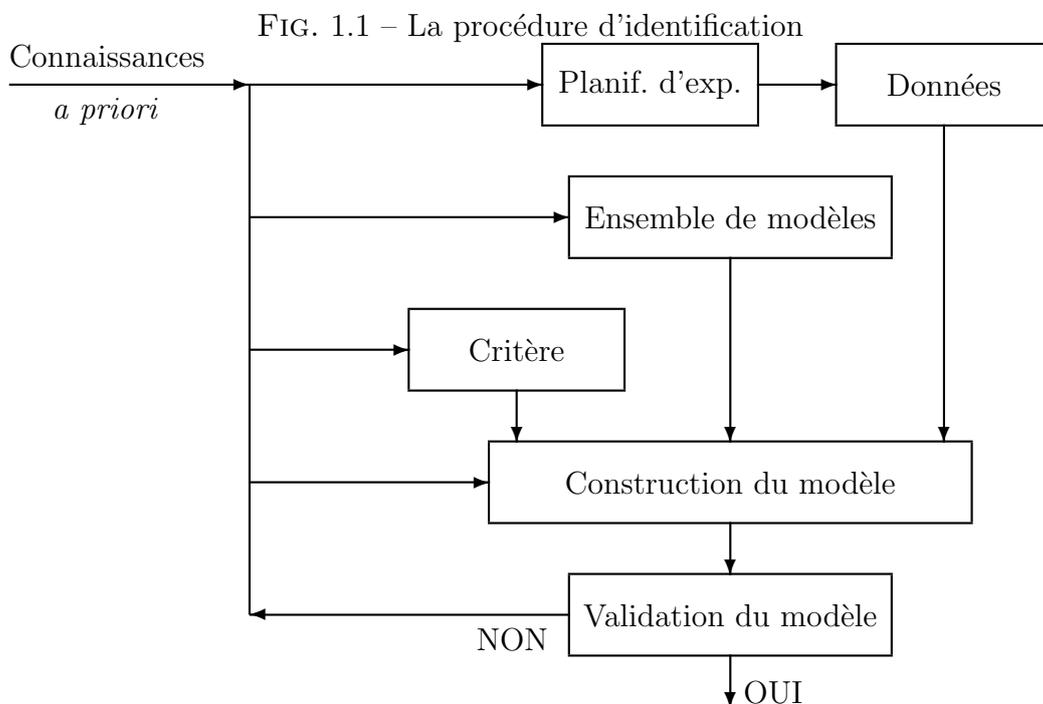
- H1.** La fonction de transfert G_o est asymptotiquement stable
 - H2.** Les modes du système sont observables et contrôlables
 - H3.** Les degrés n_a et n_b des polynômes A_o et B_o sont connus
 - H4.** $\check{u}_o, \check{u}, \check{y}$ sont ergodiques, stationnaires et de moyenne nulle
 - H5.** \check{u}_o est d'excitation persistante d'ordre suffisant, et n'est pas corrélé avec \check{u} , ni avec \check{y}
 - H6.** Les données disponibles, $\{u(t_k), y(t_k)\}_{k=1}^N$, proviennent du vrai système \mathcal{S}
-
- Hb1.** Les bruits en entrée et en sortie \tilde{u} et \tilde{y} sont non-corrélés
 - Hb2.** Les bruits en entrée et en sortie sont blancs
 - Hb3.** Le bruit en entrée est blanc, tandis que le bruit en sortie est coloré (filtre MA)
 - Hb4.** Le bruit en entrée est blanc, tandis que le bruit en sortie est coloré (filtre ARMA)
 - Hb5.** Les bruits en entrée et en sortie sont colorés (filtres ARMA)
 - Hb6.** Les bruits en entrée et en sortie ont une distribution ayant un axe de symétrie
 - Hb7.** Les bruits en entrée et en sortie sont gaussiens
-
- He1.** Le signal d'entrée non-bruité u_o a une densité spectrale rationnelle
 - He2.** Le signal d'entrée non-bruité u_o a une distribution asymétrique
 - He3.** Le signal d'entrée non-bruité u_o est non-gaussien
-
- Hs1.** $\mathcal{S} \in \mathcal{M}^*$: les fonctions de transfert $G_o, H_o^{u_o}, H_o^{\tilde{u}}$ et $H_o^{\tilde{y}}$ appartiennent à l'ensemble de modèles considéré
 - Hs2.** $G_o \in \mathcal{G}^*$ et $(H_o^{\tilde{u}}, H_o^{\tilde{y}}) \in \mathcal{H}^*$: les fonctions de transfert $G_o, H_o^{\tilde{u}}$ et $H_o^{\tilde{y}}$ appartiennent à l'ensemble de modèles considéré
 - Hs3.** $G_o \in \mathcal{G}^*$: la fonction de transfert G_o appartient à l'ensemble de modèles considéré

Chapitre 1

Introduction

1.1 L'identification des systèmes

L'élaboration de modèles de systèmes dynamiques représente un point clé en automatique et dans les sciences expérimentales. La procédure d'identification, qui suit les étapes détaillées dans la [figure 1.1](#) [VdH96, Joh93], consiste à rechercher un modèle mathématique adéquat pour un système donné à partir de données expérimentales et de connaissances disponibles *a priori*.



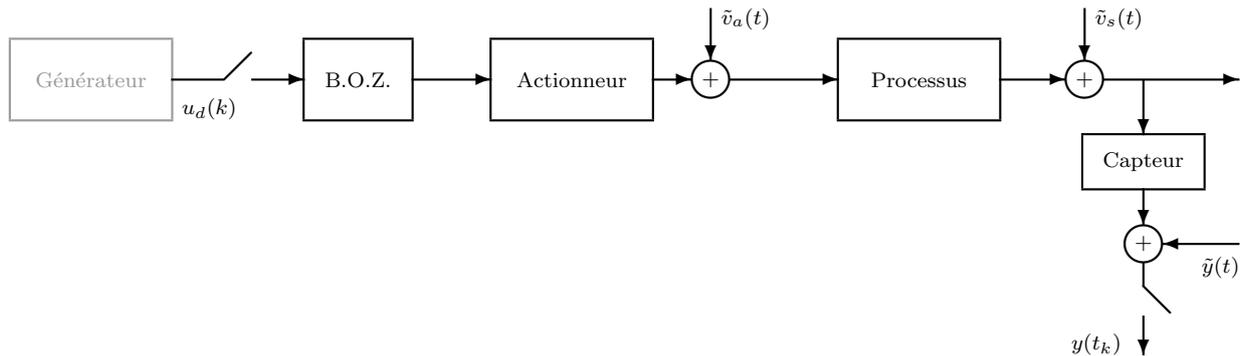
Les trois éléments principaux nécessaires lors d'une procédure d'identification sont :

- *les données* : elles peuvent être disponibles d'après les mesures d'un système fonctionnant de manière normale (on parle alors de conditions expérimentales passives [Tho05]) ; lorsque

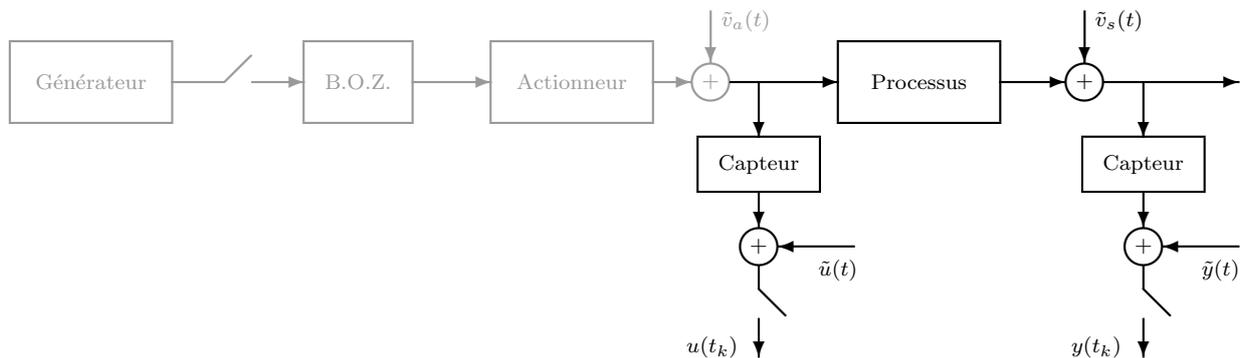
cela est possible, elles peuvent aussi provenir d'expériences particulièrement étudiées pour l'obtention d'informations spécifiques sur le système ;

- *l'ensemble de modèles* : définit l'ensemble au sein duquel le « meilleur » modèle sera obtenu. L'ensemble de modèles détermine par exemple la structure et l'ordre des modèles ;
- *le critère d'identification* : ayant à disposition des données et un ensemble de modèles, le critère d'identification spécifie dans quel sens le modèle recherché sera « meilleur ».

FIG. 1.2 – Objectifs de l'identification et modèles recherchés



(a) Synthèse d'une loi de commande : modèle entre $u_d(k)$ et $y(t_k)$



(b) Simulation : modèle entre $u(t_k)$ et $y(t_k)$

Le travail présenté dans ce document se focalise sur la partie centrale de la [figure 1.1](#) (données, ensemble de modèles, critère et construction du modèle).

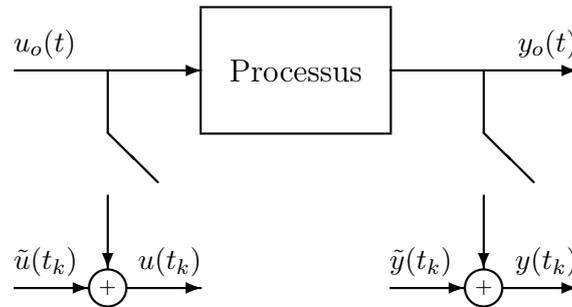
Durant les dernières décennies, de nombreuses méthodes ont été développées pour l'identification de systèmes à partir de leurs signaux d'entrée/sortie. La majorité des techniques ainsi développées suppose le signal d'entrée parfaitement connu. Une justification de cette hypothèse est fournie à la [figure 1.2](#) : si l'objectif de l'identification est, par exemple, la synthèse d'une loi de commande, la recherche du modèle s'effectue entre la sortie du générateur, connue, et la sortie du système, mesurée et entachée de bruits (d'actionneur \tilde{v}_a , de système \tilde{v}_s , de capteur \tilde{y}). En revanche, si l'objectif de l'identification est, par exemple, d'avoir une meilleure connaissance du processus, l'entrée du système n'est bien souvent accessible que par l'intermédiaire de mesures et est par conséquent entachée de bruit [PS01]. C'est cette dernière situation où l'entrée et la sortie du système sont toutes deux perturbées par des

bruits – que l’on nommera modèle « avec erreurs en les variables » – qui est étudiée dans ce document. L’abréviation correspondante sera EIV (d’après l’anglais « errors-in-variables »).

1.2 Formulation du problème EIV

Dans la suite de ce mémoire, deux types de modèles EIV sont considérés : les modèles à temps continu et les modèles à temps discret. Le modèle à temps continu étudié est de type hybride, c’est-à-dire que les bruits additifs sur les signaux d’entrée/sortie du processus à identifier sont à temps discret (voir la [figure 1.3](#)). Les disciplines scientifiques ayant recours à ce type de modèle incluent l’étude des séries temporelles, le traitement d’antenne, le traitement d’image, ou encore la modélisation environnementale [[VH02](#), [GYTG07](#)].

FIG. 1.3 – Modèle EIV à temps continu (hybride)



1.2.1 Hypothèses générales – Problème considéré

Suivant les notations employées dans l’ouvrage [[Lju99](#)], le système générant les données est noté¹ :

$$\mathcal{S} : \begin{cases} \check{y}_o(t_k) = G_o(q)\check{u}_o(t_k) = \frac{B_o(q^{-1})}{A_o(q^{-1})}\check{u}_o(t_k) \\ u(t_k) = \check{u}_o(t_k) + \check{\check{u}}(t_k) \\ y(t_k) = \check{y}_o(t_k) + \check{\check{y}}(t_k) \end{cases} \quad (1.1a)$$

où l’on note :

- les (vrais) signaux d’entrée/sortie non bruités sont respectivement notés \check{u}_o et \check{y}_o ;
- les (vrais) bruits additifs affectant les signaux d’entrée/sortie sont respectivement notés $\check{\check{u}}$ et $\check{\check{y}}$;
- les signaux d’entrée/sortie disponibles sont respectivement notés u et y .

¹On prend ici l’exemple d’un modèle à temps discret. Les notations sont néanmoins directement transposables au cas d’un modèle à temps continu.

Lorsque cela sera nécessaire, nous nous référerons aux vrais opérateurs de transfert de l'entrée non bruitée et des bruits additifs, notés :

$$\begin{cases} \check{u}_o(t_k) = H_o^{u_o}(q)e_o^{u_o}(t_k) \\ \check{u}(t_k) = H_o^{\check{u}}(q)e_o^{\check{u}}(t_k) \\ \check{y}(t_k) = H_o^{\check{y}}(q)e_o^{\check{y}}(t_k) \end{cases} \quad (1.1b)$$

où $e_o^{u_o}$, $e_o^{\check{u}}$ et $e_o^{\check{y}}$ sont des bruits blancs.

Introduisons les hypothèses générales suivantes :

- H1.** la fonction de transfert G_o est asymptotiquement stable ;
- H2.** les modes du système sont observables et contrôlables, *i.e.* les polynômes A_o et B_o n'ont pas de facteurs communs ;
- H3.** les degrés n_a et n_b des polynômes A_o et B_o sont connus ;
- H4.** \check{u}_o , \check{u} , \check{y} sont des processus aléatoires ergodiques², stationnaires au sens strict³ et de moyenne nulle ;
- H5.** \check{u}_o est d'excitation persistante d'ordre suffisant, et n'est pas corrélé avec \check{u} ni avec \check{y} ;
- H6.** les données disponibles, $\{u(t_k), y(t_k)\}_{k=1}^N$, sont générées par le vrai système \mathcal{S} .

Ces hypothèses seront supposées vraies tout au long de ce document. Au fil des chapitres s'ajouteront d'autres hypothèses, spécifiques à telle méthode ou nécessaires pour l'application de telle autre. Ces dernières n'auront de validité qu'au sein de la section dans laquelle elles auront été introduites.

Le vecteur des paramètres du processus vrai est noté θ_o . Le problème considéré peut maintenant être formulé de la manière suivante :

Problème : sous les hypothèses **H1–H6**, et ayant à disposition un jeu de données d'entrée/sortie $\{y(t_k), u(t_k)\}_{k=1}^N$, obtenir un estimateur convergent du vecteur des paramètres θ_o .

Pour les besoins de l'identification, nous définissons un ensemble de modèles pouvant être considérés comme candidats possibles pour le système, et dans lequel le « meilleur » modèle est déterminé, suivant le critère choisi. Cet ensemble de modèles est défini dans le paragraphe suivant.

²Plus précisément, on supposera les processus *ergodiques à l'ordre des statistiques utilisées*. Ainsi, lorsque les cumulants d'ordre quatre seront utilisés, les processus seront supposés ergodiques à l'ordre quatre.

³Cette hypothèse peut être adoucie en la remplaçant par *stationnaires à l'ordre des statistiques utilisées*.

1.2.2 Modèle du système « vrai »

Le modèle que l'on considère, noté \mathcal{M} , est associé à la structure suivante :

$$\begin{cases} y_o(t_k) = G(q, \boldsymbol{\theta})u_o(t_k) \\ u(t_k) = u_o(t_k) + \tilde{u}(t_k) \\ y(t_k) = y_o(t_k) + \tilde{y}(t_k) \end{cases} \quad \text{avec} \quad \begin{cases} u_o(t_k) = H^{u_o}(q, \boldsymbol{\eta}_{u_o})e^{u_o}(t_k) \\ \tilde{u}(t_k) = H^{\tilde{u}}(q, \boldsymbol{\eta}_{\tilde{u}})e^{\tilde{u}}(t_k) \\ \tilde{y}(t_k) = H^{\tilde{y}}(q, \boldsymbol{\eta}_{\tilde{y}})e^{\tilde{y}}(t_k) \end{cases} \quad (1.2)$$

où l'on utilise les notations suivantes :

- les signaux d'entrée/sortie non bruités sont respectivement notés u_o et y_o ;
- les bruits additifs affectant les signaux d'entrée/sortie sont respectivement notés \tilde{u} et \tilde{y} ;
- les signaux d'entrée/sortie disponibles sont respectivement notés u et y ;
- e^{u_o} , $e^{\tilde{u}}$ et $e^{\tilde{y}}$ désignent des bruits blancs.

Par ailleurs, puisque les bruits sont additifs, on adopte la convention suivante pour toute variable linéaire en les données mesurées : la partie non bruitée de la variable sera notée avec l'indice “ o ”, tandis que la contribution du bruit sera coiffée d'un signe “ \sim ”. Ainsi, par exemple, le signal de sortie mesuré s'écrit $y(t_k) = y_o(t_k) + \tilde{y}(t_k)$.

Formellement, la modélisation utilisée est définie comme une fonction d'un ouvert $D_{\mathcal{M}}$ de $\mathbb{R}^{n_{\Theta}}$ dans un ensemble de modèles \mathcal{M}^* :

$$\begin{aligned} \mathcal{M}: D_{\mathcal{M}} &\longrightarrow \mathcal{M}^* = \{\mathcal{M}(\boldsymbol{\Theta}) \mid \boldsymbol{\Theta} \in D_{\mathcal{M}}\} \\ \boldsymbol{\Theta} &\longmapsto \mathcal{M}(\boldsymbol{\Theta}) \end{aligned}$$

La modélisation \mathcal{M} est très générale et contient les modèles du processus, de l'entrée non bruitée et des bruits additifs sur les signaux d'entrée/sortie ; tous les coefficients de ces modèles sont regroupés au sein du vecteur des paramètres $\boldsymbol{\Theta}$. On suppose que ces modèles sont paramétrés de manière indépendante, si bien qu'il est possible de décomposer l'ensemble de modèles \mathcal{M}^* de la manière suivante :

$$D_{\mathcal{M}} = D_{\theta} \times D_{\boldsymbol{\eta}_{u_o}} \times D_{\boldsymbol{\eta}_{\tilde{u}}} \times D_{\boldsymbol{\eta}_{\tilde{y}}} \quad (1.3a)$$

$$\mathcal{M}^* = \mathcal{G}^* \cup \mathcal{E}^* \cup \mathcal{H}^* \quad (1.3b)$$

$$\mathcal{G}^* = \left\{ G(q, \boldsymbol{\theta}) = \frac{B(q^{-1}, \boldsymbol{\theta})}{A(q^{-1}, \boldsymbol{\theta})} \mid \boldsymbol{\theta} \in D_{\theta} \right\} \quad (1.3c)$$

$$\mathcal{E}^* = \left\{ H^{u_o}(q, \boldsymbol{\eta}_{u_o}) \mid \boldsymbol{\eta}_{u_o} \in D_{\boldsymbol{\eta}_{u_o}} \right\} \quad (1.3d)$$

$$\mathcal{H}^* = \left\{ (H^{\tilde{u}}(q, \boldsymbol{\eta}_{\tilde{u}}), H^{\tilde{y}}(q, \boldsymbol{\eta}_{\tilde{y}})) \mid \boldsymbol{\eta}_{\tilde{u}} \in D_{\boldsymbol{\eta}_{\tilde{u}}}, \boldsymbol{\eta}_{\tilde{y}} \in D_{\boldsymbol{\eta}_{\tilde{y}}} \right\} \quad (1.3e)$$

On a ainsi :

$$\mathcal{M}(\boldsymbol{\Theta}) = \{G(q, \boldsymbol{\theta}), H^{u_o}(q, \boldsymbol{\eta}_{u_o}), H^{\tilde{u}}(q, \boldsymbol{\eta}_{\tilde{u}}), H^{\tilde{y}}(q, \boldsymbol{\eta}_{\tilde{y}})\} \quad (1.3f)$$

$$\boldsymbol{\Theta}^T = [\boldsymbol{\theta}^T \quad \boldsymbol{\eta}_{u_o}^T \quad \boldsymbol{\eta}_{\tilde{u}}^T \quad \boldsymbol{\eta}_{\tilde{y}}^T] \in D_{\mathcal{M}} \quad (1.3g)$$

Il convient de noter que l'ensemble de ces modélisations n'est pas toujours nécessaire. En effet, toutes les méthodes présentées dans la suite ne nécessitent pas de modéliser l'entrée non bruitée ou les bruits affectant les signaux d'entrée/sortie du système pour fournir une estimation du vecteur des paramètres d'intérêt θ . Bien que, comme nous le verrons dans la suite, certaines méthodes nécessitent l'estimation de certains d'entre eux, les autres paramètres contenus dans les vecteurs η_{u_o} , $\eta_{\tilde{u}}$ et $\eta_{\tilde{y}}$ ne présentent pas un intérêt direct, et leur estimation n'est pas impérative. Lorsque le cas se présentera, nous ferons donc les hypothèses *ad hoc*.

1.2.3 Adéquation entre le modèle utilisé et le système « vrai »

Ayant ainsi défini le système à identifier (1.1a) et la modélisation choisie (1.3), plusieurs situations peuvent survenir, parmi lesquelles on citera :

Hs1. $\mathcal{S} \in \mathcal{M}^*$: situation idéale, mais peu réaliste d'un point de vue pratique. Elle signifie en effet que les fonctions de transfert G_o , $H_o^{u_o}$, $H_o^{\tilde{u}}$ et $H_o^{\tilde{y}}$ appartiennent à l'ensemble de modèles considéré ;

Hs2. $G_o \in \mathcal{G}^*$ et $(H_o^{\tilde{u}}, H_o^{\tilde{y}}) \in \mathcal{H}^*$: situation intermédiaire ne faisant pas d'hypothèse structurale sur le signal d'entrée non bruité. Les fonctions de transfert G_o , $H_o^{\tilde{u}}$ et $H_o^{\tilde{y}}$ appartiennent à l'ensemble de modèles considéré ;

Hs3. $G_o \in \mathcal{G}^*$: situation plus réaliste, puisqu'elle suppose uniquement que la fonction de transfert G_o appartient à l'ensemble de modèles considéré.

Afin de *significativement* simplifier les notations et de garder une présentation agréable, les développements effectués dans la suite de ce mémoire ne feront pas intervenir le vecteur des paramètres vrais θ_o , ni le vrai signal d'entrée non bruité ou les vrais signaux de bruits en entrée/sortie : nous travaillerons à partir des équations du modèle. Néanmoins, sous les hypothèses idoines, un estimateur convergent du vecteur des paramètres du modèle θ sera un estimateur convergent de θ_o . Ainsi, bien que le vrai système n'intervienne plus dans les développements qui suivent, nous donnons toutefois sous laquelle des hypothèses **Hs1–Hs3** les méthodes présentées fournissent une estimation convergente de θ_o .

1.3 Difficultés inhérentes à l'identification de modèles EIV

L'identification des systèmes en présence de bruits additifs sur les signaux d'entrée/sortie est reconnue comme étant un problème plus difficile que l'identification « classique »⁴. En effet, la plupart des estimateurs communément utilisés pour traiter le cas classique fournissent des estimées biaisées pour ce problème. Pour illustrer ce point, nous considérons ci-dessous un

⁴*i.e.* lorsque le signal d'entrée est parfaitement connu.

problème d'identification *a priori* très simple : l'estimation de la valeur d'une résistance à partir de mesures du courant et de la tension à ses bornes. Cet exemple est développé dans l'ouvrage [SP91].

1.3.1 Un exemple introductif : mesure d'une résistance

L'objectif est d'estimer la valeur d'une résistance R , dans laquelle passe un courant I_o constant, produisant la tension $U_o = RI_o$ à ses bornes. Supposons que l'on réalise N expériences indépendantes où l'on mesure le courant passant par la résistance et la tension à ses bornes. Ces mesures de courant et de tension sont toutes entachées d'erreurs, et les données disponibles sont en fait :

$$I(t_k) = I_o + \tilde{I}(t_k) \quad U(t_k) = U_o + \tilde{U}(t_k)$$

Supposons enfin que les bruits à des instants de mesures différents soient indépendants, que les bruits de mesure de courant et de tension soient indépendants, et que ces bruits soient de moyenne nulle. Donc, pour tout j, k , on a :

$$\begin{aligned} \mathbb{E}\{\tilde{I}(t_k)\} &= 0 & \mathbb{E}\{\tilde{U}(t_k)\} &= 0 & \mathbb{E}\{\tilde{I}(t_j)\tilde{U}(t_k)\} &= 0 \\ \mathbb{E}\{\tilde{I}(t_j)\tilde{I}(t_k)\} &= \lambda_{\tilde{I}}\delta_{j,k} & \mathbb{E}\{\tilde{U}(t_j)\tilde{U}(t_k)\} &= \lambda_{\tilde{U}}\delta_{j,k} \end{aligned}$$

où $\delta_{\cdot,\cdot}$ représente le symbole de Kronecker et $\mathbb{E}\{\cdot\}$ l'espérance mathématique. On considère un modèle ayant un paramètre inconnu R , dont l'équation est :

$$U(t_k) - RI(t_k) = v(t_k)$$

où $v(t_k)$ est « l'erreur d'équation ». En minimisant :

$$V_1 = \sum_{k=1}^N v^2(t_k) = \sum_{k=1}^N (U(t_k) - RI(t_k))^2 \quad (1.4)$$

par rapport à R , on obtient l'estimateur des moindres carrés :

$$\begin{aligned} \hat{R}_{\text{ls}} &= \frac{\frac{1}{N} \sum_{k=1}^N (U_o I_o + U_o \tilde{I}(t_k) + I_o \tilde{U}(t_k) + \tilde{I}(t_k) \tilde{U}(t_k))}{\frac{1}{N} \sum_{k=1}^N (I_o^2 + 2I_o \tilde{I}(t_k) + \tilde{I}^2(t_k))} \\ &= \frac{R + \frac{1}{N} \sum_{k=1}^N \left(R \frac{\tilde{I}(t_k)}{I_o} + \frac{\tilde{U}(t_k)}{I_o} + \frac{\tilde{I}(t_k) \tilde{U}(t_k)}{I_o^2} \right)}{1 + \frac{1}{N} \sum_{k=1}^N \left(2 \frac{\tilde{I}(t_k)}{I_o} + \frac{\tilde{I}^2(t_k)}{I_o^2} \right)} \\ &= \frac{R + \alpha}{1 + \beta} \simeq R + \alpha - \beta R - \alpha\beta + \beta^2 R + \alpha\beta^2 \end{aligned}$$

Les termes d'ordre supérieur ou égal à trois ont été négligés lors de la dernière approximation. Les hypothèses statistiques sur $\tilde{I}(t_k)$ et $\tilde{U}(t_k)$ impliquent que :

$$\mathbb{E}\{\alpha\} = 0 \quad \mathbb{E}\{\beta\} = \frac{\lambda_{\tilde{I}}}{I_o^2} \quad \mathbb{E}\{\alpha\beta\} = \frac{2}{N} \frac{\lambda_{\tilde{I}}}{I_o^2} R \quad \mathbb{E}\{\beta^2\} = \frac{4}{N} \frac{\lambda_{\tilde{I}}}{I_o^2}$$

On a ainsi :

$$\mathbb{E}\{\hat{R}_{\text{ls}}\} \simeq R \left(1 - \frac{\lambda_{\tilde{I}}}{I_o^2} + \frac{2}{N} \frac{\lambda_{\tilde{I}}}{I_o^2} \right)$$

L'estimateur des moindres carrés est donc biaisé, et même asymptotiquement biaisé puisque :

$$\lim_{N \rightarrow \infty} \mathbb{E}\{\hat{R}_{\text{ls}}\} \simeq R \left(1 - \frac{\lambda_{\tilde{I}}}{I_o^2} \right)$$

On voit donc que ce biais est uniquement dû au bruit de mesure sur le courant. Pour la variance de l'estimateur, on obtient :

$$\lambda_{\hat{R}_{\text{ls}}} = \frac{R^2}{N} \left(\frac{\lambda_{\tilde{U}}}{U_o^2} + \frac{\lambda_{\tilde{I}}}{I_o^2} \right)$$

La variance est ainsi proportionnelle à la somme des rapports bruit à signal du courant et de la tension, et elle tend vers zéro asymptotiquement. Ceci est typique de beaucoup d'estimateurs. On constate aussi que, si le bruit de mesure de la tension n'induit aucun biais, il a en revanche un impact sur la variance de l'estimateur. L'erreur sur les mesures du courant a ainsi deux conséquences sur l'estimateur des moindres carrés : une variance plus grande et un biais asymptotique.

Considérons maintenant un autre estimateur. Plutôt que de minimiser la somme des carrés des erreurs (1.4), on peut minimiser le critère suivant :

$$V_2 = \sum_{k=1}^N \left(\frac{\tilde{U}(t_k)^2}{\lambda_{\tilde{U}}} + \frac{\tilde{I}(t_k)^2}{\lambda_{\tilde{I}}} \right) = \sum_{k=1}^N \left(\frac{(U(t_k) - U_o)^2}{\lambda_{\tilde{U}}} + \frac{(I(t_k) - I_o)^2}{\lambda_{\tilde{I}}} \right)$$

si l'on suppose les variances $\lambda_{\tilde{U}}$, $\lambda_{\tilde{I}}$ connues. Ce problème de minimisation sous la contrainte $U_o = RI_o$ se résout en utilisant une technique Lagrangienne. On obtient :

$$\hat{R}_2 = \frac{\sum_{k=1}^N U(t_k)}{\sum_{k=1}^N I(t_k)}$$

On peut montrer que cet estimateur est asymptotiquement non biaisé, et que, pour N suffisamment grand :

$$\lambda_{\hat{R}_2} = \frac{R^2}{N} \left(\frac{\lambda_{\tilde{U}}}{U_o^2} + \frac{\lambda_{\tilde{I}}}{I_o^2} \right)$$

La variance sur les estimées fournies par cet estimateur est donc identique à celle sur \hat{R}_{ls} , mais l'ajout des hypothèses sur la variance des bruits permet d'obtenir une estimée qui converge vers R plutôt que vers une valeur biaisée.

1.3.2 Nécessité de poser des hypothèses supplémentaires

1.3.2.1 Estimateurs convergents

Dans l'exemple précédent, nous avons constaté que l'estimateur des moindres carrés est biaisé, mais qu'il est possible d'obtenir un estimateur asymptotiquement non biaisé en posant une hypothèse supplémentaire sur les bruits perturbant les données. Cela est typique de l'identification de modèles EIV : il est souvent crucial de formuler des hypothèses supplémentaires sur les bruits en entrée/sortie ou le signal d'entrée non bruité pour obtenir des estimées non biaisées du vecteur des paramètres, alors que dans le problème d'identification classique, ces hypothèses n'affectent habituellement que la variance sur les estimées [Söd03].

Parmi ces hypothèses supplémentaires, certaines, plus restrictives, simplifient significativement le problème, et donnent lieu à des méthodes qui ne seront pas évoquées dans ce document :

- la méthode de Koopmans-Levin (voir [Koo36, Lev64, FN85]) et la méthode des moindres carrés totaux (voir [GL80, VH02]) s'appliquent lorsque les bruits en entrée/sortie sont blancs, et ont une variance égale (ou, trivialement, lorsque le rapport des variances est connu) ;
- certaines méthodes s'appliquent sous l'hypothèse que le signal d'entrée non bruité du système est périodique [SPVG97, FGM99, SH05] ;
- dans [MKVH06] a été considérée l'utilisation de plusieurs jeux de données obtenus avec des conditions expérimentales différentes : il est supposé que le spectre du signal d'entrée non bruité change d'une expérience à l'autre, tandis que les propriétés des bruits restent identiques.

Il convient aussi de noter que ce document se focalise sur les méthodes d'estimation paramétrique dans le domaine temporel pour l'identification de modèles EIV. Les approches développées dans le domaine fréquentiel (voir par exemple [GPS95, SPVG97, BCS97, PS01, SMS03]) ne sont donc pas abordées. Par ailleurs, les méthodes fondées sur la décomposition en sous-espaces [SCE95, CS96b, CV97] n'apparaissent pas non plus dans ce document. Enfin, notons que si l'accent n'a pas été mis sur l'identification en ligne, beaucoup des algorithmes présentés peuvent néanmoins se décliner sous une forme récursive. Un panorama des algorithmes récursifs d'estimation paramétrique peut être trouvé dans [dM01], par exemple.

1.3.2.2 Identifiabilité

Nous venons de voir qu'il est crucial de formuler des hypothèses supplémentaires pour obtenir des estimées non biaisées du vecteur des paramètres. Il existe un autre problème récurrent du modèle EIV : son identifiabilité n'est pas garantie, si l'on impose uniquement des hypothèses générales telles que **H1–H5** [AD84]. Nous donnons ici une courte bibliographie sur les travaux réalisés jusqu'à présent.

L'étude de l'identifiabilité des modèles EIV a suscité beaucoup d'intérêt depuis près d'un siècle. L'attention s'est dans un premier temps portée sur le cas des systèmes statiques, les premiers travaux remontant au XIX^{ième} siècle [Adc77, Adc78, Fri34]. Ce n'est que plus tard que le cas des systèmes dynamiques a été étudié. Les travaux sont nombreux : [AD84, And85, CS96a, Dei86, Sol86, AD87, DA89, CS96a] pour n'en citer que quelques-uns. Un travail de synthèse a été récemment réalisé dans [AG06, AG08], en utilisant l'analyse spectrale. On y trouve un résultat général sur l'identifiabilité de modèles EIV paramétriques, à partir de statistiques de second ordre.

Les études mentionnées ci-dessus sont fondées uniquement sur l'utilisation des moments d'ordre deux des signaux disponibles. Or, il est possible d'utiliser les moments d'ordre supérieurs à deux pour réaliser l'identification de modèles EIV. Il est donc d'un intérêt certain de spécifier sous quelles conditions l'identifiabilité est garantie, et, bien que moins nombreux, des travaux ont néanmoins été réalisés dans ce sens. À nouveau, le cas statique a été tout d'abord étudié, et, sous l'hypothèse d'observations non-gaussiennes, il a été montré que les moments d'ordre supérieur de ces observations peuvent contenir suffisamment d'information pour garantir l'identifiabilité [Gea42, Rei50]. Ces analyses ont ensuite été étendues au cas dynamique par [Aka66], et plusieurs travaux ont été réalisés par la suite (par exemple [Dei86, DA89] en utilisant les multispectres, [Gia90] en utilisant les cumulants). Bien qu'un travail de synthèse semble être manquant dans la littérature, il est néanmoins possible de se référer au résultat [DA89, Theorem 6.1, p.61].

1.4 Organisation de la thèse et contributions

Ce mémoire est organisé de la manière suivante.

Le second chapitre de la thèse traite de l'identification de modèles EIV à temps discret, et est lui-même divisé en deux parties. La première porte sur les méthodes utilisant les statistiques d'ordre deux. L'identification de modèles EIV à temps discret à l'aide des statistiques d'ordre deux a été un domaine de recherche très actif au cours des dernières années. Aussi, nous exposons dans un premier temps les principales méthodes développées. Un accent particulier est alors mis sur les méthodes de compensation du biais de l'estimateur des moindres carrés. Des estimateurs fondés sur la technique de la variable instrumentale sont ensuite proposés. La seconde partie s'intéresse aux méthodes ayant recours aux statistiques d'ordre supérieur (à deux). Après un rapide état de l'art, des estimateurs utilisant une variable instrumentale sont étudiés : il est en effet possible de définir des vecteurs d'instruments « particuliers », conduisant à l'utilisation de cumulants d'ordre supérieur. Les estimateurs des moindres carrés et des moindres carrés itératifs fondés sur l'équation du système, qui est vérifiée par les cumulants, sont alors présentés. Enfin, une expression de la matrice de covariance asymptotique de l'estimateur des moindres carrés fondés sur les cumulants d'ordre trois est obtenue.

Le troisième chapitre traite de l'identification de modèles EIV à temps continu. Si l'identification de modèles EIV à temps discret a fait l'objet de nombreux travaux au cours des dernières années, l'identification de modèles à temps continu n'a en revanche été que très peu étudiée. Ce domaine est donc encore très largement ouvert. Dans un premier temps nous revenons sur l'intérêt particulier des méthodes directes d'identification de modèles à temps continu. Un état de l'art est ensuite dressé, au cours duquel nous revenons sur les rares méthodes existantes, utilisant toutes les statistiques d'ordre deux. Des estimateurs ayant recours aux cumulants d'ordre trois et d'ordre quatre sont ensuite proposés, en insistant sur les spécificités liées à leur mise en œuvre. Ces estimateurs permettent en particulier de s'affranchir des hypothèses structurelles sur les bruits additifs en entrée/sortie du système. Ils peuvent ainsi être utilisés lorsque ceux-ci sont tous deux colorés (et même mutuellement corrélés), cas de figure ne pouvant être traité jusqu'à présent.

Afin d'avoir une présentation cohérente, il a été choisi de ne pas séparer les contributions de la thèse des parties plus bibliographiques. Elles sont donc disséminées au fil du document. Les travaux effectués lors de la thèse ont donné lieu aux publications suivantes :

1. S. THIL, M. GILSON. Closed-loop system identification : a bayesian approach. *16th IFAC World Congress*. Prague, Czech Republic, July 2005.
2. S. THIL, M. GILSON ET H. GARNIER. Méthodes de compensation de biais pour l'identification de modèles erreurs en les variables. *Journées Doctorales du GDR MACS 2005*. Lyon, France, septembre 2005.
3. S. THIL, H. GARNIER ET M. GILSON. Une méthode pour l'identification de modèles à temps continu dans un contexte erreurs en les variables. *Conférence Internationale Francophone d'Automatique (CIFA'2006)*. Bordeaux, France, juin 2006.
4. S. THIL, H. GARNIER ET M. GILSON. A cumulant statistics-based method for continuous-time errors-in-variables model identification. *EURASIP Workshop on Total Least Squares and Errors-in-Variables Modeling*. Leuven, Belgium, August 2006.
5. H. GARNIER, P.C. YOUNG, S. THIL ET M. GILSON. Data-based Continuous-time Modelling of Environmental Systems. *STIC & Environnement*. Lyon, France, novembre 2007.
6. S. THIL, H. GARNIER, M. GILSON ET K. MAHATA. Continuous-time model identification from noisy input/output measurements using fourth-order cumulants. *46th Conference on Decision and Control (CDC'2007)*. New Orleans, LA, USA, December 2007.
7. S. THIL, H. GARNIER ET M. GILSON. Third-order cumulants based methods for continuous-time errors-in-variables model identification. *Automatica* (44), 2008. À paraître.
8. S. THIL, M. GILSON ET H. GARNIER. On instrumental variable-based methods for errors-in-variables model identification. *17th IFAC World Congress*. Seoul, Korea, July

2008.

9. S. THIL, M. HONG, T. SÖDERSTRÖM, M. GILSON ET H. GARNIER. Statistical analysis of a third-order cumulants based algorithm for discrete-time errors-in-variables identification. *17th IFAC World Congress*. Seoul, Korea, July 2008.

Chapitre 2

Identification de modèles à temps discret

2.1 Position du problème et notations

Nous nous intéressons dans ce chapitre aux méthodes dédiées à l'identification de modèles EIV à temps discret. Dans un premier temps, les notations et hypothèses générales qui sont utilisées dans la suite sont explicitées.

Conformément aux notations introduites dans le chapitre 1, nous considérons un modèle \mathcal{M} associé à la structure suivante (voir la [figure 2.1](#)) :

$$\begin{cases} y_o(t_k) = G(q, \boldsymbol{\theta})u_o(t_k) \\ u(t_k) = u_o(t_k) + \tilde{u}(t_k) \\ y(t_k) = y_o(t_k) + \tilde{y}(t_k) \end{cases} \quad (2.1a)$$

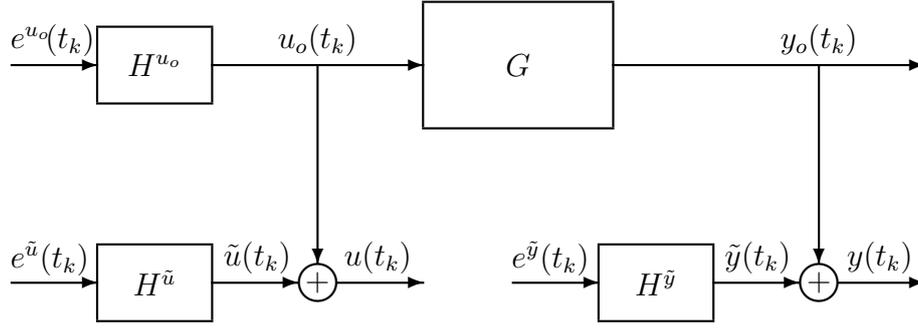
où l'on note :

- $G(q, \boldsymbol{\theta})$ l'opérateur de transfert du modèle du processus ;
- u_o, y_o les signaux d'entrée/sortie non bruités ;
- \tilde{u}, \tilde{y} les bruits de mesure sur l'entrée et la sortie ;
- u, y les signaux d'entrée/sortie mesurés.

Les signaux d'entrée non bruitée et de bruits sont modélisés par :

$$\begin{cases} u_o(t_k) = H^{u_o}(q, \boldsymbol{\eta}_{u_o})e^{u_o}(t_k) \\ \tilde{u}(t_k) = H^{\tilde{u}}(q, \boldsymbol{\eta}_{\tilde{u}})e^{\tilde{u}}(t_k) \\ \tilde{y}(t_k) = H^{\tilde{y}}(q, \boldsymbol{\eta}_{\tilde{y}})e^{\tilde{y}}(t_k) \end{cases} \quad (2.1b)$$

FIG. 2.1 – Schéma détaillé du modèle EIV à temps discret



où $e^{u_o}(t_k)$, $e^{\tilde{u}}(t_k)$, $e^{\tilde{y}}(t_k)$ désignent des bruits blancs. Les variances de $e^{\tilde{u}}(t_k)$, $e^{\tilde{y}}(t_k)$ sont notées $\lambda_{e^{\tilde{u}}}$ et $\lambda_{e^{\tilde{y}}}$. L'ensemble de modèles \mathcal{M}^* est défini comme :

$$\mathcal{M}^* = \left\{ G(q, \boldsymbol{\theta}), H^{u_o}(q, \boldsymbol{\eta}_{u_o}), H^{\tilde{u}}(q, \boldsymbol{\eta}_{\tilde{u}}), H^{\tilde{y}}(q, \boldsymbol{\eta}_{\tilde{y}}) \mid \boldsymbol{\Theta} \in D_{\boldsymbol{\Theta}} \right\} \quad (2.1c)$$

$$\mathcal{G}^* = \left\{ G(q, \boldsymbol{\theta}) = \frac{B(q^{-1}, \boldsymbol{\theta})}{A(q^{-1}, \boldsymbol{\theta})} = \frac{b_1 q^{-1} + \dots + b_{n_b} q^{-n_b}}{1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a}} \mid \boldsymbol{\theta} \in D_{\boldsymbol{\theta}} \right\} \quad (2.1d)$$

$$\mathcal{E}^* = \left\{ H^{u_o}(q, \boldsymbol{\eta}_{u_o}) \mid \boldsymbol{\eta}_{u_o} \in D_{\boldsymbol{\eta}_{u_o}} \right\} \quad (2.1e)$$

$$\mathcal{H}^* = \left\{ (H^{\tilde{u}}(q, \boldsymbol{\eta}_{\tilde{u}}), H^{\tilde{y}}(q, \boldsymbol{\eta}_{\tilde{y}})) \mid \boldsymbol{\eta}_{\tilde{u}} \in D_{\boldsymbol{\eta}_{\tilde{u}}}, \boldsymbol{\eta}_{\tilde{y}} \in D_{\boldsymbol{\eta}_{\tilde{y}}} \right\} \quad (2.1f)$$

où les vecteurs des paramètres sont donnés par :

$$\boldsymbol{\Theta}^T = [\boldsymbol{\theta}^T \quad \boldsymbol{\eta}_{u_o}^T \quad \boldsymbol{\eta}_{\tilde{u}}^T \quad \boldsymbol{\eta}_{\tilde{y}}^T] \quad (2.1g)$$

$$\boldsymbol{\theta}^T = [\mathbf{a}^T \quad \mathbf{b}^T] = [a_1 \quad \dots \quad a_{n_a} \quad b_1 \quad \dots \quad b_{n_b}] \quad (2.1h)$$

Par ailleurs, les notations suivantes sont utilisées dans la suite :

$$\bar{\mathbf{y}}^T(t_k) = [y(t_k) \quad \mathbf{y}^T(t_k)] \quad (2.2)$$

$$\mathbf{y}^T(t_k) = [y(t_{k-1}) \quad \dots \quad y(t_{k-n_a})] = \mathbf{y}_o^T(t_k) + \tilde{\mathbf{y}}^T(t_k) \quad (2.3)$$

$$\mathbf{u}^T(t_k) = [u(t_{k-1}) \quad \dots \quad u(t_{k-n_b})] = \mathbf{u}_o^T(t_k) + \tilde{\mathbf{u}}^T(t_k) \quad (2.4)$$

$$\boldsymbol{\varphi}^T(t_k) = [-\mathbf{y}^T(t_k) \quad \mathbf{u}^T(t_k)] = \boldsymbol{\varphi}_o^T(t_k) + \tilde{\boldsymbol{\varphi}}^T(t_k) \quad (2.5)$$

$$\bar{\boldsymbol{\theta}}^T = [\bar{\mathbf{a}}^T \quad \mathbf{b}^T] = [1 \quad \mathbf{a}^T \quad \mathbf{b}^T] = [1 \quad a_1 \quad \dots \quad a_{n_a} \quad b_1 \quad \dots \quad b_{n_b}] \quad (2.6)$$

Le vecteur des paramètres vrais est noté $\boldsymbol{\theta}_o$. Le modèle (2.1) peut alors être écrit sous la forme d'une régression linéaire :

$$y(t_k) = \boldsymbol{\varphi}^T(t_k) \boldsymbol{\theta} + v(t_k, \boldsymbol{\theta}) \quad (2.7a)$$

$$v(t_k, \boldsymbol{\theta}) = \tilde{y}(t_k) - \tilde{\boldsymbol{\varphi}}^T(t_k) \boldsymbol{\theta} \quad (2.7b)$$

où v est l'erreur d'équation. Enfin, on adopte les notations suivantes pour les vecteurs et les matrices de covariance :

- $\mathbf{R}_{vw} = \bar{\mathbb{E}}\{\mathbf{v}\mathbf{w}^T\}$ dénote la matrice d'intercovariance entre deux vecteurs aléatoires \mathbf{v} et \mathbf{w} ;
- $\mathbf{r}_{vs} = \bar{\mathbb{E}}\{\mathbf{v}s\}$ dénote le vecteur d'intercovariance du vecteur aléatoire \mathbf{v} et de la variable aléatoire s ;
- $r_x(\tau) = \bar{\mathbb{E}}\{x(t_k)x(t_{k+\tau})\}$ dénote la fonction d'autocovariance d'un processus aléatoire stationnaire et centré x .

où $\bar{\mathbb{E}}\{\cdot\}$ représente l'espérance généralisée, définie par [Lju99] :

$$\bar{\mathbb{E}}\{x(t_k)\} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \mathbb{E}\{x(t_k)\} \quad (2.8)$$

Les estimées des vecteurs et matrices de covariance obtenues à l'aide du jeu de données disponible seront notées $\hat{\mathbf{R}}$ et $\hat{\mathbf{r}}$.

2.2 Estimateurs fondés sur les statistiques de second ordre

Dans cette partie, nous nous intéressons aux méthodes s'appuyant sur les statistiques de second ordre pour identifier le modèle (2.1). Ce domaine de recherche a été très actif ces dernières années, et de nombreuses méthodes ont été développées. Dans un premier temps, nous rappelons les principales méthodes existantes, puis nous nous attardons plus particulièrement sur la famille des méthodes dites de compensation du biais de l'estimateur des moindres carrés. Quelques estimateurs fondés sur la technique de la variable instrumentale sont ensuite proposés, avant d'illustrer les performances des méthodes présentées à l'aide de simulations numériques.

2.2.1 Un bref tour d'horizon

2.2.1.1 Méthode de Frisch (frisch)

La méthode de Frisch a été développée dans un premier temps pour un problème classique d'estimation algébrique dans [Fri34]. Elle a été appliquée à l'identification de systèmes dynamiques par [BGS90, SD98], et a ensuite fait l'objet de plusieurs travaux, parmi lesquels on pourra citer par exemple [DGS03, Söd08].

Cette méthode s'applique lorsque les bruits en entrée et en sortie sont blancs¹ :

Hb1. \tilde{u} et \tilde{y} sont non-corrélés ;

Hb2. $H^{\tilde{u}}(q) = H^{\tilde{y}}(q) = 1$, *i.e.* les bruits additifs \tilde{u} et \tilde{y} sont blancs, de variances respectives $\lambda_{e\tilde{u}}$ et $\lambda_{e\tilde{y}}$;

¹Il convient de noter de l'hypothèse **Hb2** implique l'hypothèse **Hb1**. De manière générale, dans la suite du mémoire, nous précisons les deux hypothèses pour insister sur le fait que, lorsque les bruits ne sont plus blancs, l'hypothèse de non-corrélation est nécessaire.

Hs2. $G_o \in \mathcal{G}^*$ et $(H_o^{\tilde{u}}, H_o^{\tilde{y}}) \in \mathcal{H}^*$.

Il est nécessaire d'introduire quelques notations supplémentaires pour le vecteur de régression et le vecteur des paramètres « augmentés » :

$$\bar{\varphi}^T(t_k) = [-y(t_k) \ \varphi^T(t_k)] = [-\bar{y}^T(t_k) \ \mathbf{u}^T(t_k)] \quad (2.9)$$

$$\bar{\boldsymbol{\theta}}^T = [1 \ \boldsymbol{\theta}^T] \quad (2.10)$$

De l'équation (2.1a) on déduit :

$$\bar{\varphi}_o^T(t_k) \bar{\boldsymbol{\theta}} = 0 \quad (2.11)$$

et donc :

$$\mathbf{R}_{\bar{\varphi}_o \bar{\varphi}_o} \bar{\boldsymbol{\theta}} = 0 \quad (2.12)$$

La matrice $\mathbf{R}_{\bar{\varphi}_o \bar{\varphi}_o}$ est par conséquent singulière (semi-définie positive), avec au moins une valeur propre égale à 0. Il est en fait possible de montrer que, sous l'hypothèse **H5**, la matrice $\mathbf{R}_{\bar{\varphi}_o \bar{\varphi}_o}$ n'a qu'une seule valeur propre nulle [Söd05]. Le vecteur propre associé à cette valeur propre est $\bar{\boldsymbol{\theta}}$. L'équation (2.12) constitue le point de départ de la méthode : puisque $\mathbf{R}_{\bar{\varphi} \bar{\varphi}}$ est connue (calculée à partir des signaux mesurés), si l'on dispose d'une estimée $\mathbf{R}_{\tilde{\varphi} \tilde{\varphi}}$, il est possible d'estimer le vecteur des paramètres en exploitant la relation :

$$(\mathbf{R}_{\bar{\varphi} \bar{\varphi}} - \mathbf{R}_{\tilde{\varphi} \tilde{\varphi}}) \bar{\boldsymbol{\theta}} = 0 \quad (2.13)$$

Or, sous l'hypothèse **Hb2**, la matrice $\mathbf{R}_{\tilde{\varphi} \tilde{\varphi}}$ s'écrit :

$$\mathbf{R}_{\tilde{\varphi} \tilde{\varphi}} = \begin{bmatrix} \lambda_{e^{\tilde{y}}} \mathbf{I}_{n_a+1} & 0 \\ 0 & \lambda_{e^{\tilde{u}}} \mathbf{I}_{n_b} \end{bmatrix} \quad (2.14)$$

Il faut ainsi obtenir une estimée des variances $\lambda_{e^{\tilde{u}}}$ et $\lambda_{e^{\tilde{y}}}$. Supposons, dans un premier temps, que la variance du bruit blanc en entrée $\lambda_{e^{\tilde{u}}}$ est disponible. Alors la variance du bruit sur la sortie est déterminée de manière à ce que la matrice $(\mathbf{R}_{\bar{\varphi} \bar{\varphi}} - \mathbf{R}_{\tilde{\varphi} \tilde{\varphi}})$ apparaissant en (2.13) soit singulière. Il est en effet possible de montrer que cette variance vérifie [Söd05] :

$$\lambda_{e^{\tilde{y}}} = \sigma_{\min}(\mathbf{R}_{\bar{y}\bar{y}} - \mathbf{R}_{\bar{y}u}(\mathbf{R}_{uu} - \lambda_{e^{\tilde{u}}} \mathbf{I}_{n_b})^{-1} \mathbf{R}_{u\bar{y}}) \quad (2.15)$$

où $\sigma_{\min}(\mathbf{M})$ représente la plus petite valeur propre de \mathbf{M} . On constate par (2.15) que $\lambda_{e^{\tilde{y}}}$ est une fonction de $\lambda_{e^{\tilde{u}}}$. Donc, si l'on suppose $\lambda_{e^{\tilde{u}}}$ connue, on peut en déduire $\lambda_{e^{\tilde{y}}}$. Lorsque les deux variances sont connues, l'estimée de $\boldsymbol{\theta}$ est alors obtenue en résolvant (2.13) :

$$\left(\mathbf{R}_{\varphi\varphi} - \begin{bmatrix} \lambda_{e^{\tilde{y}}} \mathbf{I}_{n_a} & 0 \\ 0 & \lambda_{e^{\tilde{u}}} \mathbf{I}_{n_b} \end{bmatrix} \right) \boldsymbol{\theta}_{\text{frisch}}^* = \mathbf{r}_{\varphi y} \quad (2.16)$$

Il reste maintenant à obtenir une estimée de $\lambda_{e^{\tilde{u}}}$. Deux alternatives ont été proposées :

- la fonction $\lambda_{e\tilde{y}}(\lambda_{e\tilde{u}})$ est évaluée pour le modèle nominal et pour un modèle étendu, obtenu en ajoutant un paramètre à A et/ou B . Alors, idéalement, lorsque l'on trace les courbes de ces deux fonctions dans le plan $(\lambda_{e\tilde{y}}, \lambda_{e\tilde{u}})$, elles se coupent en un unique point, dont les coordonnées donnent les estimées (voir [BGS90] pour plus de détails) ;
- calculer les résidus de la méthode, et comparer leurs propriétés statistiques avec celles de résidus théoriques basés sur le modèle (voir [DGS03] pour plus de détails).

Les estimées fournies par la méthode **frisch** suivent asymptotiquement une loi gaussienne dont une expression de la variance peut être trouvée dans [Söd05] (la deuxième alternative décrite ci-dessus y est considérée).

Extension au cas d'un bruit coloré en sortie

La méthode de Frisch a été récemment étendue au cas d'un bruit coloré en sortie dans [Söd08].

Hb1. \tilde{u} et \tilde{y} sont non-corrélés ;

Hb4. $H^{\tilde{u}}(q) = 1$, $H^{\tilde{y}}(q) \neq 1$, *i.e.* le bruit en entrée \tilde{u} est blanc, de variance $\lambda_{e\tilde{u}}$, tandis que \tilde{y} est coloré (voir la [figure 2.1](#)) ;

Hs2. $G_o \in \mathcal{G}^*$ et $(H_o^{\tilde{u}}, H_o^{\tilde{y}}) \in \mathcal{H}^*$.

Comme il a été vu précédemment, la méthode **frisch** utilise les équations :

$$(\mathbf{R}_{\varphi\varphi} - \mathbf{R}_{\tilde{\varphi}\tilde{\varphi}}) \boldsymbol{\theta} = \mathbf{r}_{\varphi y} \quad (2.17)$$

$$\lambda_{e\tilde{y}} = \sigma_{\min}(\mathbf{R}_{\tilde{y}\tilde{y}} - \mathbf{R}_{\tilde{y}u}(\mathbf{R}_{uu} - \lambda_{e\tilde{u}} \mathbf{I}_{n_u})^{-1} \mathbf{R}_{u\tilde{y}}) \quad (2.18)$$

Lorsque le bruit en sortie est coloré, l'équation (2.17) devient :

$$(\mathbf{R}_{\varphi\varphi} - \mathbf{R}_{\tilde{\varphi}\tilde{\varphi}}) \boldsymbol{\theta} = \mathbf{r}_{\varphi y} - \mathbf{r}_{\tilde{\varphi}\tilde{y}} \quad (2.19)$$

Suivant la même idée que dans le cas où les bruits sont blancs, supposons que le vecteur :

$$\boldsymbol{\rho} = [r_{\tilde{y}}(0) \ \dots \ r_{\tilde{y}}(n)] \quad (2.20)$$

soit disponible, où n est une valeur à définir dans la suite. Il est alors possible de trouver la variance $\lambda_{e\tilde{u}}$ de manière à ce que la matrice $\mathbf{R}_{\tilde{\varphi}\tilde{\varphi}} - \mathbf{R}_{\tilde{\varphi}\tilde{\varphi}}$ soit singulière. On obtient alors une relation similaire à (2.18), à savoir :

$$\lambda_{e\tilde{u}} = \sigma_{\min}(\mathbf{R}_{uu} - \mathbf{R}_{u\tilde{y}}(\mathbf{R}_{\tilde{y}\tilde{y}} - \mathbf{R}_{\tilde{y}\tilde{y}}(\boldsymbol{\rho}))^{-1} \mathbf{R}_{\tilde{y}u}) \quad (2.21)$$

Enfin, l'estimation du vecteur $\boldsymbol{\rho}$ est réalisée en exploitant une propriété de la corrélation entre les résidus et les entrées retardées :

$$\bar{\mathbb{E}} \left\{ \begin{bmatrix} u(t_{k-n_b-1}) \\ \vdots \\ u(t_{k-n_b-m}) \end{bmatrix} (y(t_k) - \varphi^T(t_k) \boldsymbol{\theta}_{\text{frisch}}^*) \right\} = 0 \quad (2.22)$$

Cette relation fournit m équations. Pour que le nombre d'équations et le nombre d'inconnues soient compatibles, il faut alors choisir m et n de manière à ce que $n_a \leq n \leq m - 1$ [Söd08].

2.2.1.2 Méthodes de l'erreur de prédiction (**pem**) et du maximum de vraisemblance (**ml**)

Pour l'application de ces deux méthodes au cas EIV, le modèle est vu comme un système multivariable où u et y jouent le rôle de sorties [Söd81]. Les hypothèses nécessaires à leur application sont :

He1. le signal d'entrée non-bruité u_o a une densité spectrale rationnelle ;

Hb1. \tilde{u} et \tilde{y} sont non-corrélés ;

Hb2. $H^{\tilde{u}}(q) = H^{\tilde{y}}(q) = 1$, *i.e.* les bruits additifs \tilde{u} et \tilde{y} sont blancs, de variances respectives $\lambda_{e^{\tilde{u}}}$ et $\lambda_{e^{\tilde{y}}}$ (voir la figure 2.1) ;

Hs1. $\mathcal{S} \in \mathcal{M}^*$.

L'hypothèse **He1** est cruciale ; elle permet de considérer u_o comme la sortie d'un modèle ARMA excité par un bruit blanc :

$$u_o(t_k) = H^{u_o}(q, \boldsymbol{\eta}_{u_o}) e^{u_o}(t_k) = \frac{C^{u_o}(q^{-1}, \boldsymbol{\eta}_{u_o})}{D^{u_o}(q^{-1}, \boldsymbol{\eta}_{u_o})} e^{u_o}(t_k) \quad (2.23)$$

Le modèle EIV dans son ensemble est alors considéré comme un système excité par trois bruits blancs non corrélés e^{u_o} , \tilde{u} et \tilde{y} , et ayant un vecteur de sortie à deux dimensions $\mathbf{z} = [y \ u]^T$:

$$\begin{bmatrix} y(t_k) \\ u(t_k) \end{bmatrix} = \begin{bmatrix} \frac{B(q^{-1}, \boldsymbol{\theta}) C^{u_o}(q^{-1}, \boldsymbol{\eta}_{u_o})}{A(q^{-1}, \boldsymbol{\theta}) D^{u_o}(q^{-1}, \boldsymbol{\eta}_{u_o})} & 1 & 0 \\ \frac{C^{u_o}(q^{-1}, \boldsymbol{\eta}_{u_o})}{D^{u_o}(q^{-1}, \boldsymbol{\eta}_{u_o})} & 0 & 1 \end{bmatrix} \begin{bmatrix} e^{u_o}(t_k) \\ \tilde{y}(t_k) \\ \tilde{u}(t_k) \end{bmatrix} \quad (2.24)$$

Le modèle du signal d'entrée non-bruité $H^{u_o}(q, \boldsymbol{\eta}_{u_o})$ fait ainsi partie du modèle global, et ses coefficients doivent être estimés avec ceux de $A(q^{-1}, \boldsymbol{\theta})$ et $B(q^{-1}, \boldsymbol{\theta})$. Le vecteur des paramètres à estimer est donc $\boldsymbol{\vartheta}^T = [\boldsymbol{\theta}^T \ \boldsymbol{\eta}_{u_o}^T]$. Le modèle (2.24) est alors écrit sous forme d'état, qui est elle-même mise sous sa forme d'innovation (voir par exemple [AM79, Söd02]). Le signal d'innovation peut être interprété comme l'erreur de prédiction à un pas [Söd81], définie par :

$$\boldsymbol{\varepsilon}(t_k, \boldsymbol{\vartheta}) = \mathbf{z}(t_k) - \hat{\mathbf{z}}(t_k | t_{k-1}, \boldsymbol{\vartheta}) \quad (2.25)$$

Cette erreur de prédiction dépend des données et des matrices du modèle d'état. Le vecteur des paramètres $\boldsymbol{\vartheta}$ est estimé à partir de $\{\boldsymbol{\varepsilon}(t_k, \boldsymbol{\vartheta})\}_{k=1}^N$ en minimisant une fonction de coût $V_N(\boldsymbol{\vartheta})$:

$$\hat{\boldsymbol{\vartheta}} = \arg \min_{\boldsymbol{\vartheta}} V_N(\boldsymbol{\vartheta}) \quad (2.26)$$

Méthode à erreur de prédiction (**pem**)

Cette méthode vise à déterminer $\boldsymbol{\vartheta}$ de manière à minimiser l'erreur de prédiction $\boldsymbol{\varepsilon}(t_k, \boldsymbol{\vartheta})$. Il

n'existe pas une manière unique de définir un critère pour pénaliser l'erreur de prédiction. Une fonction de coût communément choisie est :

$$V_N(\boldsymbol{\vartheta}) = \det \left(\frac{1}{N} \sum_{k=1}^N \boldsymbol{\varepsilon}(t_k, \boldsymbol{\vartheta}) \boldsymbol{\varepsilon}^T(t_k, \boldsymbol{\vartheta}) \right) \quad (2.27)$$

L'erreur d'estimation est asymptotiquement distribuée selon une loi normale. Soit \mathbf{Q} la matrice de covariance des erreurs de prédiction. On a [SS89, Lju99] :

$$\sqrt{N} \left(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta} \right) \xrightarrow{dist} \mathcal{N}(0, \mathbf{P}_{\text{pem}}) \quad (2.28)$$

où :

$$\mathbf{P}_{\text{pem}} = \left(\bar{\mathbb{E}} \{ \boldsymbol{\psi}^T(t_k, \boldsymbol{\vartheta}) \mathbf{Q}^{-1}(\boldsymbol{\vartheta}) \boldsymbol{\psi}(t_k, \boldsymbol{\vartheta}) \} \right)^{-1} \quad (2.29)$$

$$\text{avec } \boldsymbol{\psi}(t_k, \boldsymbol{\vartheta}) = - \left(\frac{\partial \boldsymbol{\varepsilon}(t_k, \boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}} \right)^T \quad (2.30)$$

Méthode du maximum de vraisemblance (ml)

Si les données sont normalement distribuées (bruits et entrée non-bruitée), l'estimateur du maximum de vraisemblance minimise la fonction de coût suivante :

$$V_N(\boldsymbol{\vartheta}) = \frac{1}{N} \sum_{t=1}^N \ell(\boldsymbol{\varepsilon}(t_k, \boldsymbol{\vartheta}), \boldsymbol{\vartheta}, t_k) \quad (2.31)$$

où la fonction de vraisemblance est donnée par :

$$\ell(\boldsymbol{\varepsilon}(t_k, \boldsymbol{\vartheta}), \boldsymbol{\vartheta}, t_k) = \frac{1}{2} \log \det \mathbf{Q}(\boldsymbol{\vartheta}) + \frac{1}{2} \boldsymbol{\varepsilon}^T(t_k, \boldsymbol{\vartheta}) \mathbf{Q}^{-1}(\boldsymbol{\vartheta}) \boldsymbol{\varepsilon}(t_k, \boldsymbol{\vartheta}) \quad (2.32)$$

Pour le problème EIV considéré, l'estimateur du maximum de vraisemblance est plus précis que celui à erreur de prédiction. En effet, l'erreur d'estimation est asymptotiquement distribuée selon une loi normale [Söd06] :

$$\sqrt{N} \left(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta} \right) \xrightarrow{dist} \mathcal{N}(0, \mathbf{P}_{\text{ml}}) \quad (2.33)$$

où

$$\mathbf{P}_{\text{ml}} = \left(\bar{\mathbb{E}} \{ \boldsymbol{\psi}^T(t_k, \boldsymbol{\vartheta}) \mathbf{Q}^{-1}(\boldsymbol{\vartheta}) \boldsymbol{\psi}(t_k, \boldsymbol{\vartheta}) \} + \mathbf{P} \right)^{-1} \quad (2.34)$$

$$\text{où } P_{ij} = \frac{1}{2} \text{tr} \left(\mathbf{Q}^{-1} \mathbf{Q}_i \mathbf{Q}^{-1} \mathbf{Q}_j \right) \quad \text{et } \mathbf{Q}_i = \frac{\partial \mathbf{Q}}{\partial \vartheta_i} \quad (2.35)$$

En examinant (2.34) et (2.29), on remarque que le terme additionnel \mathbf{P} rend l'estimateur du maximum de vraisemblance plus précis. De plus, la borne de Cramér-Rao² est égale à \mathbf{P}_{ml} lorsque les données sont gaussiennes [Söd07].

²Voir l'annexe B pour des détails sur cette borne.

REMARQUE

Ces deux méthodes présentent l'avantage de donner des estimées précises du vecteur des paramètres, avec une erreur d'estimation proche de la borne de Cramér-Rao pour \mathbf{pem} , et même égale à cette borne pour \mathbf{ml} . Leur principal désavantage est de requérir une procédure d'optimisation numérique difficile, puisqu'à chacune de leur itération une équation de Riccati doit être résolue afin d'obtenir les erreurs de prédiction. Il est par ailleurs à noter que ces méthodes sont sensibles à l'initialisation du vecteur des paramètres : si l'on a seulement accès à une piètre estimée initiale du vecteur des paramètres, la procédure d'optimisation peut conduire à de mauvais résultats, dus à l'existence d'optimiseurs locaux [Söd07].

En théorie il est possible d'appliquer ces méthodes dans le cas de bruits colorés en entrée et en sortie. Toutefois, en plus de l'entrée non-bruitée u_o , il faut alors de surcroît modéliser les bruits \tilde{u} et \tilde{y} , et inclure ces modèles dans le modèle général (2.24). Les problèmes liés à la procédure d'optimisation numérique cités précédemment sont alors accrus.

Notons enfin que les problèmes soulevés par les optimiseurs locaux peuvent néanmoins être évités en mettant en œuvre, lorsque cela est possible, des techniques particulières de collecte des données expérimentales, voir [PW01, PW02] par exemple. ■

2.2.2 Compensation du biais de l'estimateur des moindres carrés (bcls)

Nous nous intéressons ici à la famille de méthodes dites de compensation de biais de l'estimateur des moindres carrés. Ces méthodes sont intéressantes car, sans avoir à recourir à une optimisation numérique et généralement avec un faible coût algorithmique, elles fournissent des estimations dont la précision est proche de celle obtenue avec une méthode à erreur de prédiction. Leur principe est d'estimer le biais, puis de le soustraire des paramètres estimés par moindres carrés. De nombreuses approches utilisant cette technique ont été proposées. L'estimée asymptotique de $\boldsymbol{\theta}$ au sens des moindres carrés est donnée par :

$$\boldsymbol{\theta}_{\text{ls}}^* = \arg \min_{\boldsymbol{\theta}} \bar{\mathbb{E}}\{v(t_k, \boldsymbol{\theta})^2\} = \mathbf{R}_{\varphi\varphi}^{-1} \mathbf{r}_{\varphi y} \quad (2.36)$$

où $v(t_k, \boldsymbol{\theta})$ est l'erreur d'équation du modèle, donnée par :

$$y(t_k) = \boldsymbol{\varphi}^T(t_k) \boldsymbol{\theta} + v(t_k, \boldsymbol{\theta}) \quad (2.37a)$$

$$v(t_k, \boldsymbol{\theta}) = \tilde{y}(t_k) - \tilde{\boldsymbol{\varphi}}^T(t_k) \boldsymbol{\theta} \quad (2.37b)$$

En utilisant (2.37) on obtient :

$$\boldsymbol{\theta}_{\text{ls}}^* = \mathbf{R}_{\varphi\varphi}^{-1} (\mathbf{R}_{\varphi\varphi} \boldsymbol{\theta} + \mathbf{r}_{\varphi v}) \quad (2.38)$$

$$= \boldsymbol{\theta} + \mathbf{R}_{\varphi\varphi}^{-1} (\mathbf{r}_{\tilde{\varphi}\tilde{y}} - \mathbf{R}_{\tilde{\varphi}\tilde{\varphi}} \boldsymbol{\theta}) \quad (2.39)$$

On constate ainsi que l'estimée des moindres carrés est asymptotiquement biaisée, et que le biais dépend des fonctions d'autocovariance des bruits \tilde{u} et \tilde{y} contenues dans $\mathbf{R}_{\tilde{\varphi}\tilde{\varphi}}$ et $\mathbf{r}_{\tilde{\varphi}\tilde{y}}$. Cela

suggère donc une méthode de compensation de biais basée sur l'estimation de ces fonctions. En effet, si ces fonctions sont disponibles, une estimée de $\boldsymbol{\theta}$ asymptotiquement non biaisée est donnée par :

$$\boldsymbol{\theta}_{\text{bcls}}^* = \boldsymbol{\theta}_{\text{ls}}^* - \mathbf{R}_{\varphi\varphi}^{-1} (\mathbf{r}_{\tilde{\varphi}\tilde{y}} - \mathbf{R}_{\tilde{\varphi}\tilde{\varphi}} \boldsymbol{\theta}) \quad (2.40)$$

Bien évidemment, comme l'expression du terme de biais fait aussi intervenir le vecteur des paramètres $\boldsymbol{\theta}$, l'algorithme de compensation de biais doit être itératif. On obtient ainsi la procédure suivante :

1. Initialisation : $\hat{\boldsymbol{\theta}}_{\text{bcls}}^0 = \hat{\boldsymbol{\theta}}_{\text{ls}}$;
2. Estimation de : $\hat{\mathbf{r}}_{\tilde{\varphi}\tilde{y}}^i$ et $\hat{\mathbf{R}}_{\tilde{\varphi}\tilde{\varphi}}^i$;
3. Calcul de :

$$\hat{\boldsymbol{\theta}}_{\text{bcls}}^{i+1} = \hat{\boldsymbol{\theta}}_{\text{ls}} - \hat{\mathbf{R}}_{\varphi\varphi}^{-1} \left(\hat{\mathbf{r}}_{\tilde{\varphi}\tilde{y}}^i - \hat{\mathbf{R}}_{\tilde{\varphi}\tilde{\varphi}}^i \hat{\boldsymbol{\theta}}_{\text{bcls}}^i \right)$$

4. Répéter les étapes 2 et 3 jusqu'à convergence.

Il reste maintenant à estimer les fonctions d'autocovariance des bruits \tilde{u} et \tilde{y} , ce qui représente la clé de voûte de ces algorithmes. En effet, meilleure sera cette estimation, meilleure sera l'estimée de $\boldsymbol{\theta}$ (et plus rapide sera la convergence de l'algorithme). C'est sur ce point que se distinguent les nombreuses approches développées. Nous allons tout d'abord présenter ces diverses méthodes pour le cas de bruits blancs en entrée et en sortie, puis pour le cas d'un bruit blanc en entrée et coloré en sortie.

2.2.2.1 Premier cas : bruit blanc en entrée et en sortie

On suppose que :

Hb1. \tilde{u} et \tilde{y} sont non-corrélés ;

Hb2. $H^{\tilde{u}}(q) = H^{\tilde{y}}(q) = 1$, *i.e.* les bruits additifs \tilde{u} et \tilde{y} sont blancs, de variances respectives $\lambda_{e\tilde{u}}$ et $\lambda_{e\tilde{y}}$ (voir la [figure 2.1](#)) ;

Hs2. $G_o \in \mathcal{G}^*$ et $(H_o^{\tilde{u}}, H_o^{\tilde{y}}) \in \mathcal{H}^*$.

L'équation (2.39) devient :

$$\boldsymbol{\theta}_{\text{ls}}^* = \boldsymbol{\theta} - \mathbf{R}_{\varphi\varphi}^{-1} \begin{bmatrix} \lambda_{e\tilde{y}} \mathbf{I}_{n_a} & 0 \\ 0 & \lambda_{e\tilde{u}} \mathbf{I}_{n_b} \end{bmatrix} \boldsymbol{\theta} \quad (2.41)$$

Le biais est donc entièrement déterminé par les deux variances $\lambda_{e\tilde{u}}$ et $\lambda_{e\tilde{y}}$. Pour les estimer, il faut obtenir (au moins) deux équations supplémentaires qui viendront compléter l'équation (2.40).

La première équation supplémentaire, linéaire en les inconnues $\lambda_{e\tilde{u}}$ et $\lambda_{e\tilde{y}}$, est déduite de l'expression de la corrélation des résidus de la méthode des moindres carrés :

$$J(\boldsymbol{\theta}_{\text{ls}}^*) = \bar{\mathbb{E}}\{v(t_k, \boldsymbol{\theta}_{\text{ls}}^*)^2\} \quad (2.42)$$

$$= \lambda_{e\tilde{y}} \left(1 + \mathbf{a}_{\text{ls}}^{*\text{T}} \mathbf{a} \right) + \lambda_{e\tilde{u}} \mathbf{b}_{\text{ls}}^{*\text{T}} \mathbf{b} \quad (2.43)$$

L'obtention de cette équation est détaillée à la [section B.2](#) de l'[annexe B](#). Toutes les méthodes présentées dans la suite utilisent cette équation. Elles se distinguent en revanche par la deuxième équation utilisée.

Méthode **bels**

La méthode **bels**³ a été introduite par W.X. Zheng. Une première version de la méthode a été donnée dans [Zhe98]. Dans [Zhe99] est exposée la même méthode, avec ceci de différent que les vecteurs sont mieux agencés et des calculs superflus évités, réduisant ainsi son coût algorithmique. Cette dernière version est exposée ici.

La deuxième équation nécessaire à l'estimation des variances des bruits additifs est obtenue en considérant une structure de modèle étendue. Le degré de $A(q, \boldsymbol{\theta})$ est augmenté d'une unité, le nouveau paramètre ainsi introduit ayant une valeur vraie égale à zéro. On définit l'opérateur de transfert du modèle augmenté par :

$$\bar{G}(q, \bar{\boldsymbol{\theta}}) = \frac{B(q^{-1}, \boldsymbol{\theta})}{\bar{A}(q^{-1}, \bar{\boldsymbol{\theta}})} = \frac{b_1 q^{-1} + \dots + b_{n_b} q^{-n_b}}{1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a} + a_{n_a+1} q^{-n_a-1}} \quad (2.44)$$

et le vecteur des paramètres augmenté par :

$$\bar{\boldsymbol{\theta}}^T = [\boldsymbol{\theta}^T \ a_{n_a+1}] \quad (2.45)$$

Bien que l'opérateur de transfert du modèle augmenté $\bar{G}(q, \bar{\boldsymbol{\theta}})$ soit équivalent à l'opérateur de transfert original dans le sens où $a_{n_a+1} = 0$, l'augmentation du degré du dénominateur permet d'obtenir une équation supplémentaire que l'on pourra utiliser pour estimer les variances des bruits.

REMARQUE

Notons qu'il est aussi possible d'augmenter le degré du numérateur $B(q, \boldsymbol{\theta})$. Cela mène néanmoins à des performances amoindries, tant au niveau de la convergence (voir [SHZ05]) qu'au niveau de la précision de l'estimation (voir [HSZ06a]). ■

Écrivons le modèle augmenté sous la forme d'une régression linéaire :

$$y(t_k) = \bar{\boldsymbol{\varphi}}^T(t_k) \bar{\boldsymbol{\theta}} + \tilde{y}(t_k) - \tilde{\boldsymbol{\varphi}}^T(t_k) \bar{\boldsymbol{\theta}} \quad (2.46)$$

$$\bar{\boldsymbol{\varphi}}^T(t_k) = [\boldsymbol{\varphi}^T(t_k) \ -y(t_{k-n_a-1})] \quad (2.47)$$

L'estimée asymptotique au sens des moindres carrés du vecteur augmenté $\bar{\boldsymbol{\theta}}$ s'écrit :

$$\bar{\boldsymbol{\theta}}_{ls}^* = \mathbf{R}_{\bar{\boldsymbol{\varphi}}\bar{\boldsymbol{\varphi}}}^{-1} \mathbf{r}_{\bar{\boldsymbol{\varphi}}y} \quad (2.48)$$

$$= \bar{\boldsymbol{\theta}} - \mathbf{R}_{\bar{\boldsymbol{\varphi}}\bar{\boldsymbol{\varphi}}}^{-1} \begin{bmatrix} \lambda_{e\tilde{y}} \mathbf{I}_{n_a} & 0 & 0 \\ 0 & \lambda_{e\tilde{u}} \mathbf{I}_{n_b} & 0 \\ 0 & 0 & \lambda_{e\tilde{y}} \end{bmatrix} \bar{\boldsymbol{\theta}} \quad (2.49)$$

³Pour *Bias Eliminated Least Squares*.

La forme du vecteur de régression $\bar{\varphi}(t_k)$ permet de décomposer $\mathbf{r}_{\bar{\varphi}y}$ et $\mathbf{R}_{\bar{\varphi}\bar{\varphi}}$ sous la forme :

$$\mathbf{r}_{\bar{\varphi}y} = \begin{bmatrix} \mathbf{r}_{\varphi y} \\ r_y(n_a + 1) \end{bmatrix} \quad (2.50)$$

$$\mathbf{R}_{\bar{\varphi}\bar{\varphi}} = \begin{bmatrix} \mathbf{R}_{\varphi\varphi} & \boldsymbol{\nu} \\ \boldsymbol{\nu}^T & r_y(0) \end{bmatrix} \quad (2.51)$$

avec $\boldsymbol{\nu} = \bar{\mathbb{E}}\{\boldsymbol{\varphi}(t_k)y(t_{k-n_a-1})\}$. L'inverse de cette matrice s'écrit :

$$\mathbf{R}_{\bar{\varphi}\bar{\varphi}}^{-1} = \begin{bmatrix} \mathbf{R}_{\varphi\varphi}^{-1} + \mathbf{R}_{\varphi\varphi}^{-1}\boldsymbol{\nu}\delta\boldsymbol{\nu}^T\mathbf{R}_{\varphi\varphi}^{-1} & -\mathbf{R}_{\varphi\varphi}^{-1}\boldsymbol{\nu}\delta^{-1} \\ -\delta^{-1}\boldsymbol{\nu}^T\mathbf{R}_{\varphi\varphi}^{-1} & \delta^{-1} \end{bmatrix} \quad (2.52)$$

avec $\delta = r_y(0) - \boldsymbol{\nu}^T\mathbf{R}_{\varphi\varphi}^{-1}\boldsymbol{\nu}$. Alors, en injectant les équations (2.50), (2.52) dans la première expression de l'estimateur des moindres carrés (2.48), on obtient :

$$(a_{n_a+1})_{\text{ls}}^* = \delta^{-1}r_y(n_a + 1) - \delta^{-1}\boldsymbol{\nu}^T\mathbf{R}_{\varphi\varphi}^{-1}\mathbf{r}_{\varphi y} \quad (2.53)$$

D'autre part, en injectant ces mêmes équations dans la seconde expression de l'estimateur des moindres carrés (2.49), on obtient :

$$(a_{n_a+1})_{\text{ls}}^* = a_{n_a+1} + \delta^{-1}\boldsymbol{\nu}^T\mathbf{R}_{\varphi\varphi}^{-1}\mathbf{R}_{\bar{\varphi}\bar{\varphi}}\boldsymbol{\theta} - \delta^{-1}\lambda_{e\bar{y}}a_{n_a+1} \quad (2.54)$$

$$= \delta^{-1}\boldsymbol{\nu}^T\mathbf{R}_{\varphi\varphi}^{-1}\mathbf{R}_{\bar{\varphi}\bar{\varphi}}\boldsymbol{\theta} \quad \text{car } a_{n_a+1} = 0 \quad (2.55)$$

En combinant les équations (2.53) et (2.55) on obtient :

$$\boldsymbol{\nu}^T\mathbf{R}_{\varphi\varphi}^{-1}\mathbf{R}_{\bar{\varphi}\bar{\varphi}}\boldsymbol{\theta} = r_y(n_a + 1) - \boldsymbol{\nu}^T\mathbf{R}_{\varphi\varphi}^{-1}\mathbf{r}_{\varphi y} \quad (2.56)$$

$$\iff \boldsymbol{\nu}^T\mathbf{R}_{\varphi\varphi}^{-1} \begin{bmatrix} \lambda_{e\bar{y}}\mathbf{a} \\ \lambda_{e\bar{u}}\mathbf{b} \end{bmatrix} = r_y(n_a + 1) - \boldsymbol{\nu}^T\boldsymbol{\theta}_{\text{ls}}^* \quad (2.57)$$

$$\iff \lambda_{e\bar{y}}\boldsymbol{\nu}^T\mathbf{M}_{n_a}\mathbf{a} + \lambda_{e\bar{u}}\boldsymbol{\nu}^T\mathbf{M}_{n_b}\mathbf{b} = r_y(n_a + 1) - \boldsymbol{\nu}^T\boldsymbol{\theta}_{\text{ls}}^* \quad (2.58)$$

où \mathbf{M}_{n_a} et \mathbf{M}_{n_b} dénotent respectivement les matrices formées en ne prenant que les n_a premières colonnes et les n_b dernières colonnes de la matrice $\mathbf{R}_{\varphi\varphi}^{-1}$:

$$\mathbf{R}_{\varphi\varphi}^{-1} = [\mathbf{M}_{n_a} \quad \mathbf{M}_{n_b}] \quad (2.59)$$

On obtient donc une deuxième équation linéaire en les variances des bruits, et le résultat suivant peut être énoncé :

PROPOSITION 2.2.1 ([ZHE99])

Les variances des bruits $\lambda_{e\bar{u}}$ et $\lambda_{e\bar{y}}$ satisfont

$$\lambda_{e\bar{y}}\boldsymbol{\nu}^T\mathbf{M}_{n_a}\mathbf{a} + \lambda_{e\bar{u}}\boldsymbol{\nu}^T\mathbf{M}_{n_b}\mathbf{b} = r_y(n_a + 1) - \boldsymbol{\nu}^T\boldsymbol{\theta}_{\text{ls}}^* \quad (2.60)$$

où \mathbf{M}_{n_a} , \mathbf{M}_{n_b} dénotent respectivement les matrices formées en ne prenant que les n_a premières colonnes et les n_b dernières colonnes de la matrice $\mathbf{R}_{\varphi\varphi}^{-1}$. ■

L'algorithme **bels** est alors obtenu en utilisant les deux équations (2.43) et (2.60) pour estimer les variances des bruits, et se résume ainsi [Zhe99] :

Étape 1. Initialisation par la méthode des moindres carrés à partir des données disponibles :

$$\hat{\boldsymbol{\theta}}_{\text{bels}}^0 = \hat{\boldsymbol{\theta}}_{\text{ls}}$$

Étape 2. Estimation de $\hat{\lambda}_{e\tilde{y}}^{i+1}$ et $\hat{\lambda}_{e\tilde{u}}^{i+1}$ en résolvant le système linéaire :

$$\begin{bmatrix} J(\hat{\boldsymbol{\theta}}_{\text{ls}}) \\ \hat{r}_y(n_a+1) - \hat{\boldsymbol{v}}^T \hat{\boldsymbol{\theta}}_{\text{ls}} \end{bmatrix} = \begin{bmatrix} 1 + \hat{\boldsymbol{a}}_{\text{ls}}^T \hat{\boldsymbol{a}}_{\text{bels}}^i & \hat{\boldsymbol{b}}_{\text{ls}}^T \hat{\boldsymbol{b}}_{\text{bels}}^i \\ \hat{\boldsymbol{v}}^T \hat{\boldsymbol{M}}_{n_a} \hat{\boldsymbol{a}}_{\text{bels}}^i & \hat{\boldsymbol{v}}^T \hat{\boldsymbol{M}}_{n_b} \hat{\boldsymbol{b}}_{\text{bels}}^i \end{bmatrix} \begin{bmatrix} \hat{\lambda}_{e\tilde{y}}^{i+1} \\ \hat{\lambda}_{e\tilde{u}}^{i+1} \end{bmatrix}$$

Étape 3. Calcul de $\hat{\boldsymbol{\theta}}_{\text{bels}}^{i+1}$:

$$\hat{\boldsymbol{\theta}}_{\text{bels}}^{i+1} = \hat{\boldsymbol{\theta}}_{\text{ls}} + \hat{\boldsymbol{R}}_{\varphi\varphi}^{-1} \begin{bmatrix} \hat{\lambda}_{e\tilde{y}}^{i+1} \boldsymbol{I}_{n_a} & 0 \\ 0 & \hat{\lambda}_{e\tilde{u}}^{i+1} \boldsymbol{I}_{n_b} \end{bmatrix} \hat{\boldsymbol{\theta}}_{\text{bels}}^i$$

Étape 4. Répéter les étapes 2 et 3 jusqu'à ce que la différence entre deux estimées successives satisfasse un critère d'arrêt préalablement choisi.

Les performances de cet algorithme dépendent grandement de la qualité des estimées de $\lambda_{e\tilde{u}}$ et $\lambda_{e\tilde{y}}$. Comme nous venons de le voir, leur estimation est réalisée en résolvant un système linéaire de deux équations faisant intervenir ces deux variances.

Cependant, il appert que dans un certain nombre de cas l'algorithme **bels** requiert un grand nombre d'itérations pour atteindre le critère d'arrêt [Zhe00]. Or il est toujours préférable que la vitesse de convergence d'un algorithme itératif soit grande pour qu'il soit applicable en pratique. En outre, les estimations de $\lambda_{e\tilde{u}}$ et $\lambda_{e\tilde{y}}$ ne sont pas toujours satisfaisantes, notamment quand le rapport signal à bruit (RSB) est faible. La qualité de l'estimation des paramètres en est alors affectée dans le sens où la variance sur les estimées est plus grande, bien qu'elles soient toujours non biaisées [TGG05].

La méthode **abels** (pour **augmented bels**) repose sur l'idée suivante : utiliser non plus une seule mesure supplémentaire de la sortie, mais plusieurs, de manière à avoir plus de deux équations linéaires en $\lambda_{e\tilde{u}}$ et $\lambda_{e\tilde{y}}$. La résolution au sens des moindres carrés du système linéaire surdéterminé ainsi obtenu donnera, en moins d'itérations, une estimation plus satisfaisante des deux variances. De plus, les estimées des variances étant meilleures, les variances sur les paramètres se verront diminuées, et le nouvel algorithme sera plus fiable dans le cas de faibles RSB [Zhe00].

Méthode **abels**

On augmente le degré du dénominateur de $G(q, \boldsymbol{\theta})$ non plus de un, mais d'un nombre $n_{\text{bels}} > 1$, les nouveaux paramètres ainsi introduits ayant une valeur vraie égale à zéro. L'opérateur

de transfert du modèle augmenté et le vecteur des paramètres augmenté sont donnés par :

$$\bar{G}(q, \bar{\theta}) = \frac{B(q^{-1}, \theta)}{\bar{A}(q^{-1}, \bar{\theta})} = \frac{b_1 q^{-1} + \dots + b_{n_b} q^{-n_b}}{1 + a_1 q^{-1} + \dots + a_{n_a + n_{\text{bels}}} q^{-n_a - n_{\text{bels}}}} \quad (2.61)$$

$$\bar{\theta}^T = [\theta^T \ \alpha^T] ; \ \alpha^T = [a_{n_a+1} \ \dots \ a_{n_a+n_{\text{bels}}}] \quad (2.62)$$

Le vecteur de régression s'écrit :

$$\bar{\varphi}^T(t_k) = [\varphi^T(t_k) \ \mu^T(t_k)] \quad (2.63)$$

$$\mu = [-y(t_{k-n_a-1}) \ \dots \ -y(t_{k-n_a-n_{\text{bels}}})]^T \quad (2.64)$$

De même qu'auparavant, où une équation linéaire en $\lambda_{e\tilde{u}}$ et $\lambda_{e\tilde{y}}$ était déduite de l'ajout d'un nouveau paramètre, il est ici possible de déduire n_{bels} équations supplémentaires, comme il est précisé dans la proposition suivante (dont la preuve se trouve à la [section B.3](#) de l'[annexe B](#)) :

PROPOSITION 2.2.2 ([ZHE00])

Les variances des bruits $\lambda_{e\tilde{u}}$ et $\lambda_{e\tilde{y}}$ satisfont

$$\lambda_{e\tilde{y}} \mathbf{R}_{\varphi\mu}^T \mathbf{M}_{n_a} \mathbf{a} + \lambda_{e\tilde{u}} \mathbf{R}_{\varphi\mu}^T \mathbf{M}_{n_b} \mathbf{b} = \mathbf{r}_{\mu y} - \mathbf{R}_{\varphi\mu}^T \theta_{ls}^* \quad (2.65)$$

où \mathbf{M}_{n_a} , \mathbf{M}_{n_b} représentent respectivement les matrices formées en ne prenant que les n_a premières colonnes et les n_b dernières colonnes de la matrice $\mathbf{R}_{\varphi\varphi}^{-1}$. ■

En utilisant les équations (2.43) et (2.65), on obtient désormais un système de $(n_{\text{bels}} + 1)$ équations linéaires. La résolution au sens des moindres carrés de ce système surdéterminé donne alors une estimée des deux variances $\lambda_{e\tilde{u}}$ et $\lambda_{e\tilde{y}}$. L'algorithme **abels** peut être vu comme une extension de la méthode **bels** : lorsque $n_{\text{bels}} = 1$, l'équation (2.65) est identique à l'équation (2.60).

Le nombre n_{bels} est à choisir par l'utilisateur, mais le meilleur compromis entre temps de calcul et qualité des estimées semble obtenu pour $n_{\text{bels}} = 2$ ou 3 ; d'après différents tests pratiques, de plus grandes valeurs n'offrent pas de différence significative en terme de précision des estimées, tandis que le temps de calcul est accru [Zhe00].

Méthode **fbcls**

Cette méthode diffère de la méthode **bels** en cela qu'elle ne considère pas une structure de modèles étendue pour obtenir la seconde équation nécessaire à l'estimation des variances des bruits. Cette seconde équation est en effet déduite des résidus de la méthode des moindres carrés appliquée aux données filtrées [IKYW05].

Introduisons un filtre $\alpha(q^{-1}) = \sum_{i=0}^{n_\alpha} \alpha_i q^{-i}$ avec $n_\alpha \geq \max(n_a, n_b)$ que l'on applique aux deux membres de l'équation (2.37a). L'estimateur des moindres carrés appliqué aux données

filtrées est donné par⁴ :

$$\boldsymbol{\theta}_{\text{fls}}^* = \arg \min_{\boldsymbol{\theta}} \bar{\mathbb{E}}\{v_f(t_k, \boldsymbol{\theta})^2\} = \mathbf{R}_{\varphi_f \varphi_f}^{-1} \mathbf{r}_{\varphi_f y_f} \quad (2.66)$$

$$v_f(t_k, \boldsymbol{\theta}) = y_f(t_k) - \boldsymbol{\varphi}_f^T(t_k) \boldsymbol{\theta} \quad (2.67)$$

Introduisons la notation suivante : $\mathcal{I}_n(k)$ est la matrice carrée de dimension n ayant la k^e diagonale inférieure et la k^e diagonale supérieure composées de 1, et des zéros partout ailleurs (par convention $\mathcal{I}_n(0) = \mathbf{I}_n$). Alors, de l'examen de la corrélation des résidus de la méthode des moindres carrés appliquée aux données filtrées, on déduit (voir la [section B.5](#) de l'[annexe B](#)) :

PROPOSITION 2.2.3 ([IKYW05])

Les variances des bruits $\lambda_{e\bar{u}}$ et $\lambda_{e\bar{y}}$ satisfont

$$\begin{aligned} J(\boldsymbol{\theta}_{\text{fls}}^*) &= \bar{\mathbb{E}}\{v_f(t_k, \boldsymbol{\theta}_{\text{fls}}^*)^2\} \\ &= \lambda_{e\bar{y}} \left(\sum_{i=0}^{n_\alpha} \alpha_i^2 + \mathbf{v}^T (\mathbf{a} + \mathbf{a}_{\text{fls}}^*) + \mathbf{a}^T \mathbf{H}_{n_a} \mathbf{a}_{\text{fls}}^* \right) + \lambda_{e\bar{u}} \mathbf{b}^T \mathbf{H}_{n_b} \mathbf{b}_{\text{fls}}^* \end{aligned} \quad (2.68)$$

où l'on note :

$$\mathbf{H}_{n_a} = \sum_{j=0}^{n_a-1} \sum_{i=0}^{n_\alpha-j} \alpha_i \alpha_{i+j} \mathcal{I}_{n_a}(j) \quad \mathbf{H}_{n_b} = \sum_{j=0}^{n_b-1} \sum_{i=0}^{n_\alpha-j} \alpha_i \alpha_{i+j} \mathcal{I}_{n_b}(j)$$

■

L'équation (2.68) en conjonction avec l'équation (2.43), permet alors d'obtenir une estimée des variances des deux bruits blancs [IKYW05].

REMARQUE

L'équation (2.68) peut être vue comme une généralisation de l'équation (2.43). En effet, si le filtre $\alpha(q^{-1})$ est égal à 1, on retrouve bien l'expression de la corrélation des résidus de la méthode des moindres carrés, puisqu'alors $\sum_{i=0}^{n_\alpha} \alpha_i^2 = 1$, $\mathbf{v} = 0$, $\mathbf{H}_{n_a} = \mathbf{I}_{n_a}$ et $\mathbf{H}_{n_b} = \mathbf{I}_{n_b}$. ■

Méthode pbcls

Cette méthode utilise une estimée du vecteur des paramètres $\boldsymbol{\theta}$ obtenue à l'aide d'un prédicteur à un pas de la sortie [JIJW01]. Cette estimée est donnée par :

$$\boldsymbol{\theta}_{\text{pbcls}}^* = \bar{\mathbb{E}}\{\boldsymbol{\varphi}(t_{k-1}) \boldsymbol{\varphi}^T(t_{k-1})\}^{-1} \bar{\mathbb{E}}\{\boldsymbol{\varphi}(t_{k-1}) y(t_k)\} \quad (2.69)$$

$$= \mathbf{R}_{\varphi \varphi}^{-1} \mathbf{r}_{\varphi(t_{k-1}) y(t_k)} \quad (2.70)$$

Les résidus de cette méthode s'écrivent :

$$v(t_k, \boldsymbol{\theta}_{\text{pbcls}}^*) = y(t_k) - \boldsymbol{\varphi}^T(t_{k-1}) \boldsymbol{\theta}_{\text{pbcls}}^* \quad (2.71)$$

⁴Les variables filtrées sont notées avec un f en indice.

et vérifient de manière triviale :

$$\bar{\mathbb{E}}\{\varphi(t_{k-1})v(t_k, \boldsymbol{\theta}_{pbcls}^*)\} = 0 \quad (2.72)$$

L'examen de la corrélation entre les résidus de la méthode des moindres carrés et de ceux du prédicteur à un pas de la sortie permet d'énoncer le résultat suivant (voir la [section B.4](#) de l'[annexe B](#)) :

PROPOSITION 2.2.4 ([JIJW01])

Les variances des bruits $\lambda_{e\tilde{u}}$ et $\lambda_{e\tilde{y}}$ satisfont

$$\bar{\mathbb{E}}\{v(t_{k-1}, \boldsymbol{\theta}_{ls}^*)v(t_k, \boldsymbol{\theta}_{pbcls}^*)\} = \lambda_{e\tilde{y}} \mathbf{a}^T \mathbf{a}_{pbcls}^* + \lambda_{e\tilde{u}} \mathbf{b}^T \mathbf{b}_{pbcls}^* \quad (2.73)$$

■

L'équation (2.73) en conjonction avec l'équation (2.43), permet alors d'obtenir une estimée des variances des deux bruits blancs [JIJW01].

2.2.2.2 Deuxième cas : bruit blanc en entrée et coloré en sortie

On suppose que :

Hb1. \tilde{u} et \tilde{y} sont non-corrélés ;

Hb4. $H^{\tilde{u}}(q) = 1, H^{\tilde{y}}(q) \neq 1$, *i.e.* le bruit en entrée du système est blanc, tandis que le bruit en sortie est coloré (voir la [figure 2.1](#)) ;

Hs2. $G_o \in \mathcal{G}^*$ et $(H_o^{\tilde{u}}, H_o^{\tilde{y}}) \in \mathcal{H}^*$.

L'équation (2.39) devient :

$$\boldsymbol{\theta}_{ls}^* = \boldsymbol{\theta} - \mathbf{R}_{\varphi\varphi}^{-1} \left\{ \begin{bmatrix} \mathbf{r}_{\tilde{y}\tilde{y}} \\ 0 \end{bmatrix} + \begin{bmatrix} \mathbf{R}_{\tilde{y}\tilde{y}} & 0 \\ 0 & \lambda_{e\tilde{u}} \mathbf{I}_{n_b} \end{bmatrix} \boldsymbol{\theta} \right\} \quad (2.74)$$

$$= \boldsymbol{\theta} - \mathbf{R}_{\varphi\varphi}^{-1} \begin{bmatrix} \mathbf{r}_{\tilde{y}\tilde{y}} + \mathbf{R}_{\tilde{y}\tilde{y}} \boldsymbol{\theta} \\ \lambda_{e\tilde{u}} \mathbf{b} \end{bmatrix} \quad (2.75)$$

Peu de méthodes de compensation de biais ont été étendues pour être applicables dans le cas d'un bruit blanc en entrée et d'un bruit coloré en sortie. La méthode **bels** est une notable exception. Elle est présentée dans la suite.

REMARQUE

Les deux cas « bruit coloré en entrée/blanc en sortie » et « blanc en entrée/coloré en sortie » peuvent être traités de la même manière. Une méthode applicable dans un cas peut être modifiée pour être applicable dans l'autre. Il faut toutefois noter que, d'un point de vue pratique, le deuxième cas est plus réaliste. Le bruit sur le signal de sortie peut en effet représenter non seulement le bruit de mesure ou de capteur, mais aussi les erreurs de modélisation. ■

Méthode **bels2**

Le biais asymptotique de la méthode des moindres carrés est donc déterminé par les propriétés statistiques des bruits additifs sur l'entrée et la sortie du système : la variance $\lambda_{\tilde{u}}$, le vecteur de covariance $\mathbf{r}_{\tilde{y}\tilde{y}}$ et la matrice de covariance $\mathbf{R}_{\tilde{y}\tilde{y}}$. La première étape consiste à regrouper toutes les inconnues à estimer en sus du vecteur des paramètres au sein d'un seul vecteur $\boldsymbol{\rho}_{\tilde{u},\tilde{y}}$:

$$\boldsymbol{\rho}_{\tilde{u},\tilde{y}}^{\text{T}} = [\lambda_{e\tilde{y}} \mathbf{r}_{\tilde{y}\tilde{y}}^{\text{T}} \lambda_{e\tilde{u}}] = [r_{\tilde{y}}(0) r_{\tilde{y}}(1) \dots r_{\tilde{y}}(n_a) \lambda_{e\tilde{u}}] \quad (2.76)$$

Il est alors possible d'écrire :

$$\begin{bmatrix} \mathbf{r}_{\tilde{y}\tilde{y}} + \mathbf{R}_{\tilde{y}\tilde{y}}\mathbf{a} \\ \lambda_{e\tilde{u}}\mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{M}_1 + \mathbf{M}_2(\mathbf{a}) \\ \mathbf{M}_3(\mathbf{b}) \end{bmatrix} \boldsymbol{\rho}_{\tilde{u},\tilde{y}} \quad (2.77)$$

où l'on note :

- $\mathbf{M}_1 = [0 \ \mathbf{I}_{n_a} \ 0] \in \mathbb{R}^{n_a \times (n_a+2)}$;
- $\mathbf{M}_2(\mathbf{a}) = \sum_{k=1}^{n_a} \mathcal{I}_{n_a}(k) \mathbf{a} \boldsymbol{\tau}_k^{\text{T}} \in \mathbb{R}^{n_a \times (n_a+2)}$;
- $\mathbf{M}_3(\mathbf{b}) = \mathbf{b} \boldsymbol{\tau}_{n_a+2}^{\text{T}} \in \mathbb{R}^{n_b \times (n_a+2)}$;
- $\boldsymbol{\tau}_k \in \mathbb{R}^{(n_a+2) \times 1}$ le vecteur dont la k^e composante vaut 1 et toutes les autres 0 ;
- $\mathcal{I}_n(k)$ la matrice carrée de dimension n ayant la k^e diagonale supérieure et la k^e diagonale inférieure composées de 1, et des zéros partout ailleurs, avec par convention $\mathcal{I}_n(0) = \mathbf{I}_n$ (ces matrices sont utilisées pour réécrire les matrices de Toeplitz symétriques $\mathbf{R}_{\tilde{y}\tilde{y}}$ et $\mathbf{R}_{\tilde{u}\tilde{u}}$).

Si l'on note :

$$\mathbf{M}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{M}_1 + \mathbf{M}_2(\mathbf{a}) \\ \mathbf{M}_3(\mathbf{b}) \end{bmatrix} \quad (2.78)$$

alors l'équation (2.75) devient :

$$\boldsymbol{\theta}_{\text{ls}}^* = \boldsymbol{\theta} - \mathbf{R}_{\varphi\varphi}^{-1} \mathbf{M}(\boldsymbol{\theta}) \boldsymbol{\rho}_{\tilde{u},\tilde{y}} \quad (2.79)$$

Pour la méthode **bels**, l'extension au cas d'un bruit coloré en sortie se fait plutôt naturellement. En effet, au lieu des deux variances des bruits blancs, il faut désormais estimer les n_a+2 corrélations contenues dans le vecteur $\boldsymbol{\rho}_{\tilde{u},\tilde{y}}$. Il faut donc obtenir plus de deux équations linéaires en les inconnues, ce qui est réalisé justement par la méthode **abels**, présentée auparavant.

Tout d'abord, pour un système avec bruit blanc en entrée et bruit coloré en sortie, la corrélation des résidus de la méthode des moindres carrés s'écrit (voir la [section B.2](#) de

l'annexe B) :

$$J(\boldsymbol{\theta}_{\text{ls}}^*) = \lambda_{e\tilde{y}} + \boldsymbol{\theta}_{\text{ls}}^{*\text{T}} (\mathbf{R}_{\tilde{\varphi}\tilde{\varphi}} \boldsymbol{\theta} - \mathbf{r}_{\tilde{\varphi}\tilde{y}}) - \boldsymbol{\theta}^{\text{T}} \mathbf{r}_{\tilde{\varphi}\tilde{y}} \quad (2.80)$$

$$= \lambda_{e\tilde{y}} + \boldsymbol{\theta}_{\text{ls}}^{*\text{T}} \begin{bmatrix} \mathbf{r}_{\tilde{y}\tilde{y}} + \mathbf{R}_{\tilde{y}\tilde{y}} \mathbf{a} \\ \lambda_{e\tilde{u}} \mathbf{b} \end{bmatrix} + \boldsymbol{\theta}^{\text{T}} \begin{bmatrix} \mathbf{r}_{\tilde{y}\tilde{y}} \\ 0 \end{bmatrix} \quad (2.81)$$

$$= \underbrace{\left\{ \boldsymbol{\tau}_1^{\text{T}} + \boldsymbol{\theta}_{\text{ls}}^{*\text{T}} \mathbf{M}(\boldsymbol{\theta}) + \mathbf{a}^{\text{T}} \mathbf{M}_1 \right\}}_{=\mathbf{q}(\boldsymbol{\theta}_{\text{ls}}^*, \boldsymbol{\theta})} \boldsymbol{\rho}_{\tilde{u}, \tilde{y}} \quad (2.82)$$

Pour obtenir les autres équations linéaires nécessaires, on augmente le degré du *numérateur* de $G(q, \boldsymbol{\theta})$ d'un nombre n_{bels} , les nouveaux paramètres ainsi introduits ayant une valeur vraie égale à zéro.

REMARQUE

On a vu qu'augmenter le degré du dénominateur fait apparaître des mesures supplémentaires de la sortie dans le vecteur de régression augmenté (voir l'équation (2.63), par exemple). De la même manière, augmenter le degré du numérateur fait apparaître des mesures supplémentaires de l'entrée. Celle-ci étant contaminée par un bruit blanc, il est logique de l'utiliser plutôt que la sortie, puisque, comme on le verra dans la suite, cela évite ainsi de faire apparaître la fonction d'autocorrélation du bruit coloré. ■

L'opérateur de transfert du modèle augmenté est donné par :

$$\bar{G}(q, \bar{\boldsymbol{\theta}}) = \frac{\bar{B}(q^{-1}, \bar{\boldsymbol{\theta}})}{A(q^{-1}, \boldsymbol{\theta})} = \frac{b_1 q^{-1} + \dots + b_{n_b+n_{\text{bels}}} q^{-n_b-n_{\text{bels}}}}{1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a}} \quad (2.83)$$

$$\bar{\boldsymbol{\theta}}^{\text{T}} = [\boldsymbol{\theta}^{\text{T}} \quad \boldsymbol{\beta}^{\text{T}}] \quad (2.84)$$

$$\boldsymbol{\beta}^{\text{T}} = [b_{n_b+1} \quad \dots \quad b_{n_b+n_{\text{bels}}}] = [0 \quad \dots \quad 0] \quad (2.85)$$

On écrit le modèle augmenté sous la forme d'une régression linéaire :

$$y(t_k) = \tilde{\boldsymbol{\varphi}}^{\text{T}}(t_k) \bar{\boldsymbol{\theta}} + v(t_k, \bar{\boldsymbol{\theta}}) \quad (2.86)$$

$$v(t_k, \bar{\boldsymbol{\theta}}) = \tilde{y}(t_k) - \tilde{\boldsymbol{\varphi}}^{\text{T}}(t_k) \bar{\boldsymbol{\theta}} \quad (2.87)$$

$$\tilde{\boldsymbol{\varphi}}^{\text{T}}(t_k) = [\boldsymbol{\varphi}^{\text{T}}(t_k) \quad \boldsymbol{\mu}^{\text{T}}(t_k)] \quad (2.88)$$

$$\boldsymbol{\mu}(t_k) = [u(t_{k-n_b-1}) \quad \dots \quad u(t_{k-n_b-n_{\text{bels}}})]^{\text{T}} \quad (2.89)$$

L'estimée au sens des moindres carrés pour le modèle augmenté est donnée par :

$$\bar{\boldsymbol{\theta}}_{\text{ls}}^* = \mathbf{R}_{\tilde{\boldsymbol{\varphi}}\tilde{\boldsymbol{\varphi}}}^{-1} \mathbf{r}_{\tilde{\boldsymbol{\varphi}}\tilde{y}} \quad (2.90)$$

$$= \bar{\boldsymbol{\theta}} + \mathbf{R}_{\tilde{\boldsymbol{\varphi}}\tilde{\boldsymbol{\varphi}}}^{-1} \left\{ \begin{bmatrix} \mathbf{r}_{\tilde{\boldsymbol{\varphi}}\tilde{y}} \\ \mathbf{r}_{\tilde{\boldsymbol{\mu}}\tilde{y}} \end{bmatrix} - \begin{bmatrix} \mathbf{R}_{\tilde{\boldsymbol{\varphi}}\tilde{\boldsymbol{\varphi}}} & \mathbf{R}_{\tilde{\boldsymbol{\mu}}\tilde{\boldsymbol{\varphi}}}^{\text{T}} \\ \mathbf{R}_{\tilde{\boldsymbol{\mu}}\tilde{\boldsymbol{\varphi}}} & \lambda_{e\tilde{u}} \mathbf{I}_{n_{\text{bels}}} \end{bmatrix} \bar{\boldsymbol{\theta}} \right\} \quad (2.91)$$

$$= \bar{\boldsymbol{\theta}} + \mathbf{R}_{\tilde{\boldsymbol{\varphi}}\tilde{\boldsymbol{\varphi}}}^{-1} \left\{ \begin{bmatrix} \mathbf{r}_{\tilde{\boldsymbol{\varphi}}\tilde{y}} \\ 0 \end{bmatrix} - \begin{bmatrix} \mathbf{R}_{\tilde{\boldsymbol{\varphi}}\tilde{\boldsymbol{\varphi}}} & 0 \\ 0 & \lambda_{e\tilde{u}} \mathbf{I}_{n_{\text{bels}}} \end{bmatrix} \bar{\boldsymbol{\theta}} \right\} \quad (2.92)$$

Ici apparaît clairement l'avantage d'augmenter le degré du numérateur et non du dénominateur : comme $\boldsymbol{\mu}(t_k)$ contient des versions retardées de $u(t_k)$ on a $\mathbf{r}_{\mu\tilde{y}} = 0$ et $\mathbf{R}_{\tilde{\mu}\tilde{\varphi}} = 0$ de par la blancheur de \tilde{u} et son indépendance de \tilde{y} .

REMARQUE

On constate par ailleurs que l'extension de cette méthode au cas de bruits colorés en entrée et en sortie ne peut pas se faire de cette manière, puisqu'augmenter le degré du numérateur ou du dénominateur pour obtenir des équations supplémentaires ajoute en même temps des inconnues à estimer. ■

Puis, de la même manière qu'auparavant, en exploitant la forme de la matrice $\mathbf{R}_{\tilde{\varphi}\tilde{\varphi}}$ on déduit n_{bels} équations linéaires en $\boldsymbol{\rho}_{\tilde{u},\tilde{y}}$, qui s'ajoutent à l'équation (2.82) pour obtenir le résultat suivant (dont la preuve se trouve à la section B.3 de l'annexe B) :

PROPOSITION 2.2.5 ([ZHE02])

Le vecteur $\boldsymbol{\rho}_{\tilde{u},\tilde{y}}$ satisfait le système linéaire :

$$\begin{bmatrix} \mathbf{R}_{\varphi\mu}^T \mathbf{R}_{\varphi\varphi}^{-1} \mathbf{M}(\boldsymbol{\theta}) \\ \mathbf{q}(\boldsymbol{\theta}_{ls}^*, \boldsymbol{\theta}) \end{bmatrix} \boldsymbol{\rho}_{\tilde{u},\tilde{y}} = \begin{bmatrix} \mathbf{r}_{\mu y} - \mathbf{R}_{\varphi\mu}^T \boldsymbol{\theta}_{ls}^* \\ J(\boldsymbol{\theta}_{ls}^*) \end{bmatrix} \quad (2.93)$$

■

On obtient donc un système de $(n_{\text{bels}} + 1)$ équations linéaires en les $(n_a + 2)$ inconnues contenues dans le vecteur $\boldsymbol{\rho}_{\tilde{u},\tilde{y}}$. Pour que ce système ait une unique solution, on fixe alors $n_{\text{bels}} = n_a + 1$ [Zhe02].

Méthode abels2

Il semble cependant que la méthode **bels2** ne fournisse pas toujours des estimées satisfaisantes, notamment lorsque le RSB est faible. Aussi, sur le même principe que la méthode **abels** présentée auparavant, il est possible d'étendre la méthode **bels2** : en augmentant le degré du numérateur de $n_{\text{bels}} > n_a + 1$, on obtient alors plus d'équations qu'il n'est nécessaire. On résout alors au sens des moindres carrés le système linéaire surdéterminé ainsi obtenu pour obtenir une meilleure estimation du vecteur des corrélations $\boldsymbol{\rho}_{\tilde{u},\tilde{y}}$ [TGG08b].

2.2.3 Méthodes fondées sur une variable instrumentale (iv et xiv)

Ces méthodes sont souvent utilisées car elles donnent lieu à des estimateurs simples (d'un point de vue algorithmique) pour remplacer l'estimateur des moindres carrés, biaisé. Le principe de ces méthodes remonte à [Rei41, Rei50]. Une étude circonstanciée de l'application de ces estimateurs à l'identification des systèmes est donnée dans [SS83, You84].

Rappelons brièvement le principe de ces méthodes. L'idée de la variable instrumentale est

d'utiliser un vecteur d'instruments $\mathbf{z}(t_k)$, de même dimension que $\boldsymbol{\theta}$, ayant les propriétés suivantes :

- il est corrélé avec le vecteur de régression $\boldsymbol{\varphi}(t_k)$ (voir (2.7a)) ;
- il est non-corrélé avec l'erreur d'équation $v(t_k)$ (voir (2.7b)).

On a alors :

$$0 = \bar{\mathbb{E}}\{\mathbf{z}(t_k)v(t_k)\} = \bar{\mathbb{E}}\{\mathbf{z}(t_k)(y(t_k) - \boldsymbol{\varphi}^T(t_k)\boldsymbol{\theta})\} = \mathbf{r}_{zy} - \mathbf{R}_{z\varphi}\boldsymbol{\theta} \quad (2.94)$$

L'estimateur de la variable instrumentale est donc défini par :

$$\boldsymbol{\theta}_{iv}^* = \mathbf{R}_{z\varphi}^{-1}\mathbf{r}_{zy} \quad (2.95)$$

si l'inverse de $\mathbf{R}_{z\varphi}$ existe. L'estimateur de la variable instrumentale *étendue*⁵ est obtenu en généralisant (2.95) dans deux directions : d'une part en autorisant le vecteur des instruments $\mathbf{z}(t_k)$ à être d'une dimension supérieure à celle de $\boldsymbol{\theta}$, et d'autre part en permettant de pré-filtrer les données d'entrée/sortie disponibles [SS89]. L'estimateur de la variable instrumentale étendue est donné par :

$$\boldsymbol{\theta}_{xiv}^* = \arg \min_{\boldsymbol{\theta}} \|\mathbf{R}_{z\varphi_f}\boldsymbol{\theta} - \mathbf{r}_{zy_f}\|_Q^2 \quad (2.96)$$

où l'on note :

$$\boldsymbol{\varphi}_f(t_k) = F(q)\boldsymbol{\varphi}(t_k) \quad (2.97)$$

$$y_f(t_k) = F(q)y(t_k) \quad (2.98)$$

$$\|\mathbf{x}\|_Q^2 = \mathbf{x}^T\mathbf{Q}\mathbf{x} \quad (2.99)$$

et où $F(q)$ est l'opérateur de transfert d'un filtre asymptotiquement stable et \mathbf{Q} une matrice de pondération définie positive. Notons que (2.96) n'est autre que la solution au sens des moindres carrés pondérés d'un système linéaire surdéterminé.

Le choix des instruments peut se faire de diverses manières. Parmi les choix classiques, on trouve des entrées (filtrées) ou encore des entrées et/ou des sorties retardées [SS81a]. Bien sûr, le cas du modèle EIV est rendu plus difficile par le fait que toutes les données sont corrélées avec l'un ou l'autre des bruits. Il est toutefois possible de modifier l'approche classique pour que la méthode s'applique au modèle EIV. Le choix des instruments dépend alors du modèle de bruit considéré. Supposons par exemple que :

Hb1. \tilde{u} et \tilde{y} sont non-corrélés ;

Hb2. $H^{\tilde{u}}(q) = 1$ et $H^{\tilde{y}}(q) = 1$, c'est-à-dire les bruits additifs \tilde{u} et \tilde{y} sont blancs, de variances respectives $\lambda_{e\tilde{u}}$ et $\lambda_{e\tilde{y}}$;

Hs2. $G_o \in \mathcal{G}^*$ et $(H_o^{\tilde{u}}, H_o^{\tilde{y}}) \in \mathcal{H}^*$.

⁵que l'on abrégera par *xiv* pour *eXtended Instrumental Variable*.

On peut alors définir le vecteur des instruments des manières suivantes :

$$\mathbf{z}_u^T(t_k) = [u(t_{k-d-1}) \dots u(t_{k-d-n_{z_u}})] \quad (2.100)$$

$$\mathbf{z}_y^T(t_k) = [y(t_{k-d-1}) \dots y(t_{k-d-n_{z_y}})] \quad (2.101)$$

$$\mathbf{z}_{yu}^T(t_k) = [y(t_{k-d_1-1}) \dots y(t_{k-d_1-n_{z_y}}) u(t_{k-d_2-1}) \dots u(t_{k-d_2-n_{z_u}})] \quad (2.102)$$

où d , d_1 et d_2 sont des délais dont les valeurs dépendent de l'ordre du système. Le vecteur des instruments \mathbf{z}_u est par exemple utilisé dans [SM02] avec $d = n_b$. Il semble donc y avoir une relativement grande latitude dans le choix du vecteur des intruments. Il faut néanmoins veiller à ce que le choix ne rende pas la matrice à inverser $\mathbf{R}_{z\varphi}$ singulière ou mal conditionnée ; c'est pourquoi on dit d'ordinaire qu'il est préférable de prendre un vecteur d'instruments « le plus corrélé possible » avec le vecteur de régression $\varphi(t_k)$.

La convergence et la précision asymptotique des estimateurs de la variable instrumentale sont étudiées dans [SS81a], dans le cas « classique » des modèles avec un signal d'entrée supposé parfaitement connu. Une comparaison entre l'estimateur des moindres carrés totaux et un estimateur de la variable instrumentale est également disponible dans [SM02]. Notons par ailleurs qu'il a été montré que la compensation du biais de l'estimateur des moindres carrés peut être vue comme une méthode fondée sur une variable instrumentale pour le cas de l'identification de modèles en boucle ouverte [SZS99, JJYW02] et en boucle fermée [GVdH01, GVdH05]. Il semble qu'une telle étude n'ait pas été réalisée pour le modèle EIV⁶. Dans la suite nous aurons besoin d'une notation supplémentaire, tenant compte de la dimension variable du vecteur des instruments :

$$\mathbf{u}^T(t_k, n_{z_u}) = [u(t_{k-1}) \dots u(t_{k-n_{z_u}})] \quad (2.103)$$

$$\mathbf{y}^T(t_k, n_{z_y}) = [y(t_{k-1}) \dots y(t_{k-n_{z_y}})] \quad (2.104)$$

Il est alors possible d'écrire par exemple le vecteur d'instruments \mathbf{z}_{yu} donné en (2.102) de manière compacte :

$$\mathbf{z}_{yu}^T(t_k) = [\mathbf{y}^T(t_{k-d_1}, n_{z_y}) \mathbf{u}^T(t_{k-d_2}, n_{z_u})] \quad (2.105)$$

2.2.3.1 Cas d'un bruit blanc en entrée et coloré en sortie

Les hypothèses sur les bruits sont :

Hb1. \tilde{u} et \tilde{y} sont non-corrélés ;

Hb3. $H^{\tilde{u}}(q) = 1$ et $H^{\tilde{y}}(q) = 1 + c_1^{\tilde{y}} q^{-1} + \dots + c_{n_c}^{\tilde{y}} q^{-n_c}$, c'est-à-dire \tilde{u} est blanc, de variance $\lambda_{e^{\tilde{u}}}$, tandis que \tilde{y} est obtenu en sortie d'un filtre MA ;

Hs2. $G_o \in \mathcal{G}^*$ et $(H_o^{\tilde{u}}, H_o^{\tilde{y}}) \in \mathcal{H}^*$.

⁶Il faut toutefois noter que certaines similitudes entre la méthode dels et la méthode de la variable instrumentale sont mentionnées dans [HSZ06b].

Dans la suite, les estimateurs de la variable instrumentale étendue considérés ne feront pas intervenir de pré-filtrage (donc $F(q^{-1}) = 1$), et il n'y aura pas de pondération (donc $\mathbf{Q} = \mathbf{I}$). En utilisant (2.7a), l'estimateur de la variable instrumentale étendue s'écrit :

$$\boldsymbol{\theta}_{\text{xiv}}^* = \mathbf{R}_{z\varphi}^\dagger \mathbf{r}_{zy} = (\mathbf{R}_{z\varphi}^\top \mathbf{R}_{z\varphi})^{-1} \mathbf{R}_{z\varphi}^\top \mathbf{r}_{zy} \quad (2.106)$$

$$= \boldsymbol{\theta} + (\mathbf{R}_{z\varphi}^\top \mathbf{R}_{z\varphi})^{-1} \mathbf{r}_{zv} \quad (2.107)$$

où $(\cdot)^\dagger$ dénote la pseudo-inverse d'une matrice. Pour que l'estimateur (2.106) soit non biaisé, le vecteur des instruments $\mathbf{z}(t_k)$ doit être non-corrélé avec $v(t_k)$, et « aussi corrélé que possible » avec le vecteur de régression $\boldsymbol{\varphi}(t_k)$, de sorte que la matrice à inverser (2.106) ne soit pas mal conditionnée, ce que l'on résume par :

$$\mathbf{r}_{zv} = 0 \quad (2.108)$$

$$\mathbf{R}_{z\varphi} \text{ est non-singulière} \quad (2.109)$$

Comme nous l'avons vu précédemment, il n'y a pas de manière unique de définir $\mathbf{z}(t_k)$, et dans la suite deux manières différentes de le faire sont présentées.

Première approche : méthode xiv1

La première approche a été explorée dans [SM02], considérant le cas de bruits blancs en entrée et en sortie. Puisque le vecteur d'instruments choisi est uniquement composé d'entrées retardées cette approche peut donc également être utilisée lorsque le bruit en sortie est coloré (si \tilde{u} et \tilde{y} sont non-corrélés) :

$$\mathbf{z}^\top(t_k) = \mathbf{u}^\top(t_{k-d_1}, n_{z_u}) = [u(t_{k-1-d_1}) \dots u(t_{k-d_1-n_{z_u}})] \quad (2.110)$$

Puisque \tilde{u} , \tilde{y} et u_o sont non-corrélés on a :

$$\mathbf{r}_{zv} = \bar{\mathbb{E}}\{\mathbf{z}(t_k)v(t_k)\} \quad (2.111)$$

$$= \bar{\mathbb{E}}\left\{\mathbf{z}(t_k) \left(\tilde{\mathbf{y}}^\top(t_k) \bar{\mathbf{a}} - \tilde{\mathbf{u}}^\top(t_k) \mathbf{b} \right)\right\} \quad (2.112)$$

$$= -\bar{\mathbb{E}}\{\tilde{\mathbf{z}}(t_k) \tilde{\mathbf{u}}^\top(t_k)\} \mathbf{b} \quad (2.113)$$

Ainsi, si le délai vérifie $d_1 \geq n_b$, le vecteur d'instruments $\mathbf{z}(t_k)$ satisfait à la condition (2.108). Notons enfin que la matrice de covariance asymptotique de l'estimateur $\boldsymbol{\theta}_{\text{xiv1}}^*$ est également donnée dans [SM02].

Seconde approche : méthode xiv2

La méthode xiv1 satisfait aux conditions (2.108)–(2.109). Toutefois, le conditionnement de la matrice $\mathbf{R}_{z\varphi}$ peut être amélioré en choisissant un vecteur d'instruments basé non seulement sur des entrées retardées, mais utilisant également le signal de sortie. Bien que cette idée soit classique, il semble que les travaux sur l'identification de modèles EIV à l'aide d'une variable instrumentale ne fassent intervenir que des vecteurs d'instruments composés d'entrées

retardées. Nous proposons donc d'utiliser un vecteur d'instruments contenant des entrées retardées et des sorties retardées [TGG08b] :

$$\mathbf{z}^T(t_k) = [\mathbf{y}^T(t_{k-d_2}, n_{z_y}) \mathbf{u}^T(t_{k-d_3}, n_{z_u})] \quad (2.114)$$

Puisque \tilde{u} , \tilde{y} et u_o sont non-corrélés on a :

$$\mathbf{r}_{zv} = \bar{E}\{\mathbf{z}(t_k)v(t_k)\} \quad (2.115)$$

$$= \bar{E} \left\{ \begin{array}{l} -\mathbf{y}(t_{k-d_2}, n_{z_y}) \left(\tilde{\mathbf{y}}^T(t_k)\bar{\mathbf{a}} - \tilde{\mathbf{u}}^T(t_k)\mathbf{b} \right) \\ \mathbf{u}(t_{k-d_3}, n_{z_u}) \left(\tilde{\mathbf{y}}^T(t_k)\bar{\mathbf{a}} - \tilde{\mathbf{u}}^T(t_k)\mathbf{b} \right) \end{array} \right\} \quad (2.116)$$

$$= \bar{E} \left\{ \begin{array}{l} -\tilde{\mathbf{y}}(t_{k-d_2}, n_{z_y}) \tilde{\mathbf{y}}^T(t_k)\bar{\mathbf{a}} \\ -\tilde{\mathbf{u}}(t_{k-d_3}, n_{z_u}) \tilde{\mathbf{u}}^T(t_k)\mathbf{b} \end{array} \right\} \quad (2.117)$$

Ainsi, si les délais sont tels que $d_2 \geq n_a + n_c$ et $d_3 \geq n_b$, le vecteur d'instruments $\mathbf{z}(t_k)$ satisfait à la condition (2.108). De plus, puisque le vecteur d'instruments (2.114) est plus proche du vecteur de régression, il est naturel de s'attendre à ce que la matrice $\mathbf{R}_{z\varphi}$ soit mieux conditionnée que lorsque l'on utilise le vecteur d'instruments (2.110). Les résultats obtenus à l'aide de la méthode xiv2 sont donc plus précis que ceux issus de la méthode xiv1 ; l'exemple donné en section 2.2.6 illustre ce point.

2.2.4 Compensation du biais de l'estimateur de la variable instrumentale

2.2.4.1 Méthode des moindres carrés séparables (ecls)

On suppose que :

Hb1. \tilde{u} et \tilde{y} sont non-corrélés ;

Hb2. $H^{\tilde{u}}(q) = H^{\tilde{y}}(q) = 1$, i.e. les bruits additifs \tilde{u} et \tilde{y} sont blancs, de variances respectives $\lambda_{e\tilde{u}}$ et $\lambda_{e\tilde{y}}$ (voir la figure 2.1) ;

Hs2. $G_o \in \mathcal{G}^*$ et $(H_o^{\tilde{u}}, H_o^{\tilde{y}}) \in \mathcal{H}^*$.

La méthode ecls⁷ [Ekm05] peut être vue comme une extension du concept de la compensation de biais de la méthode des moindres carrés à la méthode de la variable instrumentale. Elle est fondée sur le système (surdéterminé) suivant (à comparer avec (2.19)) :

$$(\mathbf{R}_{z\varphi} - \mathbf{R}_{z\tilde{\varphi}}(\boldsymbol{\rho})) \boldsymbol{\theta} = \mathbf{r}_{zy} - \mathbf{r}_{z\tilde{y}}(\boldsymbol{\rho}) \quad (2.118)$$

Le choix du vecteur des instruments $\mathbf{z}(t_k)$ détermine les structures de $\mathbf{R}_{z\tilde{\varphi}}$, de $\mathbf{r}_{z\tilde{y}}$, ainsi que les propriétés statistiques inconnues à estimer, regroupées dans le vecteur $\boldsymbol{\rho}$. Si, par exemple, on choisit $\mathbf{z}(t_k) = \boldsymbol{\varphi}(t_k)$ on retrouve une méthode du type bels. Si $\mathbf{z}(t_k)$ est choisi non-corrélé avec l'erreur d'équation $v(t_k)$, on obtient une méthode du type xiv. Comme l'on

⁷Pour *Extended Compensated Least Squares*.

souhaite ici estimer le vecteur des paramètres $\boldsymbol{\theta}$ et le vecteur $\boldsymbol{\rho}$, certaines des composantes de $\mathbf{z}(t_k)$ doivent être corrélées avec l'erreur d'équation $v(t_k)$.

L'équation (2.118) est résolue de la manière suivante : les estimées des vecteurs $\boldsymbol{\theta}$ et $\boldsymbol{\rho}$ sont obtenues en minimisant une fonction de coût donnée par :

$$V(\boldsymbol{\theta}, \boldsymbol{\rho}) = \|\mathbf{r}_{zy} - \mathbf{r}_{\tilde{z}\tilde{y}}(\boldsymbol{\rho}) - (\mathbf{R}_{z\varphi} - \mathbf{R}_{\tilde{z}\tilde{\varphi}}(\boldsymbol{\rho})) \boldsymbol{\theta}\|^2 \quad (2.119a)$$

$$(\boldsymbol{\theta}_{\text{ecls}}^*, \boldsymbol{\rho}_{\text{ecls}}^*) = \arg \min_{\boldsymbol{\theta}, \boldsymbol{\rho}} V(\boldsymbol{\theta}, \boldsymbol{\rho}) \quad (2.119b)$$

Puisque le problème de minimisation (2.119) est séparable, les deux vecteurs $\boldsymbol{\theta}$ et $\boldsymbol{\rho}$ peuvent être estimés séparément. Pour $\boldsymbol{\rho}$ fixé, on minimise (2.119) par rapport à $\boldsymbol{\theta}$ donne une première estimation de $\boldsymbol{\theta}$:

$$\boldsymbol{\theta}_{\text{ecls}}^*(\boldsymbol{\rho}) = (\mathbf{R}_{z\varphi} - \mathbf{R}_{\tilde{z}\tilde{\varphi}}(\boldsymbol{\rho}))^\dagger (\mathbf{r}_{zy} - \mathbf{r}_{\tilde{z}\tilde{y}}(\boldsymbol{\rho})) \quad (2.120)$$

Une estimée de $\boldsymbol{\rho}$ est alors obtenue en substituant (2.120) dans (2.119a) :

$$\boldsymbol{\rho}_{\text{ecls}}^* = \arg \min_{\boldsymbol{\rho}} V(\boldsymbol{\theta}_{\text{ecls}}^*(\boldsymbol{\rho}), \boldsymbol{\rho}) \quad (2.121)$$

Enfin, l'estimée finale du vecteur des paramètres $\boldsymbol{\theta}$ est donnée par (2.120), où l'on a substitué $\boldsymbol{\rho}$ par sa valeur obtenue précédemment.

REMARQUE

Une étude de la convergence de cette méthode, et son extension au cas de bruits colorés en entrée et en sortie peuvent être trouvées dans [EHS06]. Il est par ailleurs à noter qu'en introduisant un filtrage des données, cette méthode peut fournir des estimées ayant une précision accrue, proche de la borne de Cramér-Rao, (voir [Mah07]). ■

2.2.4.2 Méthode de compensation du biais de la variable instrumentale (bciv)

Les hypothèses d'application sont :

Hb1. \tilde{u} et \tilde{y} sont non-corrélés ;

Hb3. $H^{\tilde{u}}(q) = 1$ et $H^{\tilde{y}}(q) = 1 + c_1^{\tilde{y}}q^{-1} + \dots + c_{n_c}^{\tilde{y}}q^{-n_c}$, c'est-à-dire \tilde{u} est blanc, de variance $\lambda_{e^{\tilde{u}}}$, tandis que \tilde{y} est obtenu en sortie d'un filtre MA ;

Hs2. $G_o \in \mathcal{G}^*$ et $(H_o^{\tilde{u}}, H_o^{\tilde{y}}) \in \mathcal{H}^*$.

La précision des estimées délivrées par les méthodes de compensation de biais dépend de la qualité de l'estimation des fonctions d'autocovariance des bruits. Or cette estimation est d'autant plus délicate que le nombre de points en lesquels ces fonctions doivent être estimées est grand. Lorsque les bruits sont supposés blancs, il n'y a que leur variance à estimer, et les méthodes de compensation du biais de l'estimateur des moindres carrés fournissent des résultats satisfaisants. En revanche, lorsque le bruit en sortie est coloré, le nombre de points en lesquels sa fonction d'autocorrélation doit être estimée est directement lié à l'ordre du modèle estimé (voir (2.76)). Cela peut poser problème si celui-ci est grand.

Par ailleurs, les méthodes fondées sur une variable instrumentale sont intéressantes car elles peuvent s'appliquer dans des conditions de bruit assez générales (blanc en entrée, coloré en sortie), et sont peu coûteuses d'un point de vue algorithmique. Il n'est toutefois pas toujours aisé de choisir le vecteur des instruments, car les deux conditions qu'il doit satisfaire (2.108)–(2.109) sont bien souvent contradictoires : pour qu'il soit non-corrélé avec les bruits, il doit faire intervenir des signaux d'entrée/sortie assez retardés, ce qui le rend « moins corrélé » avec le vecteur de régression.

Dans cette partie nous proposons un moyen permettant d'éviter ces désavantages [TGG08b]. Une méthode de compensation du biais de la variable instrumentale est utilisée pour obtenir une estimée du vecteur des paramètres $\boldsymbol{\theta}$: dans une première étape la variable instrumentale permet de s'affranchir du bruit coloré en sortie du système, puis le biais induit par le bruit blanc en entrée est estimé et retranché. En autorisant l'estimateur de la variable instrumentale à être biaisé, le vecteur des instruments peut être choisi « plus corrélé » avec le vecteur de régression, tandis que la compensation de biais ne nécessite que l'estimation du bruit blanc en entrée. Il convient de noter que l'idée de compenser le biais de l'estimateur de la variable instrumentale a été initialement proposée dans [YSW93], dans le cadre de l'identification de modèles EIV à temps continu. Toutefois, afin de s'affranchir du problème de l'estimation des dérivées des signaux d'entrée/sortie, leur étude utilise une transformation bilinéaire ; il en résulte une perte de la blancheur du bruit en entrée, compliquant de la sorte l'expression du biais, ainsi que son estimation. L'algorithme proposé ci-après a l'avantage de rester très simple, et la méthode d'estimation du biais proposée est originale.

On définit le vecteur d'instruments comme :

$$\mathbf{z}^T(t_k) = [\mathbf{y}^T(t_{k-d}, n_{zy}) \quad \mathbf{u}^T(t_k, n_{zu})] \quad (2.122)$$

En utilisant (2.7a), l'estimateur de la variable instrumentale associé au vecteur d'instruments précédent s'écrit :

$$\boldsymbol{\theta}_{iv}^* = \mathbf{R}_{z\varphi}^{-1} \mathbf{r}_{zy} = \boldsymbol{\theta} + \mathbf{R}_{z\varphi}^{-1} \mathbf{r}_{zv} \quad (2.123)$$

Supposons que le délai vérifie $d \geq n_a + n_c$. Alors :

$$\mathbf{r}_{zv} = \bar{E} \begin{Bmatrix} -\tilde{\mathbf{y}}(t_{k-d}, n_{zy}) \tilde{\mathbf{y}}^T(t_k) \bar{\mathbf{a}} \\ -\tilde{\mathbf{u}}(t_k, n_{zu}) \tilde{\mathbf{u}}^T(t_k) \mathbf{b} \end{Bmatrix} \quad (2.124)$$

$$= -\lambda_{e\bar{u}} \mathbf{M}_4 \boldsymbol{\theta} \quad (2.125)$$

où :

$$\mathbf{M}_4 = \begin{bmatrix} \mathbf{0}_{n_{zy} \times n_a} & \mathbf{0}_{n_{zy} \times n_b} \\ \mathbf{0}_{n_{zu} \times n_a} & \mathbf{I}_{n_{zu} \times n_b} \end{bmatrix} \quad (2.126)$$

Donc finalement :

$$\boldsymbol{\theta}_{iv}^* = \boldsymbol{\theta} - \lambda_{e\bar{u}} \mathbf{R}_{z\varphi}^{-1} \mathbf{M}_4 \boldsymbol{\theta} \quad (2.127)$$

Si une estimée de la variance du bruit blanc en entrée $\lambda_{e\tilde{u}}$ est disponible, le biais de l'estimateur (2.127) peut alors être retranché pour obtenir une estimée non-biaisée de $\boldsymbol{\theta}$.

Estimation de $\lambda_{e\tilde{u}}$

Pour obtenir une estimée de $\lambda_{e\tilde{u}}$ l'idée est d'utiliser la corrélation entre l'erreur d'équation $v(t_k, \boldsymbol{\theta})$ et le « résidu » r_1 , défini par :

$$r_1(t_k, d, \boldsymbol{\theta}) = y(t_{k-d}) - \mathbf{z}^T(t_k) \boldsymbol{\theta} \quad (2.128)$$

$$= \bar{\mathbf{a}}^T \bar{\mathbf{y}}(t_{k-d}, n_{z_y}) - \mathbf{b}^T \mathbf{u}(t_k, n_{z_y}) \quad (2.129)$$

Alors, si $d > n_a + n_c$:

$$\begin{aligned} J_1(d, \boldsymbol{\theta}) &= \bar{\mathbb{E}}\{r_1(t_k, d, \boldsymbol{\theta})v(t_k, \boldsymbol{\theta})\} \\ &= \bar{\mathbf{a}}^T \bar{\mathbb{E}}\left\{\tilde{\mathbf{y}}(t_{k-d}, n_z) \tilde{\mathbf{y}}^T(t_k)\right\} \bar{\mathbf{a}} + \mathbf{b}^T \bar{\mathbb{E}}\{\tilde{\mathbf{u}}(t_k, n_{z_u}) \tilde{\mathbf{u}}^T(t_k)\} \mathbf{b} \\ &= \lambda_{e\tilde{u}} \mathbf{b}^T \mathbf{b} \end{aligned} \quad (2.130)$$

Ainsi, si une estimée de $\boldsymbol{\theta}$ est disponible, il est possible d'estimer la variance du bruit \tilde{u} par :

$$\hat{\lambda}_{e\tilde{u}} = \frac{J_1(d, \hat{\boldsymbol{\theta}})}{\hat{\mathbf{b}}^T \hat{\mathbf{b}}} \quad \text{pour } d > n_a + n_c \quad (2.131)$$

Amélioration par filtrage

L'estimation de $\lambda_{e\tilde{u}}$ s'appuie sur celle de \mathbf{b} . Or, c'est justement sur cette partie du vecteur des paramètres que le biais est le plus important dans le cas EIV. Il est par conséquent possible que les estimées de $\lambda_{e\tilde{u}}$ ne soient pas très précises. Cela implique des problèmes de convergence ainsi qu'une faible précision des estimées du vecteur des paramètres. Pour remédier à cela, une version filtrée de (2.131) est proposée. Ce filtrage permet d'apporter des connaissances *a priori* au problème (*i.e.* les coefficients du filtre) [TGG08b]. Définissons un filtre MA $F(q^{-1})$ par :

$$F(q^{-1}) = 1 + \sum_{k=1}^{n_f} f_k q^{-k} \quad (2.132)$$

De la même manière qu'auparavant, on s'intéresse à la corrélation entre $v(t_k, \boldsymbol{\theta})$ et un « résidu » r_2 , défini par :

$$r_2(t_k, d, \boldsymbol{\theta}) = \bar{\mathbf{a}}^T \bar{\mathbf{y}}(t_{k-d}, n_z) - \mathbf{b}^T \mathbf{u}_f(t_k, n_{z_u}) \quad (2.133)$$

où $\mathbf{u}_f(t_k, n_{z_u}) = F(q^{-1}) \mathbf{u}(t_k, n_{z_u})$. Alors, si $d > n_a + n_c$:

$$\begin{aligned} J_2(d, \boldsymbol{\theta}) &= \bar{\mathbb{E}}\{r_2(t_k, d, \boldsymbol{\theta})v(t_k, \boldsymbol{\theta})\} \\ &= \bar{\mathbf{a}}^T \bar{\mathbb{E}}\left\{\tilde{\mathbf{y}}(t_{k-d}, n_z) \tilde{\mathbf{y}}^T(t_k)\right\} \bar{\mathbf{a}} + \mathbf{b}^T \bar{\mathbb{E}}\{\tilde{\mathbf{u}}_f(t_k, n_{z_u}) \tilde{\mathbf{u}}^T(t_k)\} \mathbf{b} \\ &= \lambda_{e\tilde{u}} \mathbf{b}^T \mathbf{M}_F \mathbf{b} \end{aligned} \quad (2.134)$$

où \mathbf{M}_F est la matrice de Toeplitz suivante :

$$\mathbf{M}_F = \begin{bmatrix} f_0 & 0 & \dots & 0 \\ f_1 & f_0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ f_{n_b-1} & f_{n_b-2} & \dots & f_0 \end{bmatrix} \quad (2.135)$$

Subséquentement, si une estimée de \mathbf{b} est disponible, il est possible d'estimer la variance $\lambda_{e\tilde{u}}$ par :

$$\hat{\lambda}_{e\tilde{u}} = \frac{J_2(d, \hat{\boldsymbol{\theta}})}{\hat{\mathbf{b}}^T \mathbf{M}_F \hat{\mathbf{b}}} \quad \text{pour } d > n_a + n_c \quad (2.136)$$

REMARQUES

- la méthode **bciv** nécessite la connaissance de $\boldsymbol{\theta}$ (dans l'expression du biais et dans l'estimation de la variance du bruit en entrée). Elle doit donc être itérative ;
- puisque (2.131) est vraie pour $d > n_a + n_c$, une meilleure estimée de la variance peut être obtenue en utilisant plusieurs valeurs de d au lieu d'une seule. Le nombre de valeurs utilisées est alors noté n_{bciv} . ■

L'algorithme **bciv**

1. Initialisation : $\hat{\boldsymbol{\theta}}_{\text{bciv}}^0 = \hat{\boldsymbol{\theta}}_{\text{iv}}$
2. Itération jusqu'à convergence :
 - 2.1. Estimation de $\lambda_{e\tilde{u}}$ à l'aide de (2.131) ou (2.136) ;
 - 2.2. Calcul de $\hat{\boldsymbol{\theta}}_{\text{bciv}}^{i+1}$:

$$\hat{\boldsymbol{\theta}}_{\text{bciv}}^{i+1} = \hat{\boldsymbol{\theta}}_{\text{iv}} + \hat{\lambda}_{e\tilde{u}}^i \hat{\mathbf{R}}_{z\varphi}^{-1} \mathbf{M}_4 \hat{\boldsymbol{\theta}}_{\text{bciv}}^i$$

Dans l'exemple de simulation du [paragraphe 2.2.6.2](#), l'algorithme utilisant (2.131) pour estimer la variance de \tilde{u} est appelé **bciv**⁸, tandis que l'algorithme utilisant (2.136) est noté **fbcv**⁹ (le nombre de valeurs de $J_1(d, \boldsymbol{\theta})$ utilisées pour l'estimation est noté n_{bciv}).

2.2.5 Récapitulatif des méthodes présentées

Avant d'illustrer les performances des méthodes présentées jusqu'ici à l'aide de simulations numériques, il nous a semblé utile de donner un tableau récapitulatif ces méthodes, ainsi que leurs hypothèses, que nous rappelons ici :

Hs1. $\mathcal{S} \in \mathcal{M}^*$

Hs2. $G_o \in \mathcal{G}^*$ et $(H_o^{\tilde{u}}, H_o^{\tilde{y}}) \in \mathcal{H}^*$

⁸Pour *Bias Compensated Instrumental Variable*.

⁹Pour *Filtered Bias Compensated Instrumental Variable*.

- He1.** Le signal d'entrée non-bruité u_o a une densité spectrale rationnelle
- Hb1.** Les bruits en entrée et en sortie \tilde{u} et \tilde{y} sont non-corrélés
- Hb2.** Les bruits en entrée et en sortie sont blancs
- Hb3.** Le bruit en entrée est blanc, tandis que le bruit en sortie est coloré (filtre MA)
- Hb4.** Le bruit en entrée est blanc, tandis que le bruit en sortie est coloré (filtre ARMA)
- Hb5.** Les bruits en entrée et en sortie sont colorés (filtres ARMA)

Toutes les méthodes présentées font l'hypothèse que les bruits en entrée et en sortie du système ne sont pas corrélés (**Hb1**). Notons par ailleurs que les hypothèses **Hs1–Hs2** et **Hb2–Hb5** sont classées de la plus restrictive à la moins restrictive.

Les méthodes sont organisées en trois catégories :

- les « méthodes diverses » regroupent : la méthode de Frisch (**frisch**) [DGS03, Söd08], les méthodes de l'erreur de prédiction (**pem**) et du maximum de vraisemblance (**ml**) [Söd81] ;
- les méthodes de compensation du biais de l'estimateur des moindres carrés, à savoir **bels** [Zhe98, Zhe02], **fbcls** [IKYW05] et **pbcls** [JIJW01] ;
- les méthodes de la variable instrumentale (**iv**) [SM02, TGG08b], des moindres carrés séparables (**ecls**) [Ekm05, EHS06] et de compensation du biais de la variable instrumentale (**bciv**) [TGG08b].

Certaines méthodes sont précisées comme n'étant pas applicables¹⁰ sous certaines des hypothèses **Hb2–Hb5**. Il convient de noter que ce n'est pas forcément parce qu'il est impossible de les étendre à ce cas de figure : il se peut que ce soit simplement parce que (à notre connaissance) cela n'a pas été réalisé jusqu'à présent. On peut ainsi juger des améliorations pouvant encore être apportées aux différentes méthodes.

Par ailleurs, bien que les méthodes **pem** et **ml** soient, en théorie, applicables lorsque les bruits ne sont pas blancs, ce cas n'est pas considéré dans la littérature. Or, le nombre de paramètres à estimer sous l'hypothèse **Hb5** (à savoir $\boldsymbol{\eta}_{u_o}$, $\boldsymbol{\eta}_{\tilde{u}}$ et $\boldsymbol{\eta}_{\tilde{y}}$, en plus du vecteur des paramètres d'intérêt $\boldsymbol{\theta}$, voir (2.1)) implique une optimisation numérique délicate, dont les inconvénients habituels sont accrus à mesure que le rapport signal à bruit diminue. Sans davantage d'étude sur leur comportement lorsque les bruits ne sont pas blancs, il est plus prudent de distinguer les cas de figure pour lesquels leur efficacité est avérée (●) de ceux pour lesquels elle reste à montrer (○).

2.2.6 Simulations numériques

Les performances des algorithmes sont étudiées à l'aide de simulations de Monte Carlo (de $n_{mc} = 100$ réalisations). Pour avoir une mesure d'ensemble de la qualité d'un estimateur

¹⁰Par « ne sont pas applicables » nous entendons « ne fournissent pas une estimation convergente du vecteur des paramètres $\boldsymbol{\theta}$ ».

TAB. 2.2 – Récapitulatif des méthodes présentées qui utilisent des statistiques de second ordre, suivant leur hypothèses d'application (• = disponible dans la littérature, ◦ = en théorie, jamais expérimenté).

	Méthodes diverses			Comp. biais LS			IV et comp. biais IV		
	frisch	pem	ml	bels ¹	fbcls	pbcls	iv	ecls	bciv ¹
Hs1.		•	•						
Hs2.	•			•	•	•	•	•	•
He1.		•	•						
Hb1.	•	•	•	•	•	•	•	•	•
Hb2.	•	•	•	•	•	•	•	•	•
Hb3.	•	◦	◦	•			•	•	•
Hb4.	•	◦	◦	•			•	•	•
Hb5.		◦	◦					•	

¹ Représentent les différentes versions de ces méthodes (*i.e.* abels, fbciv, ...)

l'erreur quadratique moyenne relative sur les paramètres est calculée :

$$\text{EQMR} = \sqrt{\frac{1}{n_{mc}} \sum_{j=1}^{n_{mc}} \frac{\|\hat{\theta}_j - \theta_o\|^2}{\|\theta_o\|^2}} \quad (2.137)$$

où $\hat{\theta}_j$ représente l'estimée obtenue à la $j^{\text{ième}}$ réalisation de la simulation de Monte Carlo. Le même critère d'arrêt est adopté pour chacun des algorithmes itératifs :

$$\frac{\|\hat{\theta}^{i+1} - \hat{\theta}^i\|}{\|\hat{\theta}^i\|} < 10^{-3} \quad \text{ou} \quad i > 50 \quad (2.138)$$

Ainsi, si l'algorithme n'a pas « convergé » avant 50 itérations, il est automatiquement stoppé. Les variances des bruits perturbant l'entrée et la sortie du système sont ajustées de manière à obtenir un certain rapport signal à bruit (RSB), défini par :

$$\text{RSBu} = 10 \log_{10} \left(\frac{P_{u_o}}{P_{\hat{u}}} \right) \quad \text{RSBy} = 10 \log_{10} \left(\frac{P_{y_o}}{P_{\hat{y}}} \right) \quad (2.139)$$

où P_x représente la puissance moyenne du signal x .

2.2.6.1 Analyse comparative des différentes méthodes bcls

Nous nous intéressons tout d'abord aux algorithmes de compensation du biais de l'estimateur des moindres carrés, présentés à la [section 2.2.2](#).

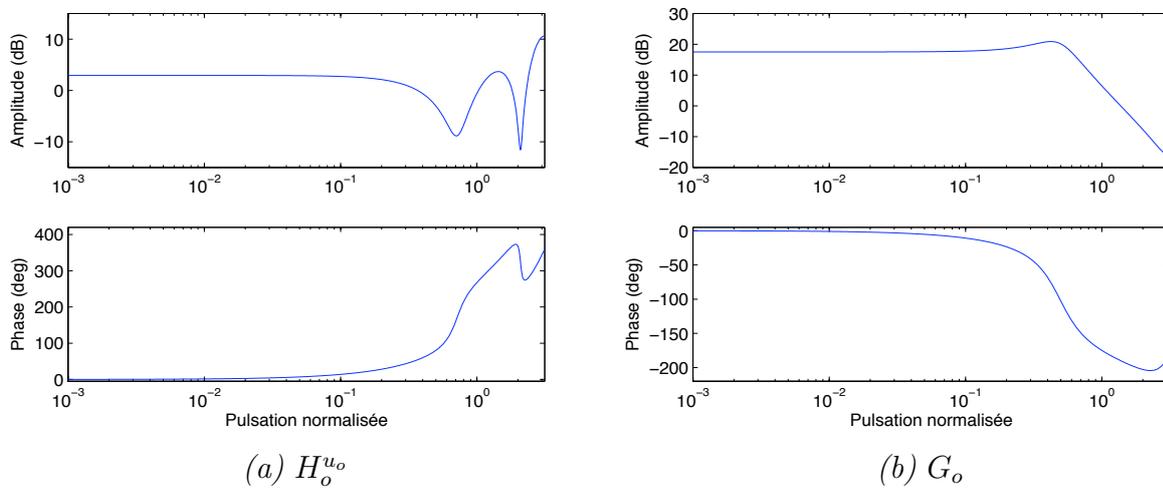
Bruit blanc en entrée et en sortie

Le système considéré est caractérisé par [Zhe99] :

$$\begin{cases} G_o(q) = \frac{1.0q^{-1} + 0.5q^{-2}}{1 - 1.5q^{-1} + 0.7q^{-2}} \\ H_o^{u_o}(q) = 1 - 0.3q^{-1} + 0.5q^{-2} - 0.7q^{-3} + 0.9q^{-4} \\ H_o^{\tilde{u}}(q) = H_o^{\tilde{y}}(q) = 1 \end{cases} \quad (2.140)$$

Les diagrammes de Bode de $H_o^{u_o}$ et de G_o sont représentés à la figure 2.2. Le vecteur des paramètres à estimer est ainsi $\theta_o^T = [-1.5 \ 0.7 \ 1 \ 0.5]$. Les bruits additifs en entrée/sortie

FIG. 2.2 – Diagrammes de Bode de $H_o^{u_o}$ et G_o



sont blancs (gaussiens), de variances fixées à $\lambda_{e_{\tilde{u}}} = 0.26$ et $\lambda_{e_{\tilde{y}}} = 1.9$, de sorte à avoir un RSB en entrée et en sortie environ égal à 10 dB. Le tableau 2.3 contient les résultats de la simulation pour des jeux de données de longueurs respectives $N = 1000$ et $N = 5000$ (« #it. » représente le nombre d'itérations nécessaires avant convergence). Pour mettre en exergue l'intérêt des algorithmes dédiés aux modèles erreurs à variables, l'identification a également été réalisée à l'aide des méthodes `iv4` et `sriv`¹¹, qui supposent le signal d'entrée non-bruité. Les diagrammes de Bode des modèles estimés à l'aide de $N = 1000$ points sont tracés à la figure 2.3.

Discussion

On constate tout d'abord que, pour ce niveau de bruit, les méthodes `sriv` et `iv4` fournissent logiquement des estimées biaisées. Ce biais concerne principalement les paramètres du numérateur de l'opérateur de transfert, tandis que les paramètres du dénominateur sont

¹¹`iv4` et `sriv` sont deux méthodes fondées sur une variable instrumentale optimale. Elles sont respectivement disponibles au sein des boîtes à outils Matlab SID et CONTSID. `iv4` est un estimateur à quatre étapes supposant un modèle de type ARARX [Lju99], tandis que `sriv` est un estimateur itératif supposant un modèle OE [You84].

TAB. 2.3 – Comparaison des méthodes bcls : moyenne des estimées obtenues \pm écart-type (bruits blancs, RSB \simeq 10dB)

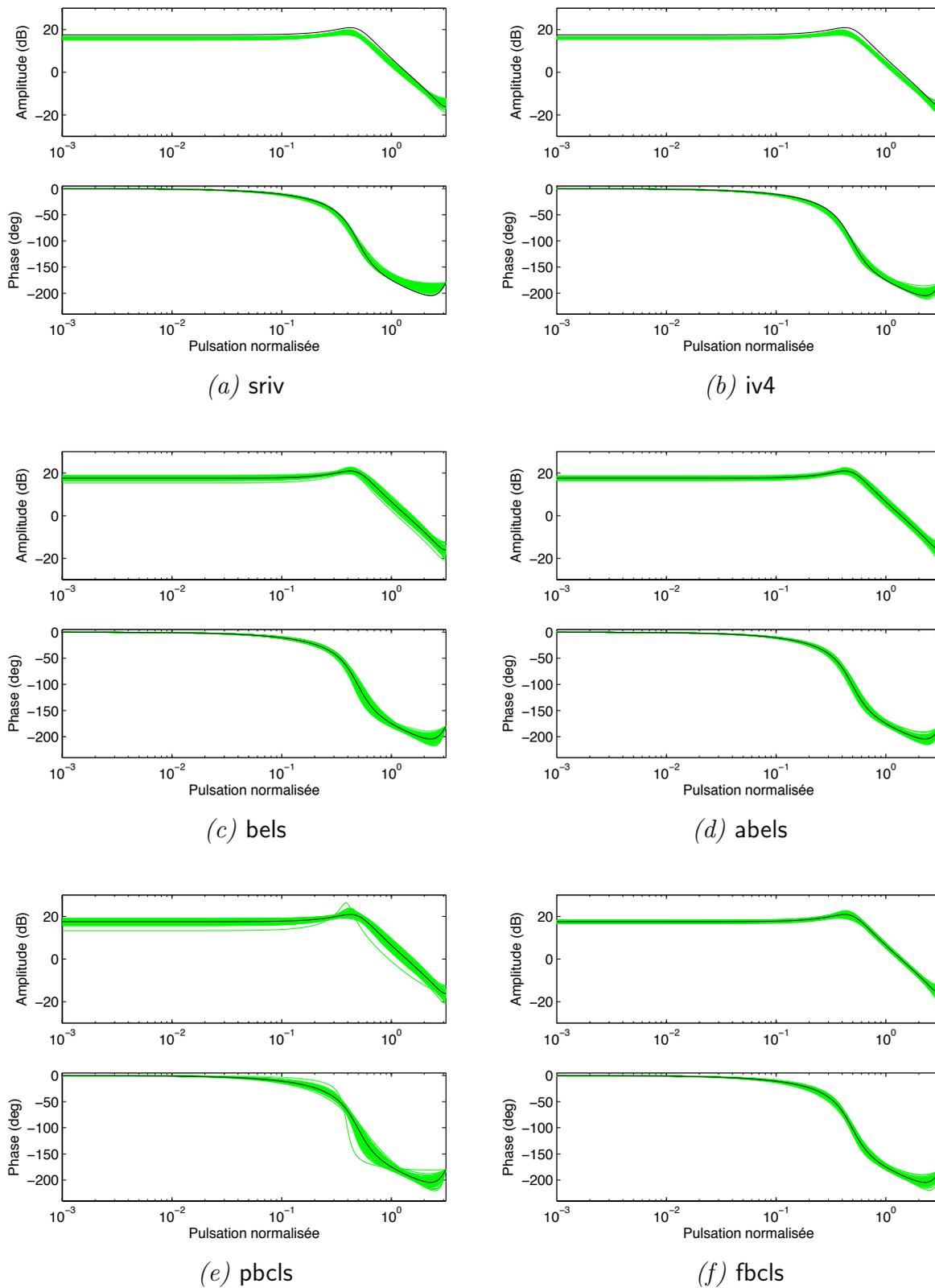
	N	$a_1 = -1.5$	$a_2 = 0.7$	$b_1 = 1.0$	$b_2 = 0.5$	$\lambda_{e_{\hat{u}}} = 0.26$	$\lambda_{e_{\hat{y}}} = 1.9$	EQMR	#it.
sriv	1000	-1.4979 ± 0.0327	0.6785 ± 0.0262	0.8838 ± 0.0482	0.2551 ± 0.0698	—	—	14.40%	8.7
	5000	-1.4959 ± 0.0154	0.6791 ± 0.0126	0.8683 ± 0.0202	0.2792 ± 0.0312	—	—	13.08%	7.3
iv4	1000	-1.4893 ± 0.0353	0.6709 ± 0.0303	0.8080 ± 0.0404	0.3515 ± 0.0557	—	—	12.93%	—
	5000	-1.4888 ± 0.0128	0.6722 ± 0.0110	0.8085 ± 0.0162	0.3587 ± 0.0258	—	—	12.13%	—
pbcls	1000	-1.4958 ± 0.0890	0.6982 ± 0.0703	1.0154 ± 0.1323	0.5148 ± 0.1566	0.2407 ± 0.2042	1.8750 ± 0.1198	11.72%	16.3
	5000	-1.5057 ± 0.0363	0.7048 ± 0.0294	0.9970 ± 0.0462	0.4913 ± 0.0687	0.2440 ± 0.0797	1.9006 ± 0.0548	4.77%	15.4
bels	1000	-1.4865 ± 0.0589	0.6912 ± 0.0442	1.0318 ± 0.1161	0.5305 ± 0.1189	0.2794 ± 0.1258	1.8672 ± 0.1109	9.35%	7.5
	5000	-1.4997 ± 0.0255	0.7001 ± 0.0197	1.0067 ± 0.0487	0.5018 ± 0.0537	0.2599 ± 0.0576	1.8957 ± 0.0509	3.96%	7.1
abels ($n_{\text{bels}}=3$)	1000	-1.4916 ± 0.0350	0.6949 ± 0.0263	1.0225 ± 0.0924	0.5190 ± 0.0795	0.2758 ± 0.0925	1.8715 ± 0.0985	6.63%	7.1
	5000	-1.5004 ± 0.0149	0.7006 ± 0.0114	1.0060 ± 0.0400	0.5007 ± 0.0352	0.2605 ± 0.0392	1.8966 ± 0.0441	2.83%	6.4
fbcls	1000	-1.5026 ± 0.0360	0.7032 ± 0.0328	1.0050 ± 0.0623	0.4987 ± 0.0703	0.2597 ± 0.0295	1.8796 ± 0.0876	5.28%	10.4
	5000	-1.5015 ± 0.0136	0.7015 ± 0.0121	1.0036 ± 0.0287	0.4980 ± 0.0335	0.2581 ± 0.0131	1.8975 ± 0.0403	2.39%	10.1

plutôt convenablement estimés. Les différents algorithmes de compensation de biais fournissent en revanche des estimées non-biaisées. On peut distinguer deux groupes d'algorithmes : **pbcls** et **bels** d'une part, **abels** et **fbcls** d'autre part. Bien que satisfaisants, les premiers offrent de moins bons résultats, car l'écart-type sur les estimées des variances des bruits blancs est plus important, ce qui implique un écart-type plus important sur les paramètres estimés.

La clé de voûte de ces algorithmes est en effet l'estimation des variances $\lambda_{e_{\hat{u}}}$ et $\lambda_{e_{\hat{y}}}$: meilleures sont ces estimées, meilleures seront les estimées des paramètres, comme en attestent les résultats contenus dans le [tableau 2.3](#). Nous avons vu qu'il existe de nombreuses manières d'estimer ces deux variances. Une perspective intéressante serait donc d'étudier s'il existe une « meilleure » manière d'utiliser les données disponibles pour estimer les variances.

Bruit blanc en entrée, coloré en sortie

Nous nous intéressons maintenant au cas d'un bruit coloré en sortie. Les méthodes étudiées sont donc **bels2** et **fbels2**, que l'on utilise pour estimer les paramètres d'un système de premier

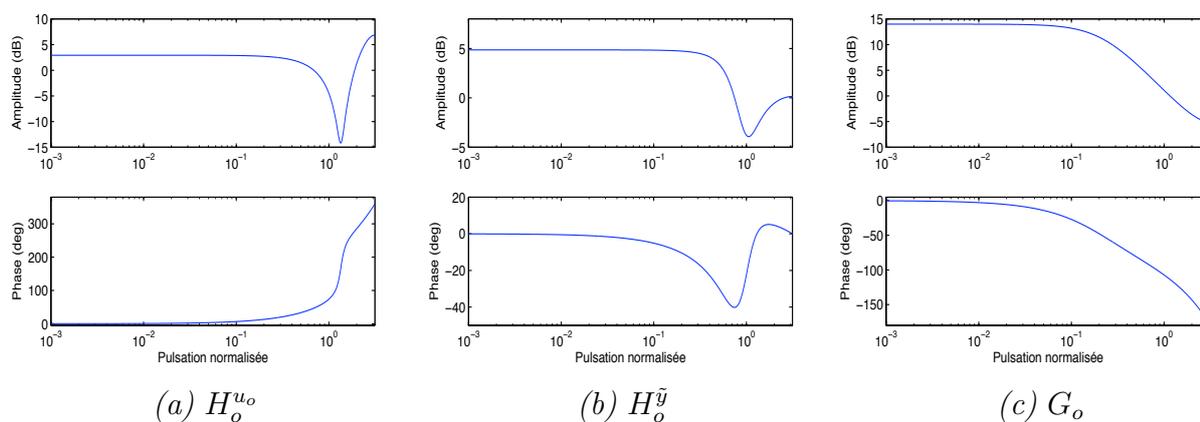
FIG. 2.3 – Diagrammes de Bode du vrai système (‘-’) et des modèles estimés (bruits blancs, $N = 1000$, $RSB \simeq 10$ dB)

ordre, donné par¹² (on constatera dans l'exemple suivant que lorsque l'ordre du système à identifier augmente, les performances de la méthode `bels2` se dégradent significativement) :

$$\begin{cases} G_o(q) = \frac{1.0q^{-1}}{1 - 0.8q^{-1}} \\ H_o^{u_o}(q) = 1 - 0.4q^{-1} + 0.8q^{-2} \\ H_o^{\tilde{y}}(q) = \frac{1 - 0.87q^{-1} + 0.57q^{-2}}{1 - q^{-1} + 0.4q^{-2}} \\ H_o^{\tilde{u}}(q) = 1 \end{cases} \quad (2.141)$$

Les diagrammes de Bode de $H_o^{u_o}$, $H_o^{\tilde{y}}$ et G_o sont donnés à la figure 2.4. Les variances des bruits sont fixées à $\lambda_{e_{\tilde{u}}} = 0.56$ et $\lambda_{e_{\tilde{y}}} = 1.15$ pour obtenir un RSB en entrée et en sortie environ égal à 5 dB. La longueur du jeu de données est fixée à $N = 1000$. À titre de comparaison, les résultats fournis par la méthode `abels`, qui suppose blancs les bruits en entrée/sortie, apparaissent également dans le tableau 2.4. Les diagrammes de Bode des modèles estimés se trouvent à la figure 2.5.

FIG. 2.4 – Diagrammes de Bode de $H_o^{u_o}$, $H_o^{\tilde{y}}$ et G_o



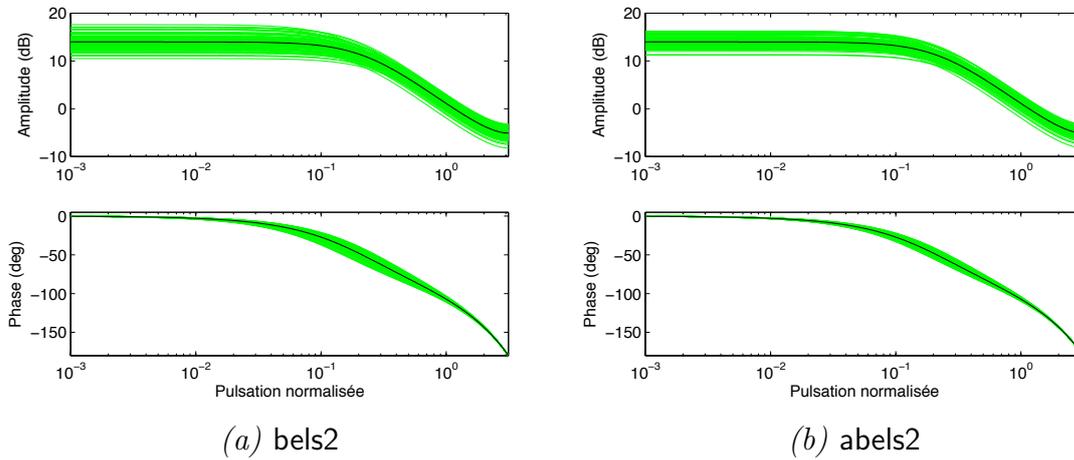
Discussion

À nouveau, on constate qu'une meilleure estimation des propriétés statistiques des bruits conduit à une précision accrue des paramètres estimés. La méthode `abels2` fournit ainsi les meilleurs résultats, offrant notamment une plus petite variance sur les estimées des fonctions d'autocovariance. Notons par ailleurs que la méthode `abels` fournit de médiocres résultats, illustrant l'implication d'une erreur de modélisation des bruits affectant l'entrée et la sortie du système. Cette méthode suppose en effet que les bruits en entrée/sortie sont blancs. On est donc dans une situation où $G_o \in \mathcal{G}^*$, mais $(H_o^{\tilde{u}}, H_o^{\tilde{y}}) \notin \mathcal{H}^*$, puisque le bruit en sortie du vrai système est coloré.

¹²Les conditions de simulation et le système s'inspirent de [Zhe02].

TAB. 2.4 – Comparaison des méthodes `bels2` et `abels2` (bruit blanc en entrée et coloré en sortie, $N = 1000$, $\text{RSB} \simeq 5\text{dB}$)

	$a_1 = -0.8$	$b_1 = 1.0$	$\lambda_{e_{\tilde{y}}} = 1.3629$	$r_{\tilde{y}(1)} = 0.3223$	$\lambda_{e_{\tilde{u}}} = 0.56$	EQMR	#it.
<code>abels</code> ($n_{\text{bels}}=3$)	-0.7830 ± 0.0198	1.2122 ± 0.0817	0.8218 ± 0.1011	–	0.8903 ± 0.1058	17.8%	4
<code>bels2</code>	-0.8011 ± 0.0288	0.9936 ± 0.1039	1.2893 ± 0.7283	0.2464 ± 0.6712	0.5440 ± 0.1709	8.39%	2
<code>abels2</code> ($n_{\text{bels}}=4$)	-0.7996 ± 0.0216	0.9905 ± 0.0994	1.3887 ± 0.4548	0.3434 ± 0.3859	0.5395 ± 0.1606	7.94%	2

FIG. 2.5 – Diagrammes de Bode du vrai système (‘–’) et des modèles estimés (bruit blanc en entrée et coloré en sortie, $N = 1000$, $\text{RSB} \simeq 5\text{ dB}$)

2.2.6.2 Analyse comparative des méthodes fondées sur une variable instrumentale

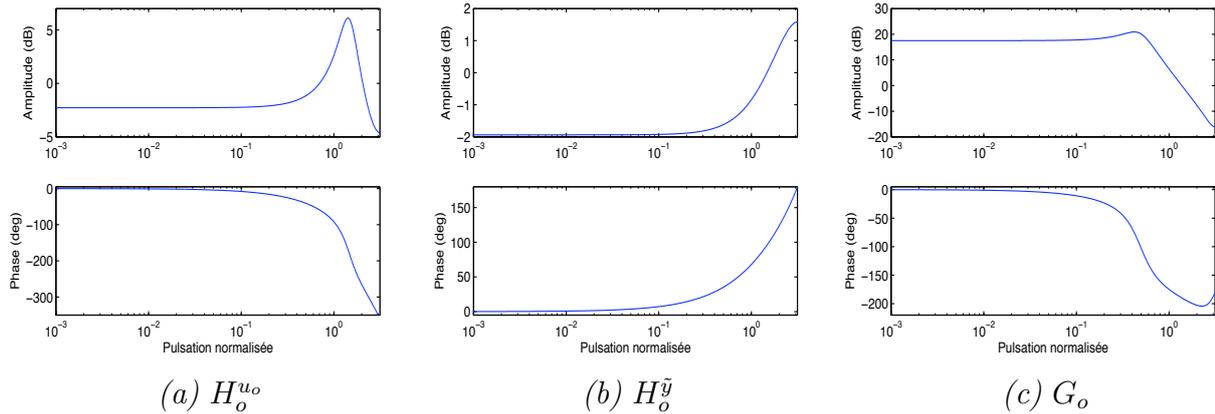
Le système considéré est [SM02] :

$$\begin{cases} G_o(q) = \frac{1.0q^{-1} + 0.5q^{-2}}{1 - 1.5q^{-1} + 0.7q^{-2}} \\ H_o^{u_o}(q) = \frac{1}{1 - 0.2q^{-1} + 0.5q^{-2}} \\ H_o^{\tilde{y}}(q) = 1 - 0.2q^{-1} \\ H_o^{\tilde{u}}(q) = 1 \end{cases} \quad (2.142)$$

Les diagrammes de Bode de $H_o^{u_o}$, $H_o^{\tilde{y}}$ et G_o se trouvent à la figure 2.6. Les variances des bruits blancs sont fixées à $\lambda_{e_{\tilde{u}}} = 0.14$ et $\lambda_{e_{\tilde{y}}} = 1.45$ pour obtenir un RSB environ égal à 10 dB en entrée et en sortie. La longueur du jeu de données est $N = 2000$, et le filtre de la méthode `fbci` est choisi comme $F(q^{-1}) = 1 + q^{-1} + q^{-2}$ (voir (2.132)). La valeur minimale

des délais est choisie pour chaque méthode utilisant une variable instrumentale. Les résultats de la simulation se trouvent dans le [tableau 2.5](#). Puisque certaines méthodes ne convergent pas toujours avant 50 itérations, ce tableau contient une ligne « C^{ce} », qui représente le pourcentage de réalisations pour lesquelles les algorithmes ont convergé. Les diagrammes de Bode sont représentés à la [figure 2.7](#).

FIG. 2.6 – Diagrammes de Bode de $H_o^{u_o}$, $H_o^{\tilde{y}}$ et G_o



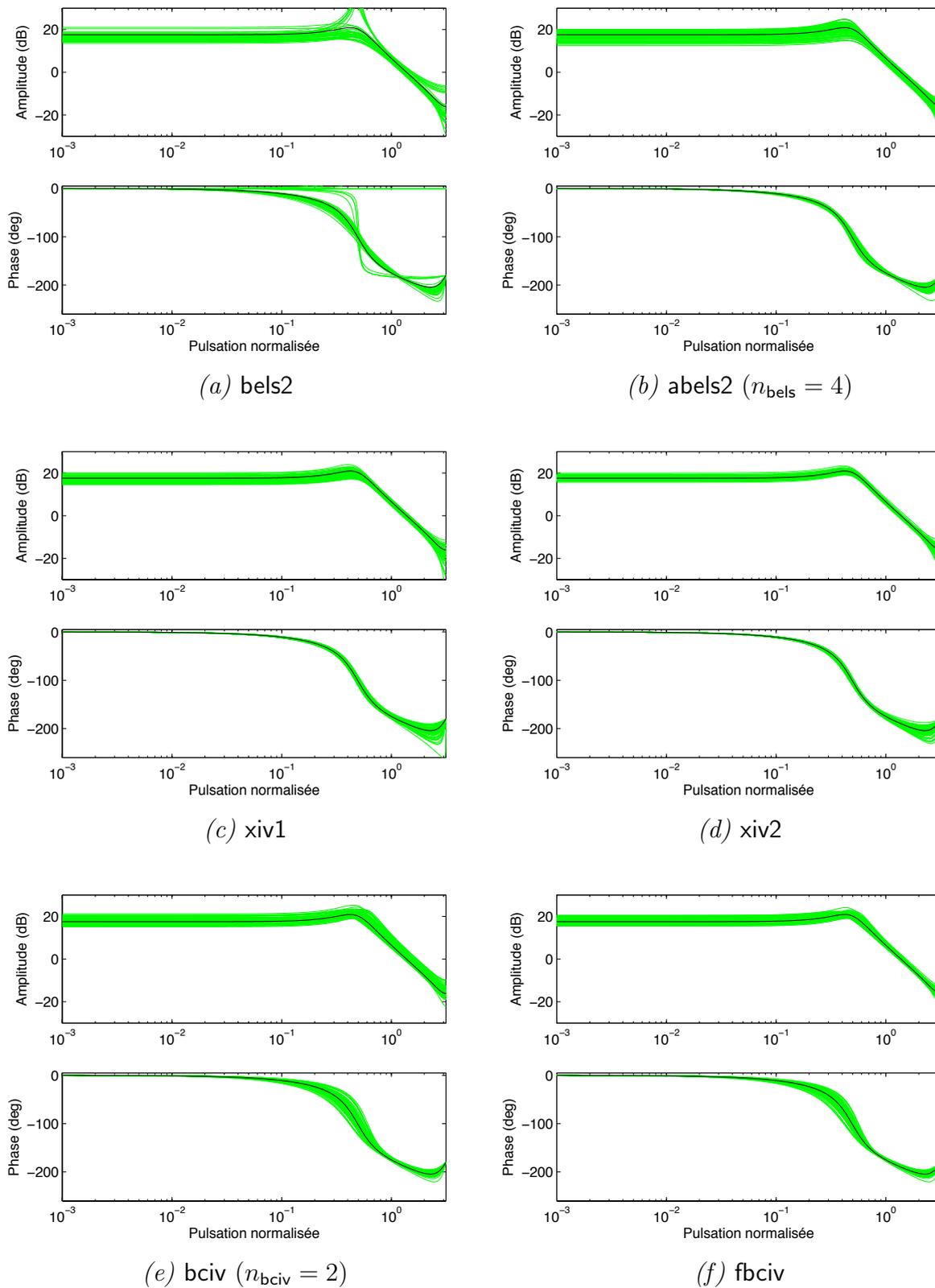
TAB. 2.5 – Comparaison des méthodes xiv, bciv et bels2 (bruit blanc en entrée et coloré en sortie, $N = 2000$, $RSB \simeq 10\text{dB}$)

		bels2	abels2		xiv1	xiv2	bciv		fbciv
			$n_{\text{bels}} = 3$	$n_{\text{bels}} = 4$			$n_{\text{bciv}} = 1$	$n_{\text{bciv}} = 2$	
a_1	-1.5	-0.696 ± 0.556	-1.488 ± 0.002	-1.488 ± 0.001	-1.476 ± 0.001	-1.499 ± 0.001	-1.462 ± 0.005	-1.490 ± 0.001	-1.498 ± 0.001
a_2	0.7	0.325 ± 0.130	0.691 ± 0.001	0.691 ± 0.001	0.681 ± 0.001	0.699 ± 0.001	0.689 ± 0.001	0.695 ± 0.001	0.697 ± 0.001
b_1	1	0.450 ± 0.279	0.973 ± 0.022	0.972 ± 0.014	0.942 ± 0.012	0.996 ± 0.008	1.305 ± 0.251	1.083 ± 0.044	1.022 ± 0.009
b_2	0.5	0.207 ± 0.061	0.491 ± 0.003	0.494 ± 0.003	0.495 ± 0.006	0.504 ± 0.007	0.629 ± 0.053	0.533 ± 0.009	0.506 ± 0.003
$\lambda_{e\tilde{u}}$	0.14	0.626 ± 12.56	0.069 ± 0.049	0.080 ± 0.028	—	—	0.361 ± 0.148	0.209 ± 0.054	0.155 ± 0.013
C^{ce}	—	47%	100%	100%	—	—	52%	88%	100%
EQMR	—	66.4%	7.7%	6.3%	6.9%	5.5%	29.0%	11.2%	5.1%

Discussion

Contrairement à ce que l'on a constaté lors de l'exemple précédent pour le cas d'un système

FIG. 2.7 – Diagrammes de Bode du vrai système (‘-’) et des modèles estimés (bruit blanc en entrée et coloré en sortie, $N = 2000$, $RSB \approx 10$ dB)



de premier ordre, la méthode `bels2` ne fournit pas ici de résultats satisfaisants. Notons en particulier la mauvaise estimation de la variance $\lambda_{e\hat{a}}$. En revanche, ses performances sont grandement améliorées par l'utilisation du système sur-déterminé suggéré dans la présentation de la méthode `abels2` (voir le [paragraphe 2.2.2.2](#)). Bien que l'estimation de la variance du bruit en entrée ne soit toujours pas satisfaisante, elle a néanmoins été améliorée, ce qui conduit à une meilleure estimation du vecteur des paramètres.

Les deux méthodes fondées sur une variable instrumentale fournissent de bons résultats, avec un écart-type raisonnable. Notons que, comme l'on s'y attendait, l'utilisation de sorties retardées dans la méthode `xiv2` permet d'obtenir de meilleurs résultats que la méthode `xiv1`. Enfin, l'algorithme de compensation du biais de la variable instrumentale `bciv` donne en premier lieu des résultats médiocres, améliorés ensuite lorsque l'on estime la variance du bruit en entrée par résolution d'un système sur-déterminé. En revanche, les résultats deviennent convaincants lorsqu'est utilisée l'étape de filtrage de l'algorithme `fbiv`. En particulier, l'estimée de $\lambda_{e\hat{a}}$ est très proche de la vraie valeur de la variance.

2.3 Estimateurs fondés sur les statistiques d'ordre supérieur

La majeure partie des techniques d'identification ont pour point commun d'utiliser les propriétés statistiques de second ordre. La (relative) simplicité d'analyse et de calcul, ou encore le théorème de la limite centrale, qui laisse penser que tout signal à tendance à être gaussien (et donc caractérisé par ses statistiques de premier et second ordre), sont quelques raisons de cet engouement pour les statistiques de second ordre. De plus, l'entrée du système est souvent considérée parfaitement connue, et l'identifiabilité est dans ce cas généralement assurée si l'on utilise les statistiques de second ordre. Or, comme nous l'avons brièvement vu dans le [paragraphe 1.3.2.2](#) du [chapitre 1](#), lorsque l'entrée est bruitée, des problèmes d'identifiabilité surviennent, et ce même dans le cas le plus simple, à savoir si l'entrée et la sortie du système sont contaminées par des bruits blancs.

Il est alors possible de suivre deux voies : soit caractériser plus précisément la classe de modèles pour laquelle les statistiques de second ordre sont en adéquation avec les données [[AD87](#), [SN87](#), [AG06](#)], soit utiliser les statistiques d'ordre supérieur¹³ (SOS). Cette dernière voie est explorée dans cette partie.

Par SOS, il faut comprendre moments et cumulants d'ordre supérieur à deux (ou multispectre dans le domaine fréquentiel [[LAC97](#)]). Dans les années 1990, ce domaine de recherche a été très populaire dans la littérature du traitement du signal (voir par exemple [[Men91](#)] et les références qui y figurent). Parmi les nombreuses raisons de cette popularité, citons le fait que les cumulants de processus gaussiens sont nuls, et que lorsque l'on observe un signal non-gaussien perturbé par un bruit indépendant gaussien, les effets de ce bruit sont potentiellement supprimés. Il est à noter que lorsque le signal d'entrée est non-gaussien, les

¹³Par « supérieur » nous entendons « supérieur à deux ».

méthodes du type maximum de vraisemblance requièrent la connaissance de la distribution de cette entrée et du type de coloration des bruits ; ce pour aboutir à une optimisation d'une fonction non-linéaire en les paramètres. Un avantage des méthodes fondées sur les SOS est que, sans connaître explicitement la distribution du signal d'entrée, elles suppriment le bruit additif gaussien (qui peut lui aussi avoir un spectre inconnu) [DG94].

Cette partie ne prétend pas être exhaustive, et de nombreux domaines d'application ne seront pas abordés, comme l'estimation de retard [NP88, Tug91, CC94], ou encore l'estimation d'ordre et la validation de modèle [Tug94, Tug95, TL99]. Nous donnons toutefois dans la suite quelques détails sur certaines méthodes paraissant être plus appropriées au problème considéré. Notons par ailleurs que les méthodes fondées sur l'analyse des multispectres et multispectres intégrés (voir [TY95] par exemple) ne seront pas étudiées.

L'annexe A contient les principales définitions et propriétés des statistiques d'ordre supérieur qui sont utiles aux méthodes étudiées dans la suite.

2.3.1 Discussion des hypothèses d'application

Avant de présenter des approches développées pour l'identification de modèles EIV, il est nécessaire de discuter les hypothèses d'application de ces méthodes. Les cumulants d'ordre trois sont nuls pour une variable aléatoire ayant une distribution symétrique, tandis que les cumulants d'ordre quatre sont nuls pour une variable aléatoire ayant une distribution gaussienne ; voir l'annexe A. Si l'on désire appliquer les SOS à l'identification des systèmes, il faut donc formuler des hypothèses sur la distribution de l'entrée non-bruitée (afin que ses cumulants soient non nuls), ainsi que sur les distributions des bruits en entrée et en sortie (pour que leurs contributions disparaissent). Les hypothèses sur les bruits sont :

- pour les cumulants d'ordre trois : ils doivent être distribués de manière symétrique ;
- pour les cumulants d'ordre quatre : ils doivent être gaussiens.

Beaucoup de méthodes d'estimation font l'hypothèse que les bruits sont gaussiens (invoquant le théorème de la limite centrale comme justification), et – à notre connaissance – aucune ne suppose que les bruits ont une distribution asymétrique. Les hypothèses sur les bruits sont donc plutôt douces et communes.

Les hypothèses sur le signal d'entrée non-bruité, en revanche, sont spécifiques aux méthodes fondées sur les SOS :

- pour les cumulants d'ordre trois : il doit avoir une distribution asymétrique ;
- pour les cumulants d'ordre quatre : il doit être non-gaussien.

Supposer le signal d'entrée non-bruité distribué de manière asymétrique ou non gaussienne est certainement restrictif, et les méthodes fondées sur les cumulants ne peuvent être appliquées à un système réel sans auparavant vérifier si cette hypothèse est vérifiée. Il est néanmoins possible de procéder à divers tests pour jauger de l'applicabilité des méthodes (condition d'identifiabilité, valeur des cumulants, etc).

Les hypothèses de travail sont rappelées au début de chaque méthode. Il faut de surcroît noter que les méthodes utilisant des SOS, bien qu'imposant des restrictions sur le signal d'entrée non-bruité, permettent bien souvent de considérer la situation de bruit la plus générale, à savoir des bruits sur l'entrée et la sortie du système colorés et mutuellement corrélés. Enfin, comme nous l'avons succinctement vu lors de la simulation numérique de la [section 2.2.6](#), une erreur de modélisation des bruits peut avoir des conséquences importantes sur les résultats fournis par les méthodes fondées sur les statistiques d'ordre deux. Les méthodes fondées sur les SOS peuvent quant à elles s'appliquer *sans changement* dans le cas de bruits blancs ou colorés, donc sans requérir la modélisation des bruits (elles s'appliquent par conséquent sous l'hypothèse $G_o \in \mathcal{G}^*$).

2.3.2 Méthodes fondées sur la minimisation de critères faisant intervenir les cumulants d'ordre supérieur

Cumulants d'ordre quatre

Dans [Tug92] est proposée une méthode fondée sur la minimisation d'une fonction de coût faisant intervenir les cumulants d'ordre quatre de l'erreur d'équation habituelle, qui est donnée par :

$$v(t_k, \boldsymbol{\theta}) = y(t_k) - \boldsymbol{\varphi}^T(t_k)\boldsymbol{\theta} \quad (2.143)$$

Elle s'applique sous les hypothèses suivantes :

He1. le signal d'entrée non-bruité u_o a une densité spectrale rationnelle ;

He3. e^{u_o} est un processus stochastique i.i.d. non gaussien ;

Hb6. \tilde{u} et \tilde{y} sont gaussiens ;

Hs3. $G_o \in \mathcal{G}^*$.

Il est alors proposé de minimiser le critère :

$$J_4(\boldsymbol{\theta}, N) = \frac{1}{N} \sum_{k=1}^N v^4(t_k, \boldsymbol{\theta}) - 3 \left(\frac{1}{N} \sum_{k=1}^N v^2(t_k, \boldsymbol{\theta}) \right)^2 \quad (2.144)$$

On reconnaît l'expression de l'estimateur du cumulant d'ordre quatre de l'erreur d'équation. Une estimée du vecteur des paramètres est alors obtenue en minimisant :

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta}} \sqrt{|J_4(\boldsymbol{\theta}, N)|} \quad (2.145)$$

Puisque cette fonction de coût est non linéaire en $\boldsymbol{\theta}$, on a alors recours à un algorithme d'optimisation itératif de type Newton-Raphson.

REMARQUE

Deux observations s'imposent :

- l'estimation du cumulants d'ordre quatre de l'erreur d'équation apparaissant dans (2.144) est réalisée en utilisant un estimateur approximativement non biaisé dans [Tug92]. Il eût été préférable d'utiliser les k -statistiques (c.f. la section A.5.2);
- lorsqu'elles existent, les dérivées de $|J_4(\boldsymbol{\theta}, N)|$ par rapport à $\boldsymbol{\theta}$ sont nulles au voisinage de la solution désirée, ce qui implique que le critère (2.144) est relativement insensible à de petites perturbations sur $\boldsymbol{\theta}$ au voisinage de la solution. Cela a non seulement une conséquence sur le conditionnement du hessien de $|J_4(\boldsymbol{\theta}, N)|$, mais cela implique aussi une grande variance sur l'estimée de $\boldsymbol{\theta}$ [Tug92]. ■

Cumulants d'ordre trois

Dans [DG94], l'identification est réalisée à l'aide d'une fonction de coût faisant intervenir les cumulants d'ordre trois des données d'entrée/sortie.

Cette méthode s'applique sous les hypothèses suivantes :

He1. le signal d'entrée non-bruité u_o a une densité spectrale rationnelle;

He2. e^{u_o} est un processus stochastique i.i.d. ayant une distribution asymétrique;

Hb6. \tilde{u} et \tilde{y} ont une distribution symétrique;

Hs3. $G_o \in \mathcal{G}^*$.

Introduisons la fonction de coût suivante :

$$\bar{V}_\vartheta = \sum_{l=-\infty}^{\infty} \bar{\mathbb{E}}\{u(t_{k+l})\varepsilon_{u,y,\vartheta}^2(t_k)\} \quad (2.146)$$

où l'erreur de prédiction généralisée $\varepsilon_{u,y,\vartheta}(t_k)$ est construite comme une combinaison linéaire d'entrées et de sorties retardées :

$$\varepsilon_{u,y,\vartheta}(t_k) = W_u(q, \boldsymbol{\vartheta})u(t_k) + W_y(q, \boldsymbol{\vartheta})y(t_k) \quad (2.147)$$

Les filtres $W_u(q, \boldsymbol{\vartheta})$ et $W_y(q, \boldsymbol{\vartheta})$ sont supposés uniformément stables en $\boldsymbol{\vartheta}$ (voir [Lju99, chap. 7] pour des détails sur différentes manières de construire ces filtres, qui déterminent ce que contient le vecteur $\boldsymbol{\vartheta}$). Il est possible de montrer que la fonction de coût (2.146) est égale, à un facteur d'échelle près, au critère des moindres carrés « idéal », c'est-à-dire défini avec les signaux non-bruités [DG94] :

$$\bar{V}_\vartheta = C_{u_o u_o u_o}(0, 0) \frac{G(0)}{C_{u_o u_o}(0)} V_\vartheta \quad (2.148)$$

$$\text{avec } V_\vartheta = \bar{\mathbb{E}}\{\varepsilon_{u_o, y_o, \vartheta}^2(t_k)\} \quad (2.149)$$

Le vecteur des paramètres $\boldsymbol{\vartheta}$ est alors obtenu par optimisation (au sens des moindres carrés) de \bar{V}_ϑ :

$$\boldsymbol{\vartheta}^* = \arg \min_{\boldsymbol{\vartheta}} V_\vartheta = \arg \text{extremum}_{\boldsymbol{\vartheta}} \bar{V}_\vartheta \quad (2.150)$$

REMARQUE

Notons que le signe de $C_{u_o u_o u_o}(0, 0)G(0)/C_{u_o u_o}(0)$ est a priori inconnu. En outre, des problèmes surviennent si V_ϑ a plus d'un minimiseur global, car alors \mathfrak{D}^* représente un ensemble de vecteurs admissibles plutôt qu'un seul. Enfin, il se trouve que l'approximation de (2.146) réalisée avec les N données disponibles nécessite un fenêtrage pour être un estimateur convergent de \bar{V}_ϑ [DG94]. ■

2.3.3 Méthodes fondées sur une variable instrumentale (tociv, fociv)

Il est possible de modifier l'approche classique de la variable instrumentale afin d'y faire intervenir les cumulants d'ordre trois [IT91] (algorithme tociv¹⁴) ou les cumulants d'ordre quatre [IS94] (algorithme fociv¹⁵).

Cumulants d'ordre trois [IT91]

Cette méthode s'applique sous les hypothèses suivantes :

He2. u_o a une distribution asymétrique ;

Hb1. \tilde{u} et \tilde{y} sont non-corrélés ;

Hb6. \tilde{u} et \tilde{y} ont une distribution symétrique ;

Hs3. $G_o \in \mathcal{G}^*$.

Le point de départ de la méthode est l'équation du modèle mise sous forme de régression linéaire :

$$y(t_k) = \boldsymbol{\varphi}^T(t_k)\boldsymbol{\theta} + v(t_k, \boldsymbol{\theta}) \quad (2.151)$$

$$v(t_k, \boldsymbol{\theta}) = \tilde{y}(t_k) - \tilde{\boldsymbol{\varphi}}^T(t_k)\boldsymbol{\theta} \quad (2.152)$$

Définissons le vecteur d'instruments comme :

$$\mathbf{z}^T(t_k) = [-\mathbf{y}^{2,T}(t_k) \mathbf{u}^{2,T}(t_k)] \quad (2.153)$$

$$= [-y^2(t_{k-1}) \dots -y^2(t_{k-n_a}) u^2(t_{k-1}) \dots u^2(t_{k-n_b})] \quad (2.154)$$

où l'on note \mathbf{v}^2 le vecteur obtenu en mettant au carré toutes les composantes du vecteur \mathbf{v} . Si l'on multiplie (2.151) par $\mathbf{z}(t_k)$, et que l'on considère l'espérance du résultat, il vient :

$$\mathbf{r}_{zy} = \mathbf{R}_{z\varphi}\boldsymbol{\theta} + \mathbf{r}_{zv} \quad (2.155)$$

Les vecteurs et matrices apparaissant dans (2.155) contiennent des cumulants d'ordre trois, et il est possible de montrer que le vecteur des instruments $\mathbf{z}(t_k)$ et l'erreur d'équation $v(t_k, \boldsymbol{\theta})$ sont non-corrélés, sous les hypothèses **H4**, **H5**, **Hb1** et **Hb6** [IT91]. De plus, si le

¹⁴tociv signifie *Third-Order Cumulants-based Instrumental Variable*.

¹⁵fociv signifie *Fourth-Order Cumulants-based Instrumental Variable*.

signal d'entrée u_o est blanc à l'ordre trois¹⁶, la matrice $\mathbf{R}_{z\varphi}$ est non-singulière de manière générique. On obtient alors l'estimateur de la variable instrumentale fondé sur les cumulants d'ordre trois :

$$\boldsymbol{\theta}_{\text{tociv}}^* = \mathbf{R}_{z\varphi}^{-1} \mathbf{r}_{zy} \quad (2.156)$$

Cumulants d'ordre quatre [IS94]

Cette méthode s'applique dans le cas de bruits gaussiens indépendants, tandis que le signal d'entrée non-bruité est supposé non-gaussien :

He3. u_o est non-gaussien ;

Hb1. \tilde{u} et \tilde{y} sont non-corrélés ;

Hb7. \tilde{u} et \tilde{y} sont gaussiens ;

Hs3. $G_o \in \mathcal{G}^*$.

Définissons le vecteur d'instruments suivant :

$$\mathbf{z}(t_k) = \begin{bmatrix} -\mathbf{y}^3(t_k) + 3\bar{\mathbf{E}}\{\mathbf{y}^2(t_k)\} \mathbf{y}(t_k) \\ \mathbf{u}^3(t_k) - 3\bar{\mathbf{E}}\{\mathbf{u}^2(t_k)\} \mathbf{u}(t_k) \end{bmatrix} \quad (2.157)$$

où l'on note \mathbf{v}^3 le vecteur obtenu en mettant au cube toutes les composantes du vecteur \mathbf{v} . On constate une ressemblance avec l'estimateur (approximativement non biaisé) du cumulant d'ordre quatre (voir (A.28) avec $x_1 = x_2 = x_3$). Si l'on multiplie (2.151) par $\mathbf{z}(t_k)$, et que l'on considère l'espérance du résultat, il vient :

$$\mathbf{r}_{zy} = \mathbf{R}_{z\varphi} \boldsymbol{\theta} + \mathbf{r}_{zv} \quad (2.158)$$

Les vecteurs et matrices apparaissant dans (2.158) contiennent des cumulants d'ordre quatre, et il est possible de montrer que le vecteur des instruments $\mathbf{z}(t_k)$ et l'erreur d'équation $v(t_k, \boldsymbol{\theta})$ sont non-corrélés, sous les hypothèses **H4**, **He3**, **Hb1** et **Hb7** [IS94]. De plus, si le signal d'entrée u_o est blanc à l'ordre quatre, la matrice $\mathbf{R}_{z\varphi}$ est non-singulière de manière générique. On obtient alors l'estimateur de la variable instrumentale fondé sur les cumulants d'ordre quatre¹⁷ :

$$\boldsymbol{\theta}_{\text{fociv}}^* = \mathbf{R}_{z\varphi}^{-1} \mathbf{r}_{zy} \quad (2.159)$$

2.3.4 Méthodes fondées sur l'équation du modèle avec cumulants

La multilinéarité des cumulants permet d'énoncer le résultat suivant :

¹⁶D'ordinaire, un bruit blanc (à temps discret) désigne une suite de variables aléatoires centrées décorréelées, c'est-à-dire un bruit blanc à l'ordre 2. Il est possible d'étendre la notion de blancheur aux ordres supérieurs à l'aide des cumulants. Un bruit blanc à l'ordre n est une suite de variables aléatoires centrées dont les cumulants d'ordre inférieur ou égal à n sont nuls, dès lors que toutes les variables aléatoires qu'ils mettent en jeu ne sont pas toutes identiques [BP90].

¹⁷fociv signifie *Fourth-Order Cumulants-based Instrumental Variable*.

PROPOSITION 2.3.1

L'équation du modèle est vérifiée par les cumulants d'ordre trois et quatre :

$$C_{uyu}(\tau_1, \tau_2) = \frac{B(q^{-1}, \boldsymbol{\theta})}{A(q^{-1}, \boldsymbol{\theta})} C_{uuu}(\tau_1, \tau_2) \quad (2.160a)$$

$$C_{uyuu}(\tau_1, \tau_2, \tau_3) = \frac{B(q^{-1}, \boldsymbol{\theta})}{A(q^{-1}, \boldsymbol{\theta})} C_{uuuu}(\tau_1, \tau_2, \tau_3) \quad (2.160b)$$

où l'opérateur retard q^{-1} agit sur τ_1 . ■

PREUVE Voir la [section B.7](#) de l'annexe B. ■

Ces équations sont à la base des méthodes développées. Dans la suite sont présentées deux méthodes que l'on applique à (2.160) : celle des moindres carrés et celle des moindres carrés itératifs. Aussi surprenant que cela puisse paraître, l'application de la méthode des moindres carrés à (2.160) n'a pas été étudiée jusqu'à présent ; seule la méthode des moindres carrés itératifs apparaît dans la littérature.

Ces deux méthodes sont ici développées avec des cumulants d'ordre trois, mais elles peuvent aisément être étendues aux cumulants d'ordre quatre. Une étude approfondie de ces méthodes est donnée dans la [section 3.3](#) du [chapitre 3](#), où leur application dans le cadre de l'identification de modèles à temps continu est étudiée. Nous omettons donc ici volontairement de nombreux aspects et détails de ces algorithmes.

Les hypothèses suivantes sont nécessaires :

He2. u_o a une distribution asymétrique ;

Hb6. \tilde{u} et \tilde{y} ont une distribution symétrique ;

Hs3. $G_o \in \mathcal{G}^*$.

REMARQUE

Dans la suite, pour simplifier les notations, nous noterons $C_{xxx}(\boldsymbol{\tau})$ les cumulants d'ordre trois et $C_{xxxx}(\boldsymbol{\tau})$ ceux d'ordre quatre (le nombre de composantes du vecteur $\boldsymbol{\tau}$ étant déterminé par l'ordre du cumulant). ■

2.3.4.1 Estimateur des moindres carrés (tocls)

À partir de (2.160) on peut définir l'erreur d'équation suivante :

$$v_{cum}(\boldsymbol{\tau}, \boldsymbol{\theta}) = A(q^{-1}, \boldsymbol{\theta})C_{uyu}(\boldsymbol{\tau}) - B(q^{-1}, \boldsymbol{\theta})C_{uuu}(\boldsymbol{\tau}) \quad (2.161)$$

$$= C_{uyu}(\boldsymbol{\tau}) - \boldsymbol{\varphi}_{cum}^T(\boldsymbol{\tau})\boldsymbol{\theta} \quad (2.162)$$

où le vecteur de régression est défini comme :

$$\boldsymbol{\varphi}_{cum}^T(\boldsymbol{\tau}) = \left[-C_{uyu}(\tau_1 - 1, \tau_2) \dots - C_{uyu}(\tau_1 - n_a, \tau_2) \right. \\ \left. C_{uuu}(\tau_1 - 1, \tau_2) \dots C_{uuu}(\tau_1 - n_b, \tau_2) \right] \quad (2.163)$$

L'estimateur des moindres carrés fondé sur les cumulants d'ordre trois minimise une fonction de coût quadratique faisant intervenir l'erreur d'équation (2.161). Par exemple, si l'on choisit de minimiser :

$$J(\tau_2, \boldsymbol{\theta}) = \frac{1}{M} \sum_{\tau_1=1}^M \frac{1}{2} v_{cum}^2(\boldsymbol{\tau}, \boldsymbol{\theta}) \quad (2.164)$$

on obtient l'estimateur des moindres carrés fondés sur les cumulants d'ordre trois¹⁸ :

$$\boldsymbol{\theta}_{\text{tocls}}^*(\tau_2) = \left(\sum_{\tau_1=1}^M \boldsymbol{\varphi}_{cum}(\boldsymbol{\tau}) \boldsymbol{\varphi}_{cum}^T(\boldsymbol{\tau}) \right)^{-1} \sum_{\tau_1=1}^M \boldsymbol{\varphi}_{cum}(\boldsymbol{\tau}) C_{uyu}(\boldsymbol{\tau}) \quad (2.165)$$

2.3.4.2 Estimateur des moindres carrés itératifs (tocls)

À partir de (2.160) on peut définir l'erreur de sortie :

$$\nu_{cum}(\boldsymbol{\tau}, \boldsymbol{\theta}) = C_{uyu}(\boldsymbol{\tau}) - \frac{B(q^{-1}, \boldsymbol{\theta})}{A(q^{-1}, \boldsymbol{\theta})} C_{uuu}(\boldsymbol{\tau}) \quad (2.166)$$

Cette erreur est non linéaire en les paramètres, et pour éviter de recourir à des méthodes d'optimisation numérique, cette erreur de sortie est réécrite sous la forme :

$$\nu_{cum}(\boldsymbol{\tau}, \boldsymbol{\theta}) = \frac{1}{A(q^{-1}, \boldsymbol{\theta})} (A(q^{-1}, \boldsymbol{\theta}) C_{uyu}(\boldsymbol{\tau}) - B(q^{-1}, \boldsymbol{\theta}) C_{uuu}(\boldsymbol{\tau})) \quad (2.167)$$

$$= A(q^{-1}, \boldsymbol{\theta}) C_{uyu}^f(\boldsymbol{\tau}) - B(q^{-1}, \boldsymbol{\theta}) C_{uuu}^f(\boldsymbol{\tau}) \quad (2.168)$$

où C_{uuu}^f et C_{uyu}^f sont les cumulants filtrés par $1/A(q^{-1}, \boldsymbol{\theta})$. On obtient ainsi une erreur d'équation linéaire en les paramètres, qui peuvent alors facilement être estimés par la méthode des moindres carrés. Comme $A(q^{-1}, \boldsymbol{\theta})$ est inconnu, on procède de façon itérative [AG96, AE97]. Cette méthode d'estimation s'inspire de la méthode de Steiglitz-McBride [SM65], appliquée aux cumulants d'ordre trois (au lieu des signaux d'entrée/sortie). La convergence de la méthode de Steiglitz-McBride est très rapide : typiquement quelques itérations suffisent [WP94]. On peut donc s'attendre à ce qu'il en soit de même ici.

2.3.5 Matrice de covariance asymptotique de l'estimateur tocls

L'estimateur $\boldsymbol{\theta}_{\text{tocls}}^*$ proposé dans la section précédente (voir (2.165)) permet d'obtenir des résultats non biaisés tout en restant simple à utiliser. Nous nous sommes attachés à trouver l'expression de la matrice de covariance sur les estimées [THS⁺08], définie par :

$$\lim_{N \rightarrow \infty} N \bar{E} \left\{ \left(\hat{\boldsymbol{\theta}}_{\text{tocls}} - \boldsymbol{\theta} \right) \left(\hat{\boldsymbol{\theta}}_{\text{tocls}} - \boldsymbol{\theta} \right)^T \right\} \quad (2.169)$$

¹⁸tocls signifie *Third-Order Cumulant-based Least Squares*.

où $\hat{\theta}_{\text{tocls}}$ est l'estimée du vecteur des paramètres obtenue à partir des données disponibles, c'est-à-dire calculée à l'aide des cumulants estimés. L'approche utilisée dans [Söd05, Hsz07] pour analyser les méthodes fondées sur les statistiques d'ordre deux est étendue au cas qui nous intéresse ici.

Posons les hypothèses d'application de la méthode tocls :

He2. u_o a une distribution asymétrique ;

Hb7. \tilde{u} et \tilde{y} ont une distribution symétrique.

Il est en outre nécessaire de faire l'hypothèse technique suivante :

Ht1. les moments d'ordre deux, quatre et six du signal d'entrée non-bruité u_o et des bruits \tilde{u} , \tilde{y} décroissent exponentiellement, c'est-à-dire que si un tel moment d'ordre $p + 1$ est donné par $\bar{E}\{x(t_k)x(t_{k+\tau_1})\dots x(t_{k+\tau_p})\}$, alors $\bar{E}\{x(t_k)x(t_{k+\tau_1})\dots x(t_{k+\tau_p})\}$ décroît exponentiellement quand $\max_k |\tau_k| \rightarrow \infty$.

Tout d'abord, réécrivons θ_{tocls}^* sous une forme matricielle :

$$\theta_{\text{tocls}}^* = (\Phi\Phi^T)^{-1} \Phi C \quad (2.170)$$

où :

$$C = \begin{bmatrix} C_{uyu}(0, 0) \\ \vdots \\ C_{uyu}(M-1, 0) \end{bmatrix} \quad \Phi^T = \begin{bmatrix} \varphi_{cum}^T(0, 0) \\ \vdots \\ \varphi_{cum}^T(M-1, 0) \end{bmatrix} \quad (2.171)$$

Les estimées de C et Φ sont notées :

$$\hat{C} = C + \tilde{C} \quad \hat{\Phi} = \Phi + \tilde{\Phi} \quad (2.172)$$

où \tilde{C} et $\tilde{\Phi}$ représentent les erreurs d'estimation correspondantes. On a :

$$\begin{aligned} \hat{\theta}_{\text{tocls}} - \theta &= (\hat{\Phi}\hat{\Phi}^T)^{-1} (\hat{\Phi}\hat{C}) - \theta \\ &= (\hat{\Phi}\hat{\Phi}^T)^{-1} (\hat{\Phi}(\hat{C} - \hat{\Phi}^T\theta)) \\ &= (\hat{\Phi}\hat{\Phi}^T)^{-1} (\Phi(\tilde{C} - \tilde{\Phi}^T\theta) + \tilde{\Phi}(\tilde{C} - \tilde{\Phi}^T\theta)) \end{aligned} \quad (2.173)$$

Supposons que N est grand, de manière à ce que $\|\tilde{C}\| \ll \|C\|$ et $\|\tilde{\Phi}\| \ll \|\Phi\|$. Le lemme suivant

LEMME 2.3.1

Soient R et \tilde{R} deux matrices carrées non singulières telles que $\|\tilde{R}\| \ll \|R\|$. Alors :

$$(R + \tilde{R})^{-1} \approx R^{-1} - R^{-1}\tilde{R}R^{-1} + \mathcal{O}(\|\tilde{R}\|^2) \quad (2.174)$$

■

PREUVE Voir la [section B.6.1](#) de l'annexe B. ■

D'après le lemme 2.3.1 on a :

$$\left(\hat{\Phi}\hat{\Phi}^T\right)^{-1} = \left(\Phi\Phi^T + \Phi\tilde{\Phi}^T + \tilde{\Phi}\Phi^T + \tilde{\Phi}\tilde{\Phi}^T\right)^{-1} \quad (2.175)$$

$$\approx \left(\Phi\Phi^T\right)^{-1} - \left(\Phi\Phi^T\right)^{-1} \left(\Phi\tilde{\Phi}^T + \tilde{\Phi}\Phi^T + \tilde{\Phi}\tilde{\Phi}^T\right) \left(\Phi\Phi^T\right)^{-1} \quad (2.176)$$

Si l'on injecte cette expression dans (2.173) et que l'on néglige les termes de second ordre, on obtient :

$$\hat{\theta}_{\text{tocls}} - \theta \approx \left(\Phi\Phi^T\right)^{-1} \Phi \left(\tilde{C} - \tilde{\Phi}^T \theta\right) \quad (2.177)$$

$$\approx \left(\Phi\Phi^T\right)^{-1} \Phi \left(\hat{C} - \hat{\Phi}^T \theta\right) \quad (2.178)$$

Soit $w = \hat{C} - \hat{\Phi}^T \theta \in \mathbb{R}^M$. Notons que Φ est une matrice déterministe, tandis que w est un terme stochastique de moyenne nulle et qui tend vers 0 lorsque N tend vers l'infini. Sous les hypothèses **H1** et **Ht1**, les cumulants d'ordre trois $C_{uyu}(\tau_1, 0)$ et $C_{uuu}(\tau_1, 0)$ décroissent exponentiellement quand τ_1 tend vers l'infini. Il est maintenant possible d'énoncer le résultat suivant :

THÉORÈME 2.3.1

Sous les hypothèses explicitées en début de section, l'estimée du vecteur des paramètres $\hat{\theta}_{\text{tocls}}$ suit une distribution gaussienne :

$$\sqrt{N}(\hat{\theta}_{\text{tocls}} - \theta) \xrightarrow{\text{dist}} \mathcal{N}(0, P) \quad (2.179)$$

où l'on note :

$$P = \left(\Phi\Phi^T\right)^{-1} \Phi \lim_{N \rightarrow \infty} N \bar{E}\{ww^T\} \Phi^T \left(\Phi\Phi^T\right)^{-1} \quad (2.180)$$

■

PREUVE Voir la section B.6.2 de l'annexe B. ■

Par souci de simplification, introduisons la notation suivante, utilisée dans la suite pour écrire les moments :

$$\mu_{x^r}(\tau) = \bar{E}\{x(t_k)x(t_{k+\tau_1}) \dots x(t_{k+\tau_{r-1}})\} \quad (2.181)$$

Supposons que les moments d'ordre deux, quatre et six du signal d'entrée non-bruité u_o et des bruits \tilde{u} , \tilde{y} décroissent exponentiellement lorsque $\tau \rightarrow \infty$. On peut désormais énoncer le lemme suivant, qui précise la forme que revêt $\lim_{N \rightarrow \infty} N \bar{E}\{ww^T\}$:

LEMME 2.3.2

On a, pour $0 \leq j_1, j_2 \leq M - 1$:

$$\begin{aligned}
\lim_{N \rightarrow \infty} N \bar{E} \{ \mathbf{w}_{j_1+1} \mathbf{w}_{j_2+1} \} &= \sum_{\tau=-\infty}^{\infty} \frac{N - |\tau|}{N^2} \left\{ \right. \\
&\sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_a} a_{l_1} a_{l_2} \left[\mu_{u_o^4}(0, \tau, \tau) \mu_{\tilde{y}^2}(\tau + j_1 - l_1 - j_2 + l_2) + 4 \mu_{u_o^2}(\tau) \mu_{\tilde{u}^2 \tilde{y}^2}(\tau, \tau + j_1 - l_1, j_2 - l_2) \right. \\
&+ \mu_{u_o^2}(0) \left(\mu_{\tilde{u}^2 \tilde{y}^2}(0, \tau + j_1 - l_1, j_2 - l_2) + \mu_{\tilde{u}^2 \tilde{y}^2}(0, j_1 - l_1, j_2 - l_2 - \tau) \right) \\
&+ \left. \mu_{\tilde{u}^4 \tilde{y}^2}(0, \tau, \tau, \tau + j_1 - l_1, j_2 - l_2) \right] \\
&- \sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_b} a_{l_1} b_{l_2} \left[\mu_{u_o^4}(0, \tau, \tau) \left(\mu_{\tilde{u} \tilde{y}}(\tau + j_1 - l_1 - j_2 + l_2) + \mu_{\tilde{u} \tilde{y}}(-\tau - j_1 - l_1 + j_2 + l_2) \right) \right. \\
&+ 4 \mu_{u_o^2}(\tau) \left(\mu_{\tilde{u}^3 \tilde{y}}(\tau, j_2 - l_2, \tau + j_1 - l_1) + \mu_{\tilde{u}^3 \tilde{y}}(\tau, \tau + j_1 - l_2, j_2 - l_1) \right) \\
&+ \mu_{u_o^2}(0) \left(\mu_{\tilde{u}^3 \tilde{y}}(0, j_2 - l_2, \tau + j_1 - l_1) + \mu_{\tilde{u}^3 \tilde{y}}(0, \tau + j_1 - l_2, j_2 - l_1) \right. \\
&+ \left. \mu_{\tilde{u}^3 \tilde{y}}(0, j_2 - l_2 - \tau, j_1 - l_1) + \mu_{\tilde{u}^3 \tilde{y}}(0, j_1 - l_2, j_2 - l_1 - \tau) \right) \\
&+ \left. \mu_{\tilde{u}^5 \tilde{y}}(0, \tau, \tau, j_2 - l_2, \tau + j_1 - l_1) + \mu_{\tilde{u}^5 \tilde{y}}(0, \tau, \tau, \tau + j_1 - l_2, j_2 - l_1) \right] \\
&+ \sum_{l_1=0}^{n_b} \sum_{l_2=0}^{n_b} b_{l_1} b_{l_2} \left[\mu_{u_o^4}(0, \tau, \tau) \mu_{\tilde{u}^2}(\tau + j_1 - l_1 - j_2 + l_2) \right. \\
&+ 4 \mu_{u_o^2}(\tau) \mu_{\tilde{u}^4}(\tau, \tau + j_1 - l_1, j_2 - l_2) + \mu_{u_o^2}(0) \left(\mu_{\tilde{u}^4}(0, \tau + j_1 - l_1, j_2 - l_2) \right. \\
&+ \left. \mu_{\tilde{u}^4}(0, j_1 - l_1, j_2 - l_2 - \tau) \right) + \left. \mu_{\tilde{u}^6}(0, \tau, \tau, \tau + j_1 - l_1, j_2 - l_2) \right] \left. \right\} \quad (2.182)
\end{aligned}$$

■

PREUVE Voir la *section B.6.3 de l'annexe B.*

■

Bien que la méthode *tocls* soit applicable lorsque les bruits sont mutuellement corrélés, par souci de simplification, nous supposons dans la suite que \tilde{u} et \tilde{y} sont indépendants. La seconde double somme apparaissant dans (2.182) disparaît, et l'on en déduit le corollaire suivant :

COROLLAIRE 2.3.1

Lorsque les bruits \tilde{u} et \tilde{y} sont indépendants, (2.182) se simplifie en :

$$\begin{aligned}
\lim_{N \rightarrow \infty} N \bar{E}\{\mathbf{w}_{j_1+1} \mathbf{w}_{j_2+1}\} &= \sum_{\tau=-\infty}^{\infty} \left\{ \sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_a} a_{l_1} a_{l_2} \mu_{\tilde{y}^2}(\tau + j_1 - j_2 - l_1 + l_2) \right. \\
&\left[\mu_{u_0^4}(0, \tau, \tau) + 4\mu_{u_0^2}(\tau) \mu_{\tilde{u}^2}(\tau) + 2\mu_{u_0^2}(0) \mu_{\tilde{u}^2}(0) + \mu_{\tilde{u}^4}(0, \tau, \tau) \right] \\
&+ \sum_{l_1=0}^{n_b} \sum_{l_2=0}^{n_b} b_{l_1} b_{l_2} \left[\mu_{\tilde{u}^6}(0, \tau, \tau, \tau + j_1 - l_1, j_2 - l_2) + 4\mu_{u_0^2}(\tau) \mu_{\tilde{u}^4}(\tau, \tau + j_1 - l_1, j_2 - l_2) \right. \\
&+ \mu_{u_0^2}(0) \left(\mu_{\tilde{u}^4}(0, \tau + j_1 - l_1, j_2 - l_2) + \mu_{\tilde{u}^4}(0, j_1 - l_1, j_2 - l_2 - \tau) \right) \\
&\left. \left. + \mu_{u_0^4}(0, \tau, \tau) \mu_{\tilde{u}^2}(\tau + j_1 - l_1 - j_2 + l_2) \right] \right\} \quad (2.183)
\end{aligned}$$

■

Pour obtenir la matrice de covariance \mathbf{P} donnée par (2.180), il reste maintenant à obtenir les moments d'ordre quatre et six apparaissant dans (2.183).

LEMME 2.3.3

Supposons que :

He1. le signal d'entrée non-bruité u_o a une densité spectrale rationnelle.

Alors les moments d'ordre quatre et six de \tilde{u} sont donnés par :

$$\begin{aligned}
\mu_{\tilde{u}^4}(\boldsymbol{\tau}) &= \mu_{e^{\tilde{u}^2}}^2 \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \left(h_{j_1}^{\tilde{u}} h_{j_1+\tau_1}^{\tilde{u}} h_{j_2}^{\tilde{u}} h_{j_2+\tau_2-\tau_3}^{\tilde{u}} + h_{j_1}^{\tilde{u}} h_{j_1+\tau_2}^{\tilde{u}} h_{j_2}^{\tilde{u}} h_{j_2+\tau_1-\tau_3}^{\tilde{u}} + h_{j_1}^{\tilde{u}} h_{j_1+\tau_3}^{\tilde{u}} h_{j_2}^{\tilde{u}} h_{j_2+\tau_2-\tau_1}^{\tilde{u}} \right) \\
&+ (\mu_{e^{\tilde{u}^4}} - 3\mu_{e^{\tilde{u}^2}}^2) \sum_{j_1=0}^{\infty} h_{j_1}^{\tilde{u}} h_{j_1+\tau_1}^{\tilde{u}} h_{j_1+\tau_2}^{\tilde{u}} h_{j_1+\tau_3}^{\tilde{u}} \quad (2.184)
\end{aligned}$$

et :

$$\begin{aligned}
\mu_{\tilde{u}^6}(0, \tau, \tau, \tau_1, \tau_2) &= \mu_{e^{\tilde{u}^2}}^3 \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \sum_{j_3=0}^{\infty} \sum_{\tau_a, \tau_b, \tau_c} h_{j_1}^{\tilde{u}} h_{j_1+\tau_a}^{\tilde{u}} h_{j_2}^{\tilde{u}} h_{j_2+\tau_b}^{\tilde{u}} h_{j_3}^{\tilde{u}} h_{j_3+\tau_c}^{\tilde{u}} \\
&+ (\mu_{e^{\tilde{u}^2}} \mu_{e^{\tilde{u}^4}} - 3\mu_{e^{\tilde{u}^2}}^3) \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \sum_{\tau_d, \tau_e, \tau_f, \tau_g} h_{j_1}^{\tilde{u}} h_{j_1+\tau_d}^{\tilde{u}} h_{j_2}^{\tilde{u}} h_{j_2+\tau_e}^{\tilde{u}} h_{j_2+\tau_f}^{\tilde{u}} h_{j_2+\tau_g}^{\tilde{u}} \\
&+ (\mu_{e^{\tilde{u}^6}} - 15\mu_{e^{\tilde{u}^2}} \mu_{e^{\tilde{u}^4}} + 30\mu_{e^{\tilde{u}^2}}^3) \sum_{j_1=0}^{\infty} (h_{j_1}^{\tilde{u}})^2 (h_{j_1+\tau}^{\tilde{u}})^2 h_{j_1+\tau_1}^{\tilde{u}} h_{j_1+\tau_2}^{\tilde{u}} \quad (2.185)
\end{aligned}$$

où τ_a, \dots, τ_g sont donnés dans les tableaux B.1 et B.1 de l'annexe B. ■

PREUVE Voir la section B.6.4 de l'annexe B. ■

Les moments de u_o peuvent être obtenus de la même manière, en supposant que u_o est lui aussi obtenu en sortie d'un filtre ARMA excité par un bruit blanc.

2.3.6 Simulations numériques

Dans cette partie, nous comparons les méthodes **tociv**, **tocls** et **tocils**, et vérifions la concordance de la distribution asymptotique des estimées de **tocls** et des résultats de simulation. Il convient de noter que l'étude des méthodes **tocls** et **tocils** est plutôt succincte (elles sont analysées plus avant dans le chapitre suivant).

Comme dans les simulations numériques précédentes, les performances des différents algorithmes sont étudiées à l'aide de simulations de Monte Carlo (de $n_{mc} = 100$ réalisations). Pour avoir une mesure d'ensemble de la qualité d'un estimateur, l'EQMR est calculée (voir (2.137)). Le même critère d'arrêt est adopté pour chacun des algorithmes itératifs :

$$\frac{\|\hat{\boldsymbol{\theta}}^{i+1} - \hat{\boldsymbol{\theta}}^i\|}{\|\hat{\boldsymbol{\theta}}^i\|} < 10^{-3} \quad \text{ou} \quad i > 50 \quad (2.186)$$

2.3.6.1 Analyse comparative des méthodes **tociv**, **tocls** et **tocils**

Le système considéré est caractérisé par¹⁹ :

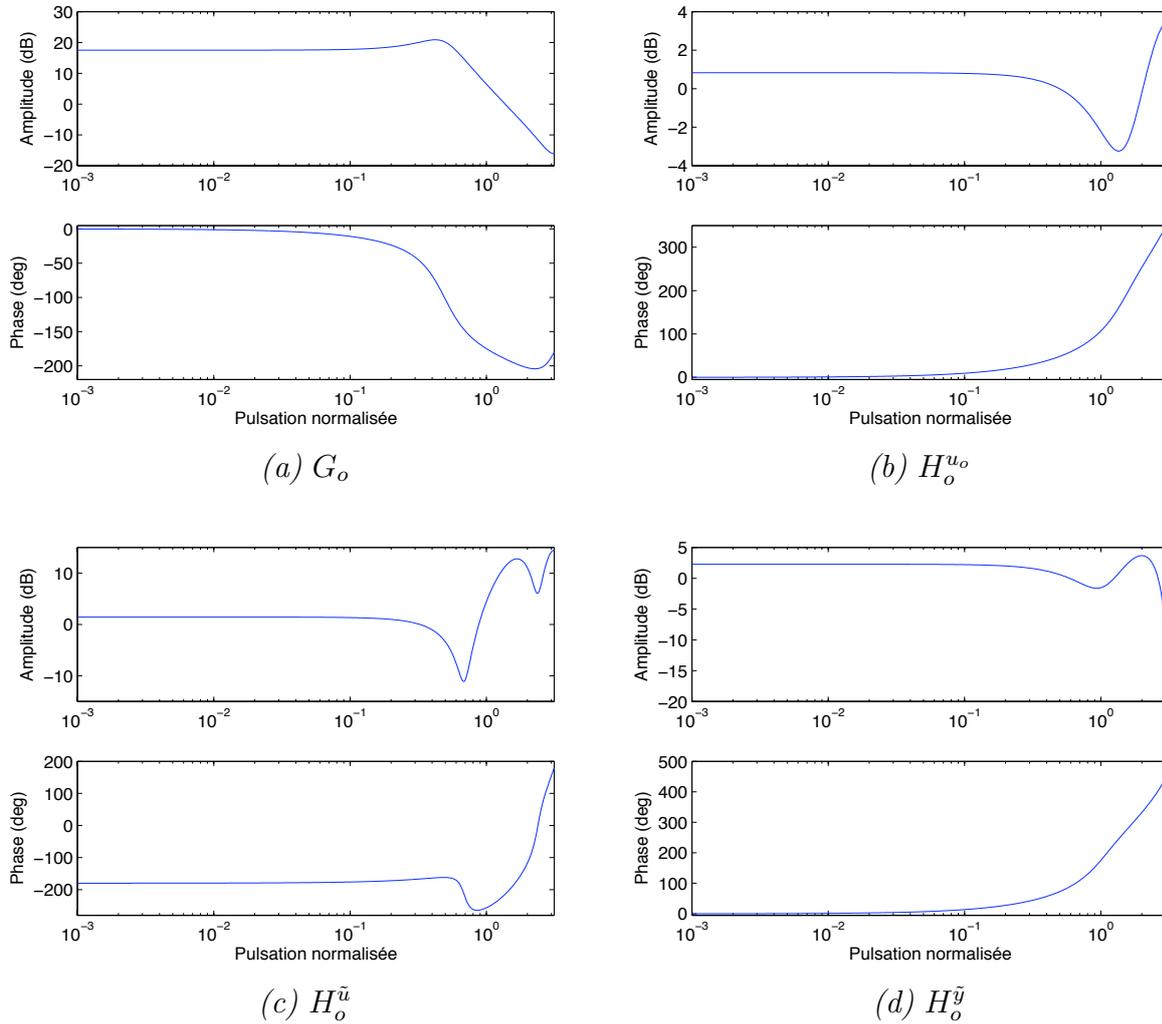
$$\begin{cases} G_o(q) = \frac{1.0q^{-1} + 0.5q^{-2}}{1 - 1.5q^{-1} + 0.7q^{-2}} \\ H_o^{u_o}(q) = 1 - 0.2q^{-1} + 0.3q^{-2} \\ H_o^{\tilde{u}}(q) = 1 - 2.33q^{-1} + 0.75q^{-2} + 0.5q^{-3} + 0.3q^{-4} - 1.4q^{-5} \\ H_o^{\tilde{y}}(q) = 1 + 0.2q^{-1} - 0.3q^{-2} + 0.4q^{-3} \end{cases} \quad (2.187)$$

Les bruits blancs gaussiens $e_o^{\tilde{u}}$ et $e_o^{\tilde{y}}$ ont une variance ajustée de manière à obtenir un RSB de 10 dB en entrée et en sortie. Les diagrammes de Bode des filtres $H_o^{u_o}$, $H_o^{\tilde{u}}$, $H_o^{\tilde{y}}$ et G_o sont représentés à la figure 2.8. Pour évaluer l'influence du nombre de données sur les résultats, les méthodes sont appliquées à trois jeux de données de longueurs respectives $N = 1000$, $N = 5000$ et $N = 10000$. Les résultats apparaissent dans le tableau 2.6 et les diagrammes de Bode des modèles estimés avec $N = 5000$ points sont représentés à la figure 2.9.

Discussion

On constate tout d'abord que la longueur du jeu de données a une grande importance, puisque lorsque $N = 1000$ les algorithmes n'offrent pas d'estimées acceptables ; le biais est non négligeable et la variance est assez grande. En revanche, à mesure que le nombre de données augmente, les estimées des cumulants d'ordre trois deviennent meilleures, ce qui conduit à de très bons résultats. Rappelons qu'avec des bruits colorés en entrée et en sortie, la plupart des méthodes utilisant les statistiques de second ordre sont condamnées à échouer. La méthode fondée sur une variable instrumentale (**tociv**) fournit des résultats avec une plus grande variance que les algorithmes **tocls** et **tocils**. Enfin, notons que les estimées obtenues par **tocils** sont un peu plus dispersées que celles de **tocls**, mais ont un plus petit biais.

¹⁹Le système considéré et les conditions de simulation s'inspirent de [AG96].

FIG. 2.8 – Diagrammes de Bode de $H_o^{u_o}$, $H_o^{\tilde{u}}$, $H_o^{\tilde{y}}$ et G_o 

2.3.6.2 Matrice de covariance asymptotique

Le système considéré est caractérisé par :

$$\begin{cases} G_o(q) = \frac{1.0q^{-1} + 0.5q^{-2}}{1 - 1.5q^{-1} + 0.7q^{-2}} \\ H_o^{\tilde{u}}(q) = \frac{1}{1 - 0.7q^{-1}} \end{cases} \quad \begin{cases} H_o^{u_o}(q) = \frac{1 + 0.7q^{-1}}{1 - 0.5q^{-1}} \\ H_o^{\tilde{y}}(q) = \frac{1 + 0.7q^{-1}}{1 - 0.7q^{-1}} \end{cases} \quad (2.188)$$

Le signal e^{u_o} en entrée de $H_o^{u_o}(q)$ est un bruit blanc exponentiel de moyenne nulle et de variance unité. Les moments d'ordre trois et quatre de e^{u_o} sont donc $\mu_{e^{u_o}3} = 2$ et $\mu_{e^{u_o}4} = 9$. Les bruits $e^{\tilde{u}}$, $e^{\tilde{y}}$ en entrée de $H_o^{\tilde{u}}(q)$ et $H_o^{\tilde{y}}(q)$ sont des bruits blancs gaussiens de moyenne nulle et de variance unité, dont les moments d'ordre quatre et six sont donc respectivement égaux à 3 et 15.

Les simulations de Monte Carlo (200 réalisations) ont été effectuées avec un jeu de données de longueur $N = 50000$. L'hyper-paramètre M est choisi égal à 5.

tandis que la matrice de covariance asymptotique théorique est :

$$P = \begin{pmatrix} 80.13 & & & \\ -58.14 & 43.51 & & \\ -12.84 & 6.56 & 20.62 & \\ 109.54 & -77.22 & -30.73 & 166.95 \end{pmatrix}$$

On constate une bonne concordance des résultats théoriques et des résultats de simulation (voir aussi le [tableau 2.7](#), qui présente les valeurs des paramètres estimés ainsi l'écart-type théorique et celui obtenu d'après les simulations).

TAB. 2.7 – Comparaison des résultats théoriques et des résultats de simulation pour la méthode `tocls` ($N = 50000$)

θ	-1.5	0.7	1.0	-0.5
$\hat{\theta}_{\text{tocls}}$	-1.49919	0.69954	0.99951	-0.49901
Sim.	± 0.00166	± 0.00089	± 0.00043	± 0.00339
Théor.	± 0.00160	± 0.00087	± 0.00041	± 0.00333

2.4 Conclusion

Ce chapitre a traité de l'identification de modèles à temps discret avec des erreurs en les variables.

Nous nous sommes dans un premier temps intéressés aux méthodes utilisant les statistiques d'ordre deux. Après un bref tour d'horizon, au cours duquel nous sommes revenus sur les principales méthodes développées, un accent particulier a été mis sur les méthodes de compensation du biais de l'estimateur des moindres carrés. Ces méthodes sont intéressantes car, sans avoir à recourir à une optimisation numérique et généralement avec un faible coût algorithmique, elles fournissent des estimées dont la précision est proche de celle obtenue avec une méthode à erreur de prédiction. Une étude des différents algorithmes de compensation du biais de l'estimateur des moindres carrés a donc été effectuée, en insistant sur la clé de voûte de ces algorithmes, à savoir la manière dont les propriétés statistiques des bruits sont estimées. Nous avons ensuite proposé des estimateurs fondés sur la technique de la variable instrumentale, ainsi que des algorithmes cherchant à compenser leur éventuel biais, selon le vecteur d'instruments choisi. En effet, puisque dans le modèle EIV les signaux d'entrée/sortie sont tous deux corrompus par un bruit additif, le vecteur d'instruments est à choisir précautionneusement. Une analyse comparative des différents algorithmes de compensation de biais, ainsi qu'une étude des performances des méthodes fondées sur une variable

instrumentale, ont alors été réalisées par l'intermédiaire d'exemples de simulation.

Dans un second temps, nous nous sommes focalisés sur les méthodes fondées sur les statistiques d'ordre supérieur. Des estimateurs fondés sur la minimisation de critères faisant intervenir les cumulants d'ordre trois et quatre sont tout d'abord évoqués. Ensuite, des estimateurs utilisant une variable instrumentale ont été étudiés : il est en effet possible de définir des vecteurs d'instruments « particuliers », conduisant à l'utilisation de cumulants d'ordre supérieur. Nous nous sommes alors intéressés aux estimateurs des moindres carrés et des moindres carrés itératifs fondés sur l'équation du système, dont on peut montrer qu'elle est vérifiée par les cumulants. Enfin, nous nous sommes attachés à trouver l'expression de la matrice de covariance asymptotique de l'estimateur des moindres carrés fondé sur les cumulants d'ordre trois, que nous avons proposé auparavant. Des exemples de simulation ont permis d'illustrer les performances des estimateurs de la variable instrumentale, des moindres carrés et des moindres carrés itératifs, ainsi que de constater la concordance entre la matrice de covariance théorique et celle obtenue par simulation.

De ces études et analyses il ressort :

- l'importance des *hypothèses structurelles* pour les méthodes fondées sur les statistiques d'ordre deux. Une erreur de modélisation des bruits ou du signal d'entrée non-bruité peut en effet conduire à une estimation défailante du vecteur des paramètres ;
- l'importance des *hypothèses distributionnelles* pour les méthodes fondées sur les statistiques d'ordre supérieur. En effet, si la structure des modèles des bruits ou de l'entrée non-bruitée n'ont dans ce cas que peu d'importance, les hypothèses sur leurs distributions sont en revanche fondamentales.

Les estimateurs fondés sur les statistiques d'ordre deux et ceux fondés sur les statistiques d'ordre supérieur ont leurs avantages et inconvénients respectifs, et il serait insensé de privilégier les uns ou les autres. Lors du choix entre les différentes méthodes, les conditions expérimentales sont alors déterminantes : longueur du jeu de données, vraisemblance des hypothèses structurelles/distributionnelles, *etc.*

Perspectives

Le travail présenté dans ce chapitre a permis d'ouvrir de nombreuses perspectives, pour beaucoup spécifiques à chacune des méthodes présentées. Certaines d'entre elles n'ont été développées que pour le cas de bruits blancs en entrée et en sortie du système. L'extension de ces méthodes à des situations plus générales est donc une première perspective, à laquelle peuvent s'ajouter les quelques autres, plus spécifiques, données ci-après :

- lors de l'étude des méthodes de compensation de biais, nous avons vu qu'il existe de nombreuses équations à partir desquelles il est possible d'estimer les variances des bruits (ou leurs fonctions d'autocovariance lorsqu'ils ne sont pas blancs). Existe-t-il des équations plus « fiables » que d'autres, permettant d'obtenir des estimées plus précises, et pourquoi ?

- nous avons constaté que les méthodes de compensation de biais utilisant les données filtrées par un filtre MA (**fbcls** et **fbciv**) semblent fournir de meilleurs résultats que les méthodes ne faisant pas intervenir de filtrage. Est-ce que le filtrage peut être vu comme un moyen « générique » d'améliorer les performances de méthodes? Il serait également intéressant d'étudier comment choisir ce filtre de manière optimale;
- la ligne de cumulants utilisée par l'estimateur de la variable instrumentale **tociv** est déterminée par le vecteur d'instruments choisi. Ce choix peut être effectué de diverses manières, et il serait intéressant d'étudier les relations entre cet estimateur et celui des moindres carrés fondés sur les cumulants (**tocls**);
- une étude approfondie de la matrice de covariance asymptotique permet-elle de justifier l'utilisation de cumulants particuliers? Par exemple, peut-on trouver une justification théorique de l'utilisation de cumulants du type C_{yyy} plutôt que C_{uyu} ?

Chapitre 3

Identification de modèles à temps continu

3.1 Introduction et position du problème

Si, comme nous avons pu le voir au chapitre précédent, l'identification de modèles EIV à temps discret a fait l'objet de nombreux travaux au cours des dernières années, l'identification de modèles à temps continu n'a en revanche été que très peu étudiée. Ce domaine reste donc très largement ouvert. Pour identifier un modèle à temps continu à partir de données d'entrée/sortie échantillonnées, deux approches principales sont possibles : l'approche indirecte et l'approche directe [UR90]. L'approche indirecte consiste à déterminer, dans un premier temps, un modèle à temps discret à l'aide de techniques maintenant conventionnelles [Lju99], [SS89], [WP94], puis à convertir ce dernier en un modèle à temps continu. L'approche directe vise quant à elle à obtenir un modèle à temps continu en une seule étape, à l'aide de méthodes spécifiquement développées pour l'identification de modèles à temps continu.

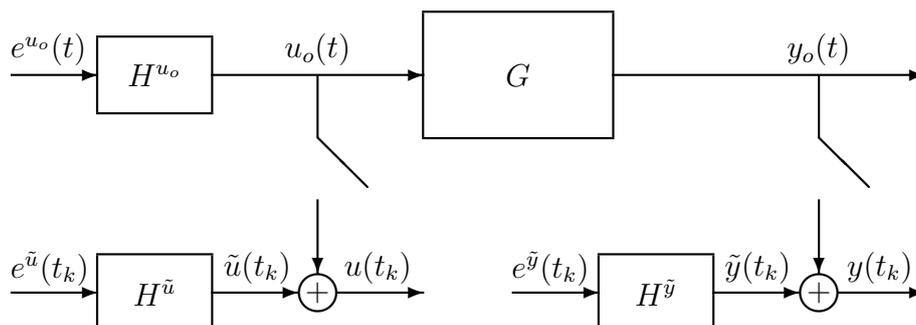
La plupart des systèmes réels régis par les lois de la physique (systèmes électroniques, mécaniques, thermiques, hydrauliques, ...) évoluent par nature continûment dans le temps. Leur comportement peut être décrit approximativement après linéarisation autour d'un point de fonctionnement par des équations différentielles ordinaires linéaires à coefficients constants. Cette représentation dans le domaine continu est donc naturelle pour l'étude des systèmes linéaires. Comparativement à l'approche indirecte, l'identification directe de modèles à temps continu présente de nombreux avantages, dont les principaux sont :

- de fournir directement un modèle à temps continu à partir de données échantillonnées. Les paramètres du modèle identifié sont fortement liés aux propriétés et aux coefficients physiques du système à identifier. Ce modèle peut être interprété physiquement beaucoup plus aisément que l'équivalent discret délivré par les approches indirecte d'identification ;
- de s'appuyer sur une paramétrisation parcimonieuse. La connaissance *a priori* de l'ordre

- relatif du système est facilement prise en compte ce qui permet d'éviter l'estimation de zéros de discrétisation rencontrés lors de l'identification de modèles à temps discret [ÅHS84, Lju03];
- de pouvoir plus naturellement traiter des données échantillonnées rapidement. Ces approches sont en effet beaucoup moins sensibles aux problèmes numériques que les approches discrètes. Elles sont par conséquent très bien adaptées aux équipements actuels d'acquisition qui permettent un échantillonnage rapide des données;
 - de pouvoir aisément traiter le cas de données échantillonnées à pas variable [HG02]. Ce type de données est souvent rencontré dans le cas des systèmes mécaniques où un échantillonnage angulaire est réalisé, ou encore dans le domaine biomédical [WFS98] ou biologique;
 - de simplifier l'application de la procédure complète d'identification de systèmes pour l'utilisateur. La théorie de l'identification de modèles à temps discret est bien établie; la pratique, quant à elle, n'est pas toujours simple à mettre en œuvre pour l'utilisateur qui se trouve souvent démuné devant les choix multiples à effectuer (signal d'excitation, période d'échantillonnage, filtrage des signaux bruités, ...) avant d'effectuer l'estimation paramétrique. L'étape de pré-filtrage est souvent primordiale dans le cas de l'identification de modèles à temps discret. Les approches directes intègrent toutes un pré-filtrage implicite des données, l'application de la méthodologie d'identification se trouve alors grandement simplifiée pour l'utilisateur.

Dans la suite du chapitre, nous allons nous intéresser à ces techniques d'identification directe de modèles EIV à temps continu. Après avoir introduit les notations utilisées, nous présentons succinctement les rares méthodes développées jusqu'à présent pour l'identification de modèles EIV à temps continu à partir de statistiques d'ordre deux. Plusieurs estimateurs fondés sur les statistiques d'ordre supérieur pour identifier les modèles EIV à temps continu sont ensuite proposés. Enfin, avant de conclure, les performances des algorithmes proposés sont illustrées à l'aide de simulations numériques.

FIG. 3.1 – Schéma détaillé du modèle EIV à temps continu



On considère un système à temps continu, linéaire et invariant. Selon les notations introduites au chapitre 1, les signaux d'entrée/sortie non-bruités sont liés par :

$$\check{y}_o(t) = G_o(p)\check{u}_o(t) \quad (3.1)$$

où p est l'opérateur différentiel et $G_o(p)$ l'opérateur de transfert du vrai système. On suppose que les signaux $u_o(t)$ et $y_o(t)$ sont échantillonnés aux instants $\{t_k = kT_e\}_{k=1}^N$, et que les signaux échantillonnés mesurés sont perturbés par des bruits à temps discret $\tilde{u}(t_k)$ et $\tilde{y}(t_k)$. Le système à temps continu générant les données est donc caractérisé par :

$$\mathcal{S} : \begin{cases} \check{y}_o(t) = G_o(p)\check{u}_o(t) \\ u(t_k) = \check{u}_o(t_k) + \tilde{u}(t_k) \\ y(t_k) = \check{y}_o(t_k) + \tilde{y}(t_k) \end{cases} \quad (3.2)$$

Les modèles paramétriques recherchés ont une forme hybride, au sens où le modèle du système est à temps continu, tandis que les modèles des bruits en entrée/sortie sont à temps discret. Considérons par conséquent la structure suivante, dont le schéma est donné à la [figure 3.1](#) :

$$\mathcal{M}(\Theta) : \begin{cases} y_o(t) = G(p, \theta)u_o(t) \\ u(t_k) = u_o(t_k) + \tilde{u}(t_k) \\ y(t_k) = y_o(t_k) + \tilde{y}(t_k) \end{cases} \text{ avec } \begin{cases} u_o(t) = H^{u_o}(p, \boldsymbol{\eta}_{u_o})e^{u_o}(t) \\ \tilde{u}(t_k) = H^{\tilde{u}}(q, \boldsymbol{\eta}_{\tilde{u}})e^{\tilde{u}}(t_k) \\ \tilde{y}(t_k) = H^{\tilde{y}}(q, \boldsymbol{\eta}_{\tilde{y}})e^{\tilde{y}}(t_k) \end{cases} \quad (3.3a)$$

où $e^{u_o}(t)$ est la dérivée d'un processus de Wiener de variance incrémentale unité, et $e^{\tilde{u}}(t_k)$, $e^{\tilde{y}}(t_k)$ désignent des bruits blancs à temps discret, dont les variances sont respectivement notées $\lambda_{e^{\tilde{u}}}$ et $\lambda_{e^{\tilde{y}}}$. Les ensembles de modèles sont ainsi :

$$\mathcal{M}^* = \left\{ G(p, \theta), H^{u_o}(p, \boldsymbol{\eta}_{u_o}), H^{\tilde{u}}(q, \boldsymbol{\eta}_{\tilde{u}}), H^{\tilde{y}}(q, \boldsymbol{\eta}_{\tilde{y}}) \mid \Theta \in D_\Theta \right\} \quad (3.3b)$$

$$\mathcal{G}^* = \left\{ G(p, \theta) = \frac{B(p, \theta)}{A(p, \theta)} = \frac{b_0 + b_1p + \dots + b_{n_b}p^{n_b}}{a_0 + a_1p + \dots + a_{n_a-1}p^{n_a-1} + p^{n_a}} \mid \theta \in D_\theta \right\} \quad (3.3c)$$

$$\mathcal{E}^* = \left\{ H^{u_o}(p, \boldsymbol{\eta}_{u_o}) \mid \boldsymbol{\eta}_{u_o} \in D_{\boldsymbol{\eta}_{u_o}} \right\} \quad (3.3d)$$

$$\mathcal{H}^* = \left\{ (H^{\tilde{u}}(q, \boldsymbol{\eta}_{\tilde{u}}), H^{\tilde{y}}(q, \boldsymbol{\eta}_{\tilde{y}})) \mid \boldsymbol{\eta}_{\tilde{u}} \in D_{\boldsymbol{\eta}_{\tilde{u}}}, \boldsymbol{\eta}_{\tilde{y}} \in D_{\boldsymbol{\eta}_{\tilde{y}}} \right\} \quad (3.3e)$$

Les vecteurs des paramètres sont donnés par :

$$\boldsymbol{\theta}^T = [\mathbf{a}^T \ \mathbf{b}^T] = [a_0 \ \dots \ a_{n_a-1} \ b_0 \ \dots \ b_{n_b}] \quad (3.4)$$

$$\Theta^T = [\boldsymbol{\theta}^T \ \boldsymbol{\eta}_{u_o}^T \ \boldsymbol{\eta}_{\tilde{u}}^T \ \boldsymbol{\eta}_{\tilde{y}}^T] \quad (3.5)$$

3.2 Estimateurs fondés sur les statistiques de second ordre

3.2.1 Succinct état de l'art

Dans cette partie, nous revenons sur les rares travaux effectués jusqu'à présent pour identifier directement un modèle à temps continu à partir de données d'entrée/sortie échantillonnées, perturbées par des bruits additifs. Plutôt que de décrire chacune des méthodes, nous donnons ici un tableau regroupant leurs principales caractéristiques : elles sont en effet lourdes à présenter, tout en semblant être d'une efficacité limitée.

Un des problèmes majeurs de l'identification directe de modèles à temps continu est la nécessité d'obtenir les dérivées des signaux d'entrée/sortie. Toute tentative utilisant une approximation numérique de la dérivation résulte en effet en une accentuation des bruits. Ce problème est habituellement évité en réalisant un pré-traitement des données. On peut distinguer trois classes de pré-traitement [Men99, GMR03] :

- les filtres linéaires ;
- les méthodes du filtre intégral linéaire ;
- les fonctions de modulation.

Dans la majeure partie des cas, la présentation des méthodes à temps continu met l'accent sur les aspects de mise en œuvre. Les méthodes sont en effet rendues plus complexes par le pré-traitement utilisé. Par exemple, l'expression du biais de la méthode des moindres carrés, dans le cas de bruits blancs en entrée et en sortie du système, dépend de leurs variances et du vecteur des paramètres recherché. Lorsque l'on effectue un pré-traitement à base de filtrage, l'expression de ce biais dépend alors également d'une manière non triviale du filtre employé.

Néanmoins, la mise en œuvre des méthodes peut être effectuée de différentes manières, et une manière plus générale de présenter les méthodes est de séparer la méthode d'estimation proprement dite du pré-traitement utilisé. Dans le tableau ci-dessous, les travaux sont classés suivant la méthode d'estimation employée, l'hypothèse structurelle sous laquelle elles fournissent une estimation convergente de θ_o , les hypothèses sur les bruits en entrée/sortie (E/S) du système, le pré-traitement choisi, et les inconnues supplémentaires à estimer en sus du vecteur des paramètres θ .

Méthode d'estimation	Compens. biais LS	Compens. biais IV	LS séparables
Hyp. structurelle	Hs2	Hs2	Hs2
Bruits en E/S	Blanc/blanc	Blanc/coloré	Blanc/blanc
Pré-traitement	Filtre intégral lin.	Filtre intégral lin.	Fonctions de mod.
Inconnues suppl.	$\lambda_{e\tilde{u}}, \lambda_{e\tilde{y}}$	$\lambda_{e\tilde{u}}$	$\lambda_{e\tilde{u}}, \lambda_{e\tilde{y}}$
Référence	[SYW91]	[YSW93]	[SLMM06]

REMARQUES

- les méthodes de compensation du biais de l'estimateur des moindres carrés présentées au chapitre précédent estiment les variances $\lambda_{e\tilde{u}}$, $\lambda_{e\tilde{y}}$ en résolvant un système linéaire. La méthode de [SYW91] procède de la même manière, et la première équation de ce système est identique (elle provient de la corrélation des résidus de la méthode des moindres carrés). La deuxième équation est en revanche originale et utilise la corrélation des résidus d'un estimateur fondé sur une variable instrumentale. Cela montre une fois de plus qu'il existe de nombreuses manières d'estimer les variances des bruits blancs perturbant les signaux d'entrée/sortie ;
- dans [SLMM06] sont également exposées les méthodes de l'erreur de prédiction et du maximum de vraisemblance. Le modèle à temps continu est dans un premier temps représenté sous forme d'état ; ce modèle d'état est ensuite échantillonné. À partir de ce modèle d'état à temps discret, il est alors possible de définir les estimateurs de l'erreur de prédiction et du maximum de vraisemblance. Il convient toutefois de noter que les matrices du modèle d'état à temps discret dépendent d'une manière plus complexe de $\boldsymbol{\theta}$, le vecteur des paramètres du modèle à temps continu, impliquant une fonction plus délicate encore à optimiser. Ces deux méthodes présentent en effet des risques inhérents (optimiseurs locaux), d'autant plus grands que le rapport signal-à-bruit est faible [SLMM06]. Par ailleurs, la méthode à erreur de prédiction nécessite une étape intermédiaire. En effet, les erreurs de prédiction sont déterminées par le gain du filtre de Kalman, et celui-ci est insensible à une multiplication par une constante des matrices de covariance des bruits. Pour lever ce degré de liberté, il est possible de fixer un des paramètres des bruits à une valeur constante lors de l'optimisation, et ensuite d'utiliser également la valeur minimale du critère. Voir [Söd06] pour des détails. ■

3.2.2 Méthode fondée sur les moindres carrés séparables (eivsvf)

Le principe de cette méthode est d'appliquer l'approche traditionnelle du filtre des variables d'état (voir [You81] ou la section 3.3.3) au modèle EIV, puis d'analyser les effets des bruits en entrée et en sortie du système sur le filtre des variables d'état pour obtenir une estimation convergente du vecteur des paramètres [MG05, MG06].

Les hypothèses d'application de cette méthode sont :

Hb1. Les bruits en entrée et en sortie \tilde{u} et \tilde{y} sont non-corrélés ;

Hb2. Les bruits en entrée et en sortie sont blancs ;

Hs2. $G_o \in \mathcal{G}^*$ et $(H_o^{\tilde{u}}, H_o^{\tilde{y}}) \in \mathcal{H}^*$.

Soit :

$$F(p) = \sum_{j=0}^{n_a+1} f_j p^j \quad (3.6)$$

un filtre linéaire invariant. Pré-filtrons les données d'entrée/sortie non-bruitées pour obtenir les signaux w et r :

$$\left(\sum_{j=0}^{n_a+1} f_j p^j \right) r(t) = u_o(t) \quad \left(\sum_{j=0}^{n_a+1} f_j p^j \right) w(t) = y_o(t) \quad (3.7)$$

Il est immédiat que les signaux w et r vérifient l'équation différentielle du système :

$$\sum_{j=0}^{n_a} a_j w^{(j)}(t) = \sum_{j=0}^{n_b} b_j r^{(j)}(t) \quad (3.8)$$

ce que l'on peut écrire sous la forme :

$$\phi_o^T(t) \bar{\theta} = 0 \quad (3.9)$$

où l'on note ($x^{(n)}$ représente la $n^{\text{ième}}$ dérivée de x) :

$$\phi_o^T(t) = [-w^{(0)}(t) \dots -w^{(n_a)}(t) \ r^{(0)}(t) \dots r^{(n_b)}(t)] \quad (3.10)$$

$$\bar{\theta}^T = [a_0 \dots a_{n_a+1} \ 1 \ b_0 \dots b_{n_b}] \quad (3.11)$$

Introduisons les matrices suivantes, définies à l'aide des signaux échantillonnés :

$$\Phi = \sum_{k=1}^N \phi(t_k) \phi^T(t_k) \quad (3.12)$$

$$\Phi_o = \sum_{k=1}^N \phi_o(t_k) \phi_o^T(t_k) \quad (3.13)$$

$$\tilde{\Phi} = \sum_{k=1}^N \mathbb{E} \left\{ \tilde{\phi}(t_k) \tilde{\phi}^T(t_k) \right\} \quad (3.14)$$

Alors, si N est assez grand, on peut décomposer Φ en une partie bruitée et une partie non-bruitée¹, et par conséquent :

$$\Phi^T \bar{\theta} = \tilde{\Phi}^T \bar{\theta} \quad (3.15)$$

Il est en outre possible de montrer que :

PROPOSITION 3.2.1 ([MG06])

Soit $\{t_k\}$ une suite strictement croissante de \mathbb{R}^+ . Supposons que les moments d'ordre quatre des bruits blancs \tilde{u} et \tilde{y} soient bornés. Alors :

$$\lim_{N \rightarrow \infty} \frac{1}{t_N} \left(\Phi - \Phi_o - \tilde{\Phi} \right) = 0 \quad p.s.$$

■

¹Si N est assez grand $\Phi \simeq \Phi_o + \tilde{\Phi}$ car les termes croisés tendent vers zéro lorsque $N \rightarrow \infty$.

Par ailleurs, l'analyse de l'implantation numérique du filtre des variables d'état permet d'isoler la contribution des bruits. En effet, en utilisant la proposition précédente et l'équation (3.15) il vient :

$$\Phi\bar{\theta} = \check{\Phi}\Sigma\bar{\theta} + \mathcal{O}(\gamma_N) \quad (3.16)$$

où :

- $\check{\Phi}$ est une matrice diagonale par bloc dont les éléments sont donnés par l'implantation numérique du filtre des variables d'état, mis sous forme d'état. Ces éléments sont dépendants de l'hypothèse réalisée sur l'évolution du signal entre les instants d'échantillonnage (voir [MG05, MG06] pour plus de détails) ;
- γ_N est un terme stochastique qui converge presque sûrement vers 0 lorsque $N \rightarrow \infty$;
- $\Sigma = \begin{bmatrix} \lambda_{e\bar{u}}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \lambda_{e\bar{y}}\mathbf{I} \end{bmatrix}$ est la matrice contenant la contribution des bruits blancs.

Pour exploiter (3.16) il est encore nécessaire de s'affranchir de la matrice Σ , inconnue. Pour ce faire, l'idée est de pré-filtrer les données avec deux filtres différents : on obtient par conséquent deux équations du type (3.16), qu'il est possible d'utiliser pour estimer conjointement le vecteur des paramètres et les variances des bruits blancs. En effet, si l'on note $\phi_j(t_k)$ le vecteur de régression associé au $j^{\text{ième}}$ filtre ($j = 1, 2$) et :

$$\Phi_{ij} = \sum_{k=1}^N \phi_i(t_k)\phi_j^T(t_k) \quad (3.17)$$

on obtient :

$$\frac{1}{t_N} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} \bar{\theta} & \mathbf{0} \\ \mathbf{0} & \bar{\theta} \end{bmatrix} = \frac{1}{t_N} \begin{bmatrix} \check{\Phi}_{11} & \check{\Phi}_{12} \\ \check{\Phi}_{21} & \check{\Phi}_{22} \end{bmatrix} \begin{bmatrix} \Sigma\bar{\theta} & \mathbf{0} \\ \mathbf{0} & \Sigma\bar{\theta} \end{bmatrix} + \mathcal{O}(\gamma_N) \quad (3.18)$$

Il est possible d'écrire cette équation sous la forme :

$$\Psi\bar{\theta} = \check{\Psi}\Sigma\bar{\theta} + \mathcal{O}(\gamma_N) \quad (3.19)$$

avec :

$$\Psi = \frac{1}{t_N} [\Phi_{11}^T \ \Phi_{12}^T \ \Phi_{21}^T \ \Phi_{22}^T]^T \quad \check{\Psi} = \frac{1}{t_N} [\check{\Phi}_{11}^T \ \check{\Phi}_{12}^T \ \check{\Phi}_{21}^T \ \check{\Phi}_{22}^T]^T \quad (3.20)$$

L'équation (3.19) est alors résolue au sens des moindres carrés pour obtenir des estimées de $\bar{\theta}$, $\lambda_{e\bar{u}}$ et $\lambda_{e\bar{y}}$:

$$[\bar{\theta}^T \ \lambda_{e\bar{u}} \ \lambda_{e\bar{y}}]^T = \arg \min_{\bar{\theta}, \lambda_{e\bar{u}}, \lambda_{e\bar{y}}} \|\Psi\bar{\theta} - \check{\Psi}\Sigma\bar{\theta}\|^2 \quad (3.21)$$

Puisque (3.19) est bilinéaire en Σ et $\bar{\theta}$, le problème d'optimisation (3.21) peut être résolu de manière itérative. L'algorithme de minimisation est donné ci-dessous.

1. Choisir une valeur initiale appropriée $\hat{\theta}^0$ et fixer $i = 0$;

2. Dans (3.21), fixer $\hat{\boldsymbol{\theta}} = [1 \ \hat{\boldsymbol{\theta}}^i]$ puis minimiser par rapport à $\lambda_{e\bar{u}}$ et $\lambda_{e\bar{y}}$;
3. Fixer $\lambda_{e\bar{u}}$ et $\lambda_{e\bar{y}}$ aux valeurs estimées à l'étape précédente, et minimiser (3.21) par rapport à $\boldsymbol{\theta}$;
4. Incrémenter i et revenir à l'étape 2. tant que

$$\left\| \begin{bmatrix} \hat{\boldsymbol{\theta}}^i - \hat{\boldsymbol{\theta}}^{i-1} \\ \hat{\lambda}_{e\bar{u}}^i - \hat{\lambda}_{e\bar{u}}^{i-1} \\ \hat{\lambda}_{e\bar{y}}^i - \hat{\lambda}_{e\bar{y}}^{i-1} \end{bmatrix} \right\|$$

est supérieure à un critère d'arrêt préalablement défini.

REMARQUES

- les estimations réalisées au cours des étapes 2. et 3. sont effectuées par moindres carrés classiques ;
- une étude de la convergence de cet estimateur, une expression de la matrice de covariance des estimées et une méthode permettant d'obtenir une initialisation « appropriée » sont également disponibles dans [MG06]. ■

3.3 Estimateurs fondés sur les statistiques d'ordre supérieur

Comme nous venons de le voir, l'identification de modèles EIV à temps continu a fait l'objet de peu de travaux jusqu'à présent. Hormis la méthode de compensation du biais de la variable instrumentale [YSW93], les méthodes font l'hypothèse que les bruits perturbant les signaux d'entrée/sortie sont blancs. Cette hypothèse de blancheur permet non seulement de simplifier l'analyse, mais aussi d'éliminer certains problèmes d'identifiabilité lorsque l'on a recours aux statistiques d'ordre deux [AD84, AG06]. En outre, le pré-traitement des données résulte bien souvent en la perte de la blancheur des bruits et mène ainsi à une complexité accrue. Les approches fondées sur les statistiques d'ordre supérieur (SOS) sont donc intéressantes d'une part pour leur « insensibilité » aux bruits, indépendamment de leur structure, et d'autre part pour s'affranchir des problèmes d'identifiabilité.

Si plusieurs méthodes utilisant les SOS ont été proposées dans le cadre de l'identification de modèles EIV à temps discret (voir la section 2.3.1 du chapitre 2), aucune n'a en revanche été développée pour le cas des modèles à temps continu. Nous nous intéressons ici particulièrement à la méthode mentionnée au section 2.3.4, et présentons son extension à l'identification de modèles EIV à temps continu, ainsi que les spécificités liées à la mise en œuvre des algorithmes proposés.

3.3.1 Méthodes fondées sur l'équation du modèle avec cumulants

La multilinéarité des cumulants permet d'énoncer le résultat suivant :

PROPOSITION 3.3.1 ([TGG08A])

L'équation du modèle est vérifiée par les cumulants d'ordre trois et quatre :

$$C_{uyu}(\tau_1, \tau_2) = \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} C_{uuu}(\tau_1, \tau_2) \quad (3.22a)$$

$$C_{uyuu}(\tau_1, \tau_2, \tau_3) = \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} C_{uuuu}(\tau_1, \tau_2, \tau_3) \quad (3.22b)$$

où l'opérateur différentiel p représente $\frac{\partial}{\partial \tau_1}$. ■

PREUVE Voir la [section B.7](#) de l'[annexe B](#). ■

REMARQUE

Comme au [chapitre 2](#), pour simplifier les notations, nous noterons $C_{xxx}(\boldsymbol{\tau})$ les cumulants d'ordre trois et $C_{xxxx}(\boldsymbol{\tau})$ ceux d'ordre quatre (le nombre de composantes du vecteur $\boldsymbol{\tau}$ étant déterminé par l'ordre du cumulant). ■

Ces équations sont à la base des méthodes développées : à partir de (3.22) il est possible d'utiliser toute méthode « classique » pour obtenir une estimée du vecteur des paramètres $\boldsymbol{\theta}$. En outre, puisque les cumulants d'ordre trois (resp. quatre) sont nuls pour les signaux distribués de manière symétrique (resp. gaussiens), si l'on pose les hypothèses idoines sur le signal d'entrée non-bruitée et sur les bruits, (3.22) mène à :

$$C_{u_o y_o u_o}(\boldsymbol{\tau}) = \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} C_{u_o u_o u_o}(\boldsymbol{\tau}) \quad (3.23a)$$

$$C_{u_o y_o u_o u_o}(\boldsymbol{\tau}) = \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} C_{u_o u_o u_o u_o}(\boldsymbol{\tau}) \quad (3.23b)$$

On voit donc – si les hypothèses adéquates sont vérifiées – que l'utilisation de SOS permet de travailler non plus avec les signaux d'entrée/sortie bruités, mais avec les cumulants, théoriquement insensibles aux bruits. Toutefois, en pratique, les cumulants sont estimés à partir des N échantillons de données disponibles. Par conséquent, on a :

$$\hat{C}_{uyu}(\boldsymbol{\tau}) = \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} \hat{C}_{uuu}(\boldsymbol{\tau}) + \varepsilon_3(\boldsymbol{\tau}, \boldsymbol{\theta}) \quad (3.24a)$$

$$\hat{C}_{uyuu}(\boldsymbol{\tau}) = \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} \hat{C}_{uuuu}(\boldsymbol{\tau}) + \varepsilon_4(\boldsymbol{\tau}, \boldsymbol{\theta}) \quad (3.24b)$$

avec :

$$\varepsilon_3(\boldsymbol{\tau}, \boldsymbol{\theta}) = \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} \tilde{C}_{uuu}(\boldsymbol{\tau}) - \tilde{C}_{uyu}(\boldsymbol{\tau}) \quad (3.25a)$$

$$\varepsilon_4(\boldsymbol{\tau}, \boldsymbol{\theta}) = \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} \tilde{C}_{uuuu}(\boldsymbol{\tau}) - \tilde{C}_{uyuu}(\boldsymbol{\tau}) \quad (3.25b)$$

où $\tilde{C}(\boldsymbol{\tau}) = C(\boldsymbol{\tau}) - \hat{C}(\boldsymbol{\tau})$ représente l'erreur d'estimation de $C(\boldsymbol{\tau})$. Puisque les estimées des cumulants sont non-biaisées et convergentes (voir la section A.5.2 de l'annexe A), il vient pour $j = 3, 4$:

$$\lim_{N \rightarrow \infty} \varepsilon_j(\boldsymbol{\tau}, \boldsymbol{\theta}) = 0 \quad \text{p.s.} \quad (3.26)$$

Dans la suite sont présentés deux estimateurs utilisés pour estimer $\boldsymbol{\theta}$ à partir de (3.22) : celui des moindres carrés et celui des moindres carrés itératifs [TGGM07, TGG08a].

REMARQUE

Il est possible de travailler avec d'autres cumulants, comme par exemple :

$$C_{yyy}(\boldsymbol{\tau}) = \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} C_{yuy}(\boldsymbol{\tau}) \quad (3.27a)$$

$$C_{yyyy}(\boldsymbol{\tau}) = \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} C_{yuyy}(\boldsymbol{\tau}) \quad (3.27b)$$

S'il paraît plus approprié de travailler avec les cumulants faisant intervenir le plus possible le signal de sortie (ils contiennent plus d'information sur la dynamique du système), il n'y a toutefois pas de justification théorique à ce jour. Les développements effectués dans la suite s'appuient sur (3.22). Ils sont néanmoins directement transposables à d'autres cumulants, comme par exemple ceux de (3.27). ■

3.3.2 Méthodes fondées sur les cumulants d'ordre trois

Les hypothèses suivantes sont nécessaires :

He2. u_o a une distribution asymétrique ;

Hb6. \tilde{u} et \tilde{y} ont une distribution symétrique ;

Hs3. $G_o \in \mathcal{G}^*$.

3.3.2.1 Estimateur des moindres carrés (tocls)

L'équation (3.24a) peut être écrite sous la forme d'une régression linéaire :

$$\hat{C}_{uyu}^{(n_a)}(\boldsymbol{\tau}) = \hat{\boldsymbol{\Phi}}^T(\boldsymbol{\tau})\boldsymbol{\theta} + A(p, \boldsymbol{\theta})\varepsilon_3(\boldsymbol{\tau}, \boldsymbol{\theta}) \quad (3.28)$$

où le vecteur de régression est donné par :

$$\hat{\boldsymbol{\Phi}}^T(\boldsymbol{\tau}) = \left[-\hat{C}_{uyu}^{(0)}(\boldsymbol{\tau}) \quad \dots \quad -\hat{C}_{uyu}^{(n_a-1)}(\boldsymbol{\tau}) \quad \hat{C}_{uuu}^{(0)}(\boldsymbol{\tau}) \quad \dots \quad \hat{C}_{uuu}^{(n_b)}(\boldsymbol{\tau}) \right] \quad (3.29)$$

$$\text{avec} \quad \hat{C}_{uyu}^{(j)}(\boldsymbol{\tau}) = \frac{\partial^j}{\partial \tau_1^j} \hat{C}_{uyu}(\boldsymbol{\tau}) \quad (3.30)$$

On définit alors l'erreur d'équation suivante :

$$v_{cum}(\boldsymbol{\tau}, \boldsymbol{\theta}) = \hat{C}_{uyu}^{(n_a)}(\boldsymbol{\tau}) - \hat{\boldsymbol{\Phi}}^T(\boldsymbol{\tau})\boldsymbol{\theta} \quad (3.31)$$

$$= B(p, \boldsymbol{\theta})\tilde{C}_{uuu}(\boldsymbol{\tau}) - A(p, \boldsymbol{\theta})\tilde{C}_{uyu}(\boldsymbol{\tau}) \quad (3.32)$$

Minimiser la fonction de coût ci-après par rapport au vecteur des paramètres $\boldsymbol{\theta}$:

$$V(\tau_2, \boldsymbol{\theta}, M) = \frac{1}{M} \sum_{\tau_1=0}^{M-1} \frac{1}{2} v_{cum}^2(\boldsymbol{\tau}, \boldsymbol{\theta}) \quad (3.33)$$

conduit à l'estimateur `tocls`², à savoir :

$$\hat{\boldsymbol{\theta}}_{\text{tocls}}(\tau_2, M) = \left[\frac{1}{M} \sum_{\tau_1=0}^{M-1} \hat{\boldsymbol{\Phi}}(\boldsymbol{\tau}) \hat{\boldsymbol{\Phi}}^T(\boldsymbol{\tau}) \right]^{-1} \left[\frac{1}{M} \sum_{\tau_1=0}^{M-1} \hat{\boldsymbol{\Phi}}(\boldsymbol{\tau}) \hat{C}_{uyu}^{(n_a)}(\boldsymbol{\tau}) \right] \quad (3.34)$$

Le « bruit » apparaissant dans le membre de droite de l'erreur d'équation (3.32) est composé des erreurs d'estimation \tilde{C}_{uuu} et \tilde{C}_{uyu} filtrées par A et B . Ce bruit n'est par conséquent pas blanc, et l'on ne s'attend donc pas à ce que l'estimée $\hat{\boldsymbol{\theta}}_{\text{tocls}}$ soit non biaisée, à partir d'un nombre fini de données. Toutefois, lorsque N tend vers l'infini, (3.22) fait uniquement intervenir des termes sans erreur : l'estimée $\hat{\boldsymbol{\theta}}_{\text{tocls}}$ est donc asymptotiquement non biaisée.

Le rôle de l'hyper-paramètre M apparaissant dans la fonction de coût (3.33) est de prévenir l'utilisation de cumulants ayant un grand délai. Il est en effet possible de calculer cette fonction de coût avec toutes les données disponibles, *i.e.* en fixant la borne supérieure de la somme de l'équation (3.33) à N . Cependant, puisque les cumulants sont estimés par des moyennes, lorsque le délai augmente, ces estimées sont calculées avec moins de points, devenant ainsi moins fiables (voir la section A.5 de l'annexe A). Notons qu'il est nécessaire que $M \geq n_a + n_b + 1$ afin que la matrice à inverser dans (3.34) ne soit pas singulière.

REMARQUE

Il est possible de choisir la fonction de coût de différentes manières ; (3.33) peut notamment être remplacée par :

$$V_2(\boldsymbol{\theta}, M_1, M_2) = \frac{1}{M_1 M_2} \sum_{\tau_1=0}^{M_1-1} \sum_{\tau_2=0}^{M_2-1} \frac{1}{2} v_{cum}^2(\boldsymbol{\tau}, \boldsymbol{\theta}) \quad (3.35)$$

ce qui mène à l'estimateur :

$$\hat{\boldsymbol{\theta}}_{V_2}(M_1, M_2) = \left[\frac{1}{M_1 M_2} \sum_{\tau_1=0}^{M_1-1} \sum_{\tau_2=0}^{M_2-1} \hat{\boldsymbol{\Phi}}(\boldsymbol{\tau}) \hat{\boldsymbol{\Phi}}^T(\boldsymbol{\tau}) \right]^{-1} \left[\frac{1}{M_1 M_2} \sum_{\tau_1=0}^{M_1-1} \sum_{\tau_2=0}^{M_2-1} \hat{\boldsymbol{\Phi}}(\boldsymbol{\tau}) \hat{C}_{uyu}^{(n_a)}(\boldsymbol{\tau}) \right] \quad (3.36)$$

Cette question est liée au choix des lignes de cumulants : par définition les cumulants d'ordre trois (de processus stationnaires) sont bi-dimensionnels, dans le sens où ils dépendent de deux variables (ici τ_1 et τ_2). À partir des données disponibles il est donc possible d'obtenir une matrice constituée des cumulants d'ordre trois calculés avec différents délais. Le second estimateur $\hat{\boldsymbol{\theta}}_{V_2}(M_1, M_2)$ peut alors être interprété comme la solution de la minimisation de

²Pour *Third-Order Cumulants-based Least Squares algorithm*.

l'erreur d'équation (3.32) en considérant une sous-matrice de dimension $M_1 \times M_2$, tandis que le premier estimateur $\hat{\boldsymbol{\theta}}_{\text{tocls}}(\tau_2, M)$ est obtenu en ne considérant qu'une ligne de cette matrice; d'où sa dépendance de τ_2 . Lors des simulations numériques, le second estimateur $\hat{\boldsymbol{\theta}}_{V_2}(M_1, M_2)$ n'a pas fourni d'amélioration significative, tout en ayant un coût algorithmique plus important (voir aussi la section A.5.1).

Au final, l'estimateur $\hat{\boldsymbol{\theta}}_{\text{tocls}}$ fait intervenir une ligne de cumulants très simple : le vecteur obtenu en posant $\tau_2 = 0$ et $0 \leq \tau_1 \leq M$. L'influence de l'horizon de calcul des cumulants M est discuté à la section 3.4. ■

3.3.2.2 Estimateur des moindres carrés itératifs (tocils)

Les équations (3.24b)-(3.25b) définissent une erreur de sortie donnée par :

$$\nu_{\text{cum}}(\boldsymbol{\tau}, \boldsymbol{\theta}) = \hat{C}_{uyu}(\boldsymbol{\tau}) - \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} \hat{C}_{uuu}(\boldsymbol{\tau}) \quad (3.37)$$

$$= \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})} \tilde{C}_{uuu}(\boldsymbol{\tau}) - \tilde{C}_{uyu}(\boldsymbol{\tau}) \quad (3.38)$$

Cette erreur est non-linéaire en les paramètres. Plutôt que d'effectuer une optimisation non-linéaire, l'approche proposée est fondée sur une procédure des moindres carrés itératifs. Cette approche a été suggérée pour l'identification de modèle EIV à temps discret [AG96], et est étendue ici au cas des modèles à temps continu.

L'erreur de sortie (3.37) peut s'écrire :

$$\nu_{\text{cum}}(\boldsymbol{\tau}, \boldsymbol{\theta}) = \frac{1}{A(p, \boldsymbol{\theta})} \left(A(p, \boldsymbol{\theta}) \hat{C}_{uyu}(\boldsymbol{\tau}) - B(p, \boldsymbol{\theta}) \hat{C}_{uuu}(\boldsymbol{\tau}) \right) \quad (3.39)$$

$$= A(p, \boldsymbol{\theta}) \hat{C}_{uyu, f}(\boldsymbol{\tau}) - B(p, \boldsymbol{\theta}) \hat{C}_{uuu, f}(\boldsymbol{\tau}) \quad (3.40)$$

où $\hat{C}_{uuu, f}$ et $\hat{C}_{uyu, f}$ représentent les cumulants filtrés par $1/A(p, \boldsymbol{\theta})$. On obtient ainsi une erreur d'équation, linéaire en les paramètres, qui peuvent alors être estimés par la méthode des moindres carrés.

Puisque $A(p, \boldsymbol{\theta})$ est inconnu, on procède de manière itérative pour transformer l'erreur d'équation (3.40) en l'erreur de sortie (3.37). Soit $\hat{\boldsymbol{\theta}}^i$ l'estimée de $\boldsymbol{\theta}$ obtenue lors de la $i^{\text{ième}}$ itération. À chaque itération, $\hat{\boldsymbol{\theta}}^{i+1}$ est calculée par moindres carrés, en utilisant les cumulants filtrés par $1/A(p, \hat{\boldsymbol{\theta}}^i)$, en vérifiant auparavant la stabilité du filtre. Dans la mesure où les paramètres convergent vers une valeur constante, on a :

$$\frac{A(p, \hat{\boldsymbol{\theta}}^{i+1})}{A(p, \hat{\boldsymbol{\theta}}^i)} \longrightarrow 1 \qquad \frac{B(p, \hat{\boldsymbol{\theta}}^{i+1})}{A(p, \hat{\boldsymbol{\theta}}^i)} \longrightarrow \frac{B(p, \hat{\boldsymbol{\theta}}^{i+1})}{A(p, \hat{\boldsymbol{\theta}}^{i+1})}$$

Par conséquent, l'erreur d'équation (3.40) tend vers l'erreur d'équation (3.37). Cette méthode itérative s'inspire de celle de Steiglitz-McBride [SM65]. Il y a néanmoins une différence majeure : la méthode de Steiglitz-McBride applique l'estimateur des moindres carrés itératifs

sur les données d'entrée/sortie mesurées plutôt que sur les cumulants. Elle est par conséquent applicable sous des hypothèses assez restrictives (bruit blanc en sortie et aucun bruit en entrée) comme il a été montré dans [SS81b]. La méthode exposée ci-avant applique cette procédure itérative aux cumulants d'ordre trois, et comme ceux-ci sont insensibles aux bruits ayant une distribution symétrique, ces restrictions ne sont pas nécessaires. La convergence de l'algorithme de Steiglitz-McBride est très rapide : typiquement quelques itérations suffisent [WP94]. On peut donc s'attendre à ce qu'il en soit de même ici.

L'erreur d'équation (3.40) peut être mise sous la forme d'une régression linéaire :

$$\nu_{cum}(\boldsymbol{\tau}, \boldsymbol{\theta}) = \hat{C}_{uyu,f}^{(n_a)}(\boldsymbol{\tau}) - \hat{\boldsymbol{\Phi}}_f^T(\boldsymbol{\tau})\boldsymbol{\theta} \quad (3.41)$$

où le vecteur de régression est donné par :

$$\hat{\boldsymbol{\Phi}}_f^T(\boldsymbol{\tau}) = \left[-\hat{C}_{uyu,f}^{(0)}(\boldsymbol{\tau}) \quad \dots \quad -\hat{C}_{uyu,f}^{(n_a-1)}(\boldsymbol{\tau})\hat{C}_{uuu,f}^{(0)}(\boldsymbol{\tau}) \quad \dots \quad \hat{C}_{uuu,f}^{(n_b)}(\boldsymbol{\tau}) \right] \quad (3.42)$$

La minimisation du critère suivant à chaque itération :

$$V^i(\tau_2, \boldsymbol{\theta}, M) = \frac{1}{M} \sum_{\tau_1=0}^{M-1} \frac{1}{2} (\nu_{cum}^i(\boldsymbol{\tau}, \boldsymbol{\theta}))^2 \quad (3.43)$$

où ν_{cum}^i est donné par :

$$\nu_{cum}^i(\boldsymbol{\tau}, \boldsymbol{\theta}) = A(p, \boldsymbol{\theta}) \frac{\hat{C}_{uyu}(\boldsymbol{\tau})}{A(p, \hat{\boldsymbol{\theta}}^i)} - B(p, \boldsymbol{\theta}) \frac{\hat{C}_{uuu}(\boldsymbol{\tau})}{A(p, \hat{\boldsymbol{\theta}}^i)} \quad (3.44)$$

mène à l'estimateur **tocils**³ :

$$\hat{\boldsymbol{\theta}}_{tocils}(\tau_2, M) = \left[\frac{1}{M} \sum_{\tau_1=0}^{M-1} \hat{\boldsymbol{\Phi}}_f(\boldsymbol{\tau}) \hat{\boldsymbol{\Phi}}_f^T(\boldsymbol{\tau}) \right]^{-1} \left[\frac{1}{M} \sum_{\tau_1=0}^{M-1} \hat{\boldsymbol{\Phi}}_f(\boldsymbol{\tau}) \hat{C}_{uyu,f}^{(n_a)}(\boldsymbol{\tau}) \right] \quad (3.45)$$

3.3.3 Mise en œuvre des algorithmes **tocls** et **tocils**

Les dérivées des cumulants sont nécessaires pour construire les vecteurs de régression (3.29), (3.42). Nous avons choisi de les calculer en utilisant l'approche du filtre des variables d'état, dont le principe est rappelé dans cette partie (voir [You64, You81] pour plus de détails). Nous ne mentionnons ici que le cas des cumulants d'ordre trois, mais les cumulants d'ordre quatre sont traités de manière similaire.

L'approche du filtre des variables d'état permet de reconstruire les dérivées (jusqu'à un ordre n_a) d'un signal en le passant dans le banc de filtres suivant :

$$F_n(p) = p^n \left(\frac{\lambda}{p + \lambda} \right)^{n_a}, \quad 0 \leq n \leq n_a \quad (3.46)$$

³Pour *Third-Order Cumulants-based Iterative Least Squares algorithm*.

où λ est un hyper-paramètre symbolisant la pulsation de coupure du filtre. Intuitivement, elle peut être choisie pour privilégier une bande fréquentielle particulière, et d'une manière générale, il est préférable de la choisir légèrement supérieure à la borne supérieure de la bande passante du système à identifier.

Ainsi, les approximations des dérivées filtrées d'un signal x sont obtenues en sortie du banc de filtres :

$$x_f^{(n)}(t) = F_n(p)x(t) \quad (3.47)$$

$$= [f_n \star x](t) \quad (3.48)$$

où \star représente l'opérateur de convolution et $f_n(t)$ la réponse impulsionnelle du filtre linéaire $F_n(s)$:

$$f_n(t) = \mathcal{L}^{-1}[F_n(s)] \quad (3.49)$$

avec \mathcal{L} la transformée de Laplace.

Les vecteurs de régression (3.29)-(3.42) sont donc obtenus en utilisant l'approche du filtre des variables d'état. Or, la multilinéarité des cumulants implique qu'un cumulant filtré peut être obtenu à partir du signal filtré [LAC97, NP93] :

$$\begin{aligned} C_{xxx,f}^{(n)}(\boldsymbol{\tau}) &= F_n(p)C_{xxx}(\boldsymbol{\tau}) \quad (3.50) \\ &= \int_{\mathbb{R}} f_n(\tau)C_{xxx}(\tau_1 - \tau, \tau_2, \tau_3)d\tau \\ &= \int_{\mathbb{R}} f_n(\tau)\bar{\mathbb{E}}\{x(t)x(t + \tau_1 - \tau)x(t + \tau_2)\}d\tau \\ &= \bar{\mathbb{E}}\left\{x(t)\left(\int_{\mathbb{R}} f_n(\tau)x(t + \tau_1 - \tau)d\tau\right)x(t + \tau_2)\right\} \\ &= \bar{\mathbb{E}}\left\{x(t)x_f^{(n)}(t + \tau_1)x(t + \tau_2)\right\} \quad (3.51) \end{aligned}$$

On obtient donc :

$$C_{xxx,f}^{(n)}(\boldsymbol{\tau}) = C_{xx_f^{(n)}x}(\boldsymbol{\tau}) \quad (3.52)$$

L'utilisation de l'approche du filtre des variables d'état, c'est-à-dire l'application aux deux membres de (3.22) des filtres $\{F_n(p)\}_{n=0}^{n_{max}}$, permet ainsi de transférer le problème de l'estimation des dérivées de cumulants en un problème d'estimation des dérivées des signaux d'entrée/sortie, tâche bien connue de l'identification de modèle à temps continu [GMR03].

Pour l'algorithme des moindres carrés, le banc de filtres invariants (3.46) est utilisé pour filtrer les signaux d'entrée/sortie. Leurs dérivées dans la bande fréquentielle d'intérêt sont ainsi obtenues, à partir desquelles les dérivées des cumulants sont calculées. En revanche, pour l'algorithme des moindres carrés itératifs, le banc de filtres (3.46) n'est utilisé que lors de l'étape d'initialisation (l'algorithme des moindres carrés itératifs est donc initialisé avec

l'algorithme des moindres carrés). Lors des itérations suivantes, comme une estimée de θ est disponible, il est possible d'utiliser le banc de filtres suivant :

$$F_n(p, \hat{\theta}^i) = \frac{p^n}{A(p, \hat{\theta}^i)}, \quad 0 \leq n \leq n_a \quad (3.53)$$

Notons que ce banc de filtres présente l'avantage de réaliser la dérivation ainsi que le filtrage par $1/A(p, \theta)$ en une seule étape.

On peut désormais donner les étapes principales des algorithmes *tocls* et *tocils*.

L'algorithme *tocls*

1. Définir un banc de filtre des variables d'état d'ordre n_a :

$$F_n(p) = p^n \left(\frac{\lambda}{p + \lambda} \right)^{n_a}, \quad 0 \leq n \leq n_a$$

Générer les dérivées $u_f^{(n)}(t_k)$ et $y_f^{(n)}(t_k)$ des signaux d'entrée/sortie par :

$$\begin{cases} u_f^{(n)}(t_k) = F_n(p)u(t_k), & 0 \leq n \leq n_b \\ y_f^{(n)}(t_k) = F_n(p)y(t_k), & 0 \leq n \leq n_a \end{cases}$$

2. Calculer les cumulants d'ordre trois et leurs dérivées :

$$\hat{C}_{uuu,f}^{(n)}(\tau) = \frac{1}{N - \mu} \sum_{k=1}^{N-\mu} u(t_k)u_f^{(n)}(t_{k+\tau_1})u(t_{k+\tau_2}), \quad 0 \leq n \leq n_b$$

$$\hat{C}_{uyu,f}^{(n)}(\tau) = \frac{1}{N - \mu} \sum_{k=1}^{N-\mu} u(t_k)y_f^{(n)}(t_{k+\tau_1})u(t_{k+\tau_2}), \quad 0 \leq n \leq n_a$$

où $\mu = \max(\tau_1, \tau_2)$;

3. Construire le vecteur de régression :

$$\hat{\Phi}_f^T(\tau) = \left[-\hat{C}_{uyu,f}^{(0)}(\tau) \dots -\hat{C}_{uyu,f}^{(n_a-1)}(\tau) \quad \hat{C}_{uuu,f}^{(0)}(\tau) \dots \hat{C}_{uuu,f}^{(n_b)}(\tau) \right]$$

Calculer l'estimée *tocls* (3.34) :

$$\hat{\theta}_{\text{tocls}}(\tau_2, M) = \left(\frac{1}{M} \sum_{\tau_1=0}^{M-1} \hat{\Phi}_f(\tau) \hat{\Phi}_f^T(\tau) \right)^{-1} \left(\frac{1}{M} \sum_{\tau_1=0}^{M-1} \hat{\Phi}_f(\tau) \hat{C}_{uyu,f}^{(n_a)}(\tau) \right)$$

L'algorithme *tocils*

1. Initialisation. La première estimée $\hat{\theta}_{\text{tocils}}^0(\tau_2, M)$ est donnée par l'algorithme *tocls* :

$$\hat{\theta}_{\text{tocils}}^0(\tau_2, M) = \hat{\theta}_{\text{tocls}}(\tau_2, M)$$

2. $(i + 1)^{\text{ième}}$ étape. Comme une estimée $\hat{\boldsymbol{\theta}}_{\text{tocils}}^i(\tau_2, M)$ est disponible, procéder aux mêmes étapes que pour l'algorithme *tocls*, mais en utilisant le banc de filtres suivant :

$$F_n(p, \hat{\boldsymbol{\theta}}_{\text{tocils}}^i(\tau_2, M)) = \frac{p^n}{A(p, \hat{\boldsymbol{\theta}}_{\text{tocils}}^i(\tau_2, M))}, \quad 0 \leq n \leq n_a$$

pour obtenir $\hat{\boldsymbol{\theta}}_{\text{tocils}}^{i+1}(\tau_2, M)$;

3. Itérer jusqu'à ce que la différence entre deux estimées successives satisfasse le critère d'arrêt choisi.

3.3.4 Méthodes fondées sur les cumulants d'ordre quatre (**focls**, **focils**)

Nous proposons, dans cette partie, d'adoucir la restriction sur la distribution du signal d'entrée non bruité en utilisant les cumulants d'ordre quatre. En effet, l'hypothèse sur u_o est dans ce cas simplement d'être non gaussien (au lieu d'avoir une distribution asymétrique). Le pendant est une hypothèse légèrement plus forte sur les bruits, qui doivent quant à eux être gaussiens (au lieu d'avoir une distribution symétrique). Il est ainsi nécessaire de supposer que :

He3. u_o est non-gaussien ;

Hb7. \tilde{u} et \tilde{v} sont gaussiens ;

Hs3. $G_o \in \mathcal{G}^*$.

Les développements exposés dans la [section 3.3.2](#) restent vrais, et ne sont donc pas répétés ici. Les deux algorithmes obtenus sont appelés **focls** et **focils**⁴.

Les principales différences entre le cas des cumulants d'ordre trois et celui des cumulants d'ordre quatre sont liées aux aspects pratiques : l'estimation de cumulants et le choix des lignes de cumulants. Ayant à disposition N échantillons de données d'entrée/sortie $\{u(t_k), y(t_k)\}_{k=1}^N$, les cumulants d'ordre quatre sont estimés à l'aide des k -statistiques (voir

⁴pour *Fourth-Order Cumulants-based (Iterative) Least Squares*.

l'annexe A) :

$$\hat{C}_{xxxx}(\boldsymbol{\tau}) = \frac{1}{(N-1)(N-2)(N-3)} \left\{ \begin{aligned} & N(N+1) \sum_{k=1}^N x(t_k)x(t_k+\tau_1)x(t_k+\tau_2)x(t_k+\tau_3) \\ & - 3(N-1) \sum_{k=1}^N x(t_k)x(t_k+\tau_1) \sum_{k=1}^N x(t_k+\tau_2)x(t_k+\tau_3) \\ & - 3(N-1) \sum_{k=1}^N x(t_k)x(t_k+\tau_2) \sum_{k=1}^N x(t_k+\tau_1)x(t_k+\tau_3) \\ & - 3(N-1) \sum_{k=1}^N x(t_k)x(t_k+\tau_3) \sum_{k=1}^N x(t_k+\tau_1)x(t_k+\tau_2) \end{aligned} \right\} \quad (3.54)$$

Nous avons déjà mentionné à la fin du [paragraphe 3.3.2.1](#) qu'il est nécessaire de choisir des lignes de cumulants plutôt que d'utiliser tous les cumulants disponibles. Lorsque les cumulants d'ordre trois ont été étudiés, nous n'avons en effet utilisé qu'une partie de la « matrice » contenant les cumulants d'ordre trois calculés avec différents délais. Ici le nombre de cumulants disponibles est encore augmenté : un ordre de plus signifie une dimension supplémentaire, et on obtient ainsi un cube contenant les cumulants d'ordre quatre calculés avec différents délais. De la même manière qu'auparavant, on se limite à n'utiliser qu'une partie de ce cube : d'une part on évite l'utilisation de cumulants ayant un délai trop grand (à l'aide de l'horizon de calcul M), et d'autre part uniquement une partie de ce « sous-cube » est utilisée (à l'aide de la ligne de cumulants choisie). Au final, une ligne de cumulants très simple est à nouveau choisie : le vecteur obtenu en posant $\tau_2 = \tau_3 = 0$ et $0 \leq \tau_1 \leq M$.

Ci-dessous sont données les étapes principales des deux algorithmes focls et focils.

L'algorithme focls

1. Définir un banc de filtre des variables d'état d'ordre n_a :

$$F_n(p) = p^n \left(\frac{\lambda}{p + \lambda} \right)^{n_a}, \quad 0 \leq n \leq n_a$$

Générer les dérivées $u_f^{(n)}(t_k)$ et $y_f^{(n)}(t_k)$ des signaux d'entrée/sortie par :

$$\begin{cases} u_f^{(n)}(t_k) = F_n(p)u(t_k), & 0 \leq n \leq n_b \\ y_f^{(n)}(t_k) = F_n(p)y(t_k), & 0 \leq n \leq n_a \end{cases}$$

2. À l'aide des k -statistiques (3.54), calculer les cumulants d'ordre quatre et leurs dérivées pour $0 \leq n \leq n_a$ à partir des dérivées des signaux d'entrée/sortie obtenues précédemment ;
3. Construire le vecteur de régression :

$$\hat{\boldsymbol{\Phi}}_f^T(\boldsymbol{\tau}) = \left[-\hat{C}_{uyuu,f}^{(0)}(\boldsymbol{\tau}) \quad \dots \quad -\hat{C}_{uyuu,f}^{(n_a-1)}(\boldsymbol{\tau}) \quad \hat{C}_{uuuu,f}^{(0)}(\boldsymbol{\tau}) \quad \dots \quad \hat{C}_{uuuu,f}^{(n_b)}(\boldsymbol{\tau}) \right]$$

Calculer l'estimée *focls* :

$$\hat{\theta}_{focls}(\tau_2, \tau_3, M) = \left(\frac{1}{M} \sum_{\tau_1=0}^{M-1} \hat{\Phi}_f(\tau) \hat{\Phi}_f^T(\tau) \right)^{-1} \left(\frac{1}{M} \sum_{\tau_1=0}^{M-1} \hat{\Phi}_f(\tau) \hat{C}_{uyuu,f}^{(n_a)}(\tau) \right)$$

L'algorithme focils

1. *Initialisation.* La première estimée $\hat{\theta}_{focils}^0(\tau_2, \tau_3, M)$ est donnée par l'algorithme *focls* :

$$\hat{\theta}_{focils}^0(\tau_2, \tau_3, M) = \hat{\theta}_{focls}(\tau_2, \tau_3, M)$$

2. $(i + 1)^{ième}$ étape. Comme une estimée $\hat{\theta}_{focils}^i(\tau_2, \tau_3, M)$ est disponible, procéder aux mêmes étapes que dans l'algorithme *focls*, mais en utilisant le banc de filtres suivant :

$$F_n(p, \hat{\theta}_{focils}^i(\tau_2, \tau_3, M)) = \frac{p^n}{A(p, \hat{\theta}_{focils}^i(\tau_2, \tau_3, M))}, \quad 0 \leq n \leq n_a$$

pour obtenir $\hat{\theta}_{focils}^{i+1}(\tau_2, \tau_3, M)$;

3. Itérer jusqu'à ce que la différence entre deux estimées successives satisfasse le critère d'arrêt choisi.

3.4 Simulations numériques

Le système suivant est considéré tout au long des simulations numériques [CVJ99] :

$$G_o(p) = \frac{p - 1}{p^2 + 2p + 1} \quad (3.55)$$

Les performances des différents algorithmes sont étudiées à l'aide de simulations de Monte Carlo (de $n_{mc} = 100$ réalisations). Une mesure d'ensemble de la qualité des estimateurs est donnée par l'erreur quadratique moyenne relative sur les paramètres :

$$EQMR = \sqrt{\frac{1}{n_{mc}} \sum_{j=1}^{n_{mc}} \frac{\|\hat{\theta}_j - \theta_o\|^2}{\|\theta_o\|^2}} \quad (3.56)$$

où $\hat{\theta}_j$ représente l'estimée obtenue à la $j^{ième}$ réalisation de la simulation de Monte Carlo. Par ailleurs, le même critère d'arrêt est adopté pour chacun des algorithmes itératifs :

$$\frac{\|\hat{\theta}^{i+1} - \hat{\theta}^i\|}{\|\hat{\theta}^i\|} < 10^{-3} \quad \text{ou} \quad i > 50 \quad (3.57)$$

Ainsi, si l'algorithme n'a pas « convergé » avant 50 itérations, il est automatiquement stoppé. Les variances des bruits perturbant l'entrée et la sortie du système sont ajustées de manière à obtenir un certain rapport signal à bruit (RSB), défini par :

$$RSBu = 10 \log_{10} \left(\frac{P_{u_o}}{P_{\hat{u}}} \right) \quad RSB_y = 10 \log_{10} \left(\frac{P_{y_o}}{P_{\hat{y}}} \right) \quad (3.58)$$

où P_x représente la puissance moyenne du signal x . Les signaux d'entrée/sortie non-bruités sont échantillonnés avec une période d'échantillonnage $T_e = 50ms$.

3.4.1 Estimateurs fondés sur les cumulants d'ordre trois

Cette section est divisée en trois parties, dans lesquelles sont étudiés certains aspects des algorithmes `tocls` et `tocils` :

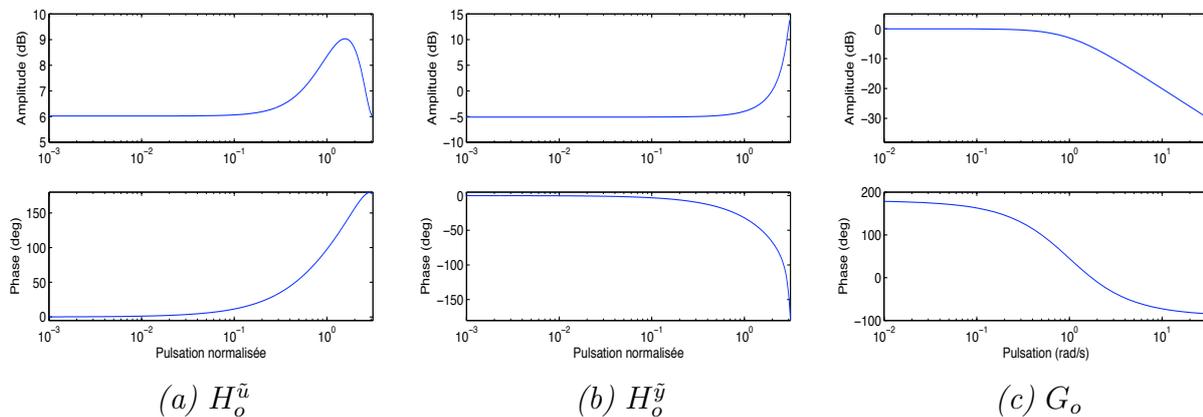
1. les performances des algorithmes proposés sont analysées et comparées à celles des méthodes `srvc` et `ivsvf`⁵ [You70, YJ80, YGG08], toutes deux disponibles dans la boîte à outils logicielle `CONTSID` ;
2. l'influence de l'horizon de calcul M sur la qualité de l'estimation est ensuite étudiée ;
3. dans la dernière partie, l'influence de λ , la pulsation de coupure du filtre des variables d'état, est analysée.

Le signal d'entrée non-bruité et les bruits en entrée/sortie du système sont définis par :

$$\begin{cases} \check{u}_o(t_k) = H_o^{u_o}(q)e_o^{u_o}(t_k) = (1 - 0.2q^{-1} + 0.3q^{-2}) e_o^{u_o}(t_k) \\ \check{\tilde{u}}(t_k) = H_o^{\tilde{u}}(q)e_o^{\tilde{u}}(t_k) = (1 + 2q^{-1} - q^{-2}) e_o^{\tilde{u}}(t_k) \\ \check{\tilde{y}}(t_k) = H_o^{\tilde{y}}(q)\check{\tilde{u}}(t_k) = \frac{\alpha}{1 + 0.8q^{-1}}\check{\tilde{u}}(t_k) \end{cases} \quad (3.59)$$

où $e_o^{u_o}$ est un bruit blanc suivant une loi du χ^2 avec deux degrés de liberté⁶, $e_o^{\tilde{u}}$ est un bruit blanc suivant une distribution uniforme et α est une constante. La variance de $e_o^{\tilde{u}}(t_k)$ et la constante α sont alors ajustées de manière à obtenir un RSB égal à 5 dB en entrée et en sortie du système. Les bruits sont donc colorés et mutuellement corrélés. Une partie des signaux d'entrée/sortie bruités et non-bruités est représentée à la [figure 3.3](#).

FIG. 3.2 – Diagrammes de Bode de $H_o^{\tilde{u}}$, $H_o^{\tilde{y}}$ et G_o



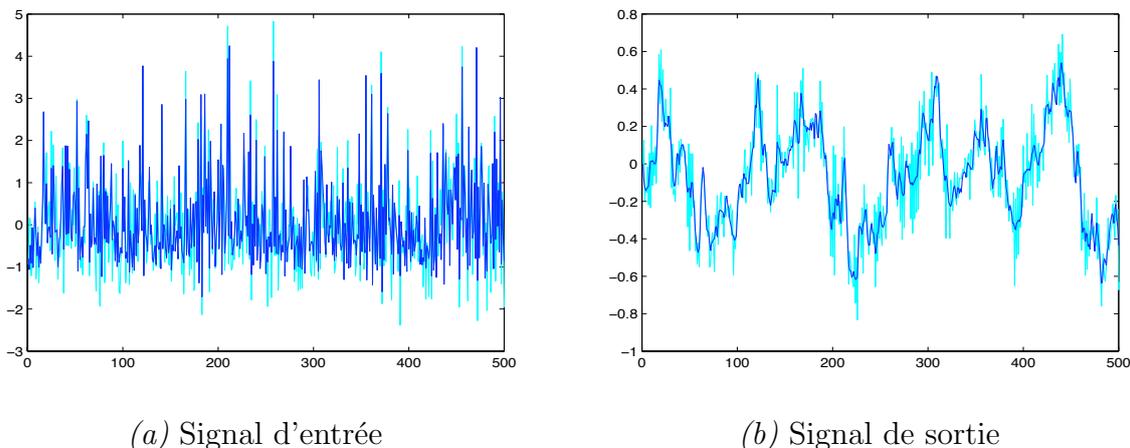
3.4.1.1 Analyse des performances

On se focalise dans un premier temps sur les performances des algorithmes `tocls`, `tocils` en comparaison avec celles de `ivsvf` et `srvc`. Bien que ces deux méthodes ne soient pas

⁵Les acronymes signifient *Simplified Refined Instrumental Variable for Continuous-time models* et *Instrumental Variable-based State Variable Filter*.

⁶La loi du χ^2 est choisie uniquement pour obtenir un signal d'entrée ayant une distribution asymétrique.

FIG. 3.3 – Partie des signaux d’entrée/sortie bruités et non-bruités



spécifiquement développées pour l’identification de modèle EIV, ils font preuve d’une certaine robustesse vis-à-vis du bruit sur le signal d’entrée. Ce phénomène peut être expliqué par l’étape de filtrage qui leur est inhérente, puisque le filtrage des données a un effet régularisant [MBR05].

Les méthodes fondées sur les SOS sont connues pour nécessiter une grande quantité de données pour atteindre une bonne précision [Men91]. Pour juger de l’influence de la longueur du jeu de données, les algorithmes ont été appliqués pour $N = 1000$ et $N = 5000$.

Le tableau 3.1 contient la moyenne et l’écart-type des estimées, l’EQMR, ainsi que λ , M et $\#it$, représentant respectivement la pulsation de coupure du filtre des variables d’état, l’horizon de calcul des cumulants intervenant dans les méthodes `tocls`, `tocils` et le nombre moyen d’itérations. Conformément à la théorie, les méthodes `ivsvf` et `srvc` fournissent des résultats biaisés. Lorsque le nombre de données augmente, le biais des deux méthodes ne décroît pas significativement (voire augmente), ce qui confirme qu’elles ne fournissent pas d’estimées convergentes dans un contexte EIV, par opposition aux algorithmes fondés sur les SOS.

Le biais sur les estimées obtenues par `tocls` et `tocils` est en effet réduit, tandis que l’écart-type est quant à lui supérieur : les estimateurs fondés sur les SOS sont connus pour fournir des résultats plus dispersés que ceux obtenus à l’aide de méthodes utilisant les statistiques de second ordre [DG94]. Cela se voit confirmé ici : la variance est environ doublée par rapport aux résultats des méthodes `ivsvf` ou `srvc`, mais reste toutefois très faible.

La méthode `srvc` fournit des résultats acceptables lorsque $N = 1000$, proches de ceux des algorithmes `tocls` et `tocils` en terme d’EQMR en raison d’une variance plus faible. Toutefois, lorsque $N = 5000$, ces deux derniers algorithmes donnent de meilleurs résultats. Enfin, si l’on compare ces deux algorithmes, on constate que `tocils` fournit de meilleures estimées que `tocls` en terme de biais (indiquant par conséquent que l’erreur minimisée par `tocils` est « plus blanche » que celle minimisée par `tocls`), mais que leur variance est légèrement plus grande.

Les diagrammes de Bode du vrai système et des modèles estimés (voir la [figure 3.4](#)) corroborent cette analyse.

Enfin, on peut vérifier que l'algorithme `tocils` converge rapidement, le nombre moyen d'itérations étant inférieur à 4.

TAB. 3.1 – Comparaison des méthodes `ivsvf`, `srivc`, `tocls` et `tocils` (bruits colorés, RSB \simeq 5dB)

	λ	M	N	$a_1 = 2$	$a_0 = 1$	$b_1 = 1$	$b_0 = -1$	EQMR	#it.
<code>ivsvf</code>	3	–	1000	1.884 ± 0.100	0.907 ± 0.088	0.696 ± 0.024	-0.973 ± 0.061	14.2%	–
			5000	1.842 ± 0.055	0.941 ± 0.040	0.656 ± 0.013	-0.971 ± 0.025	13.8%	–
<code>srivc</code>	–	–	1000	1.807 ± 0.123	0.973 ± 0.075	0.773 ± 0.033	-0.785 ± 0.065	13.1%	7.1
			5000	1.900 ± 0.072	0.943 ± 0.034	0.765 ± 0.018	-0.821 ± 0.028	11.6%	5.2
<code>tocls</code>	3	50	1000	1.861 ± 0.191	0.989 ± 0.124	0.942 ± 0.055	-0.969 ± 0.113	10.7%	–
			5000	1.939 ± 0.087	0.992 ± 0.064	0.982 ± 0.030	-0.987 ± 0.042	4.8%	–
<code>tocils</code>	–	50	1000	1.932 ± 0.219	0.977 ± 0.139	0.954 ± 0.068	-0.980 ± 0.147	11.2%	3.5
			5000	1.992 ± 0.104	0.992 ± 0.065	0.994 ± 0.037	-0.999 ± 0.056	4.9%	2.9

3.4.1.2 Influence de l'horizon de calcul M

Afin d'étudier l'influence de l'hyper-paramètre M sur la qualité des estimées, les algorithmes proposés sont appliqués avec différentes valeurs de M , le nombre d'échantillons de données disponibles étant $N = 5000$. La moyenne et l'écart-type sur les paramètres estimés sont représentés en fonction de M à la [figure 3.5](#) pour la méthode `tocls` et à la [figure 3.6](#) pour la méthode `tocils`. À la [figure 3.7](#), l'EQMR est représentée en fonction de M .

On peut constater que lorsque M est très petit ($M < 20$), les estimées obtenues sont mauvaises, puisque les cumulants ne sont alors estimés qu'à partir d'un nombre d'échantillons insuffisant. Cependant, lorsque M s'accroît, les estimées des cumulants deviennent meilleures, conduisant à une estimation précise du vecteur des paramètres. Notons que lorsque M devient très grand ($M > 80$), un léger biais apparaît pour les deux méthodes `tocls` et `tocils`. La [figure 3.7](#) résume ce phénomène : l'EQMR diminue brutalement, se stabilise, puis augmente lentement lorsque M devient grand. Les meilleures valeurs de M dépendent du système et des conditions expérimentales d'une manière restant à déterminer. Les deux méthodes ne sont néanmoins pas très sensibles à cet hyper-paramètre, tant que l'on ne le choisit pas de valeur extrême.

FIG. 3.4 – Diagrammes de Bode du vrai système (‘-’) et des modèles estimés (bruits colorés, $N = 5000$, $\text{RSB} \simeq 5\text{dB}$)

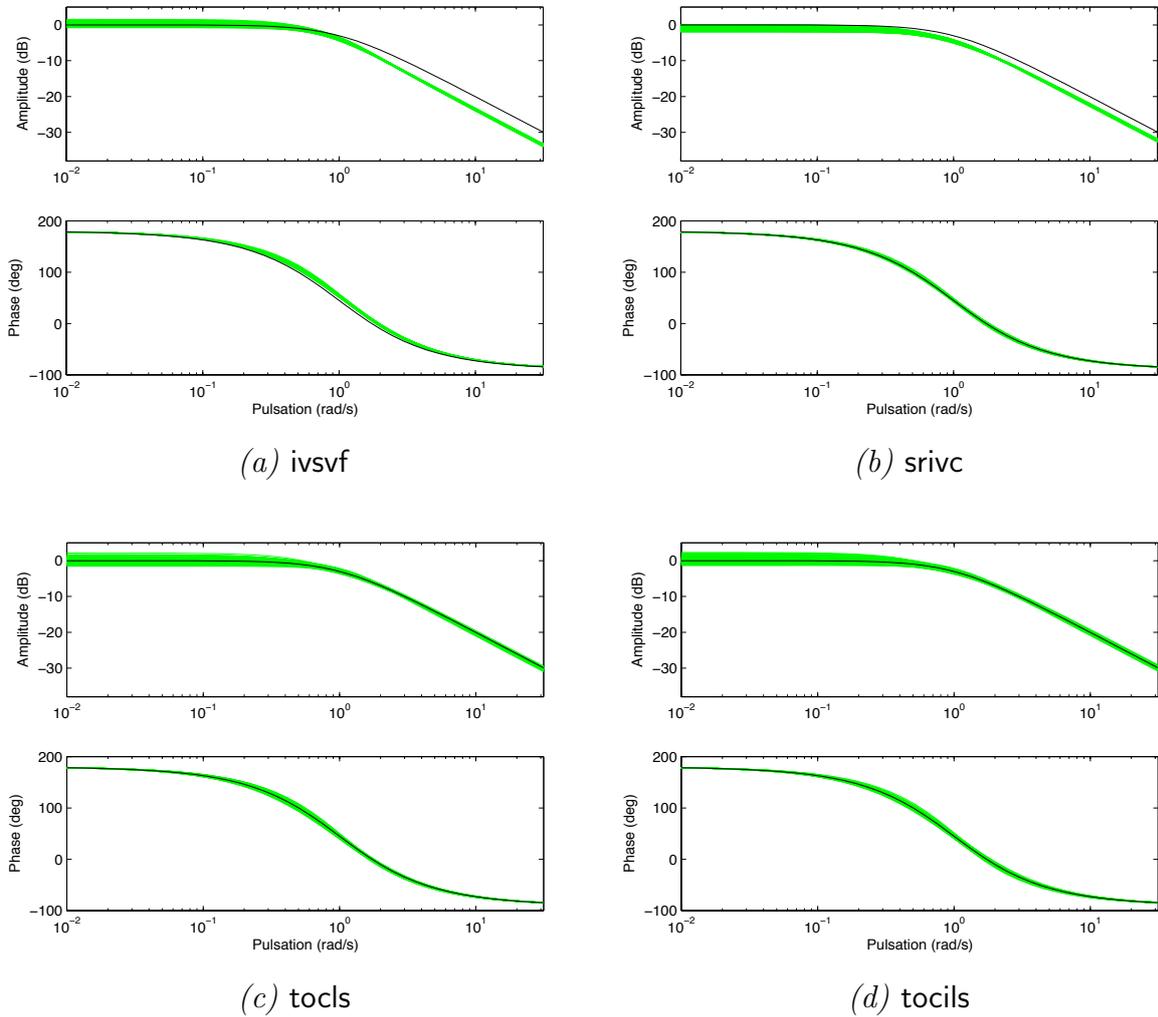


FIG. 3.5 – tocls – Moyenne des estimées et écart-type en fonction de M (bruits colorés, $N = 5000$, $\text{RSB} \simeq 5\text{dB}$)

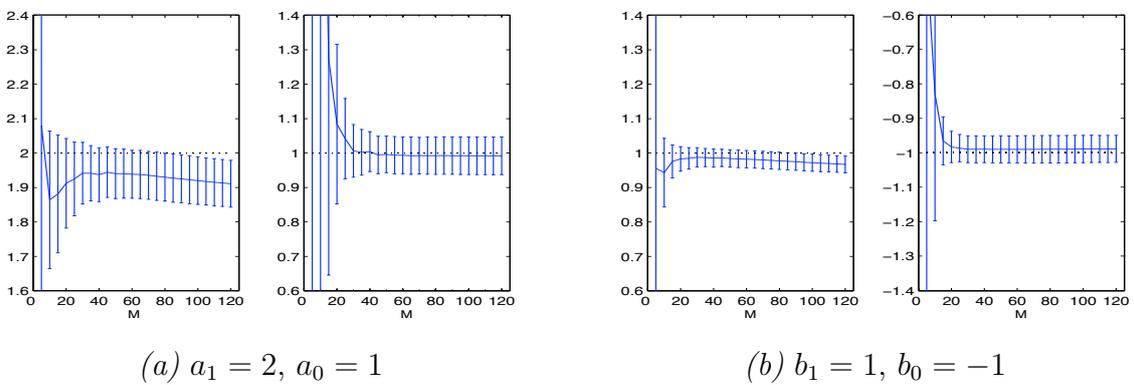


FIG. 3.6 – *tocils* – Moyenne des estimées et écart-type en fonction de M (bruits colorés, $N = 5000$, $\text{RSB} \simeq 5\text{dB}$)

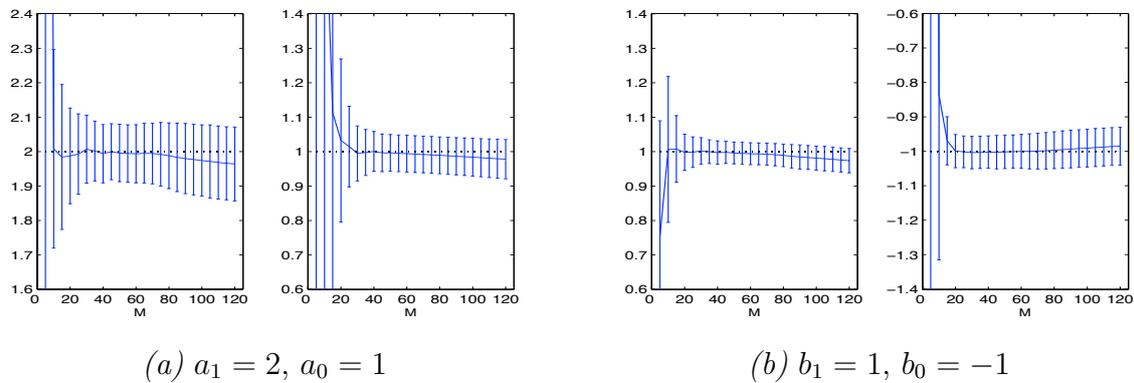
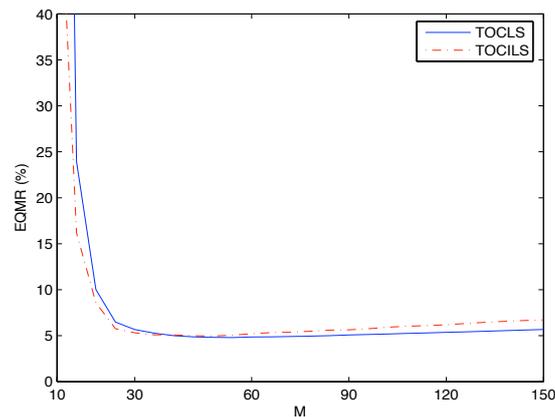


FIG. 3.7 – EQMR en fonction de l'horizon de calcul M



3.4.1.3 Influence de la pulsation de coupure du filtre des variables d'état λ

Dans ce paragraphe la robustesse des algorithmes par rapport à la pulsation de coupure du filtre des variables d'état est étudiée. La valeur de M est fixée à 50 et les algorithmes sont appliqués avec $0.4 \leq \lambda \leq 4.9$, le nombre d'échantillons étant $N = 5000$. Les résultats de la simulation sont donnés dans le [tableau 3.2](#).

On constate que cet hyper-paramètre affecte les performances de la méthode *tocls* : comme l'on pouvait s'y attendre, plus la bande passante du filtre des variables d'état est proche de celle du système, plus l'EQMR est petite. En revanche, puisque λ n'est utilisé que pour l'initialisation de la méthode *tocils*, il n'influence que le nombre d'itérations avant convergence de cet algorithme : plus l'initialisation est proche de la vraie valeur des paramètres, plus la convergence est rapide.

TAB. 3.2 – Influence de la pulsation de coupure λ (bruits colorés, $M = 50$, $N = 5000$)

	λ	$a_1 = 2$	$a_0 = 1$	$b_1 = 1$	$b_0 = -1$	EQMR	#it.
tocls	0.4	2.010 ± 0.123	0.999 ± 0.061	1.004 ± 0.106	-0.989 ± 0.062	6.5%	–
	1.3	1.998 ± 0.098	0.999 ± 0.056	0.995 ± 0.033	-0.990 ± 0.053	4.6%	–
	2.2	1.976 ± 0.093	0.998 ± 0.056	0.989 ± 0.027	-0.986 ± 0.050	4.4%	–
	3.1	1.940 ± 0.096	1.000 ± 0.055	0.981 ± 0.028	-0.978 ± 0.050	5.0%	–
	4.0	1.884 ± 0.102	1.006 ± 0.055	0.969 ± 0.031	-0.964 ± 0.050	6.4%	–
	4.9	1.806 ± 0.110	1.019 ± 0.056	0.951 ± 0.035	-0.945 ± 0.052	8.8%	–
tocils	0.4						2.7
	1.3						2.3
	2.2	2.002 ± 0.102	0.999 ± 0.057	0.997 ± 0.040	-0.990 ± 0.055	4.8%	2.8
	3.1						3.0
	4.0						3.1
	4.9						3.1

3.4.2 Estimateurs fondés sur les cumulants d'ordre quatre

Dans cette partie, le même système (3.55) est considéré. Le signal d'entrée non-bruité \check{u}_o est défini comme une somme de sinus :

$$\check{u}_o(t) = \sin(t) + \sin(1.9t) + \sin(2.1t) + \sin(2.3t) \quad (3.60)$$

Les signaux d'entrée/sortie sont échantillonnés uniformément avec une période d'échantillonnage $T_e = 50\text{ms}$. Une partie des signaux d'entrée/sortie bruités et non-bruités est représentée à la figure 3.8.

3.4.2.1 Analyse des performances

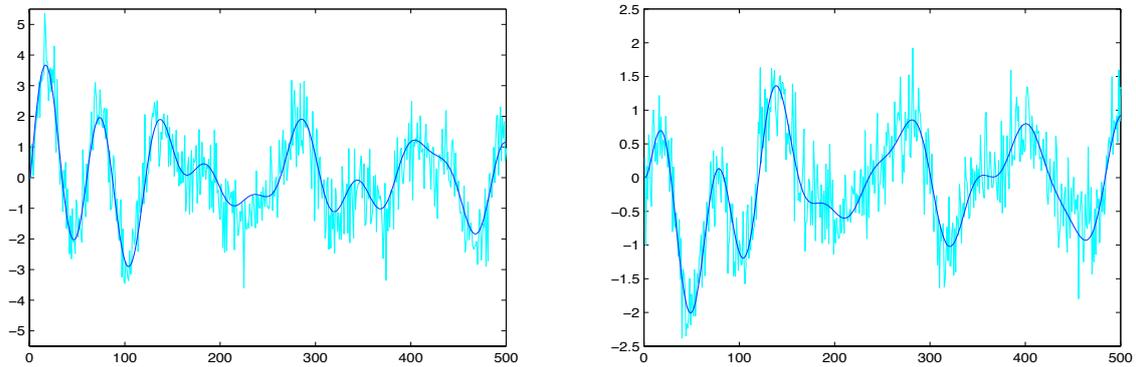
On se focalise dans un premier temps sur les performances des algorithmes focls et focils en les comparant à celles de la méthode eivsvf. Les tableaux 3.3 et 3.4 contiennent la moyenne et l'écart-type des estimées, l'EQMR, ainsi que λ , M et #it, représentant respectivement la pulsation de coupure du filtre des variables d'état, l'hyper-paramètre des méthodes focls, focils et le nombre moyen d'itérations avant convergence. Les diagrammes de Bode du vrai système et des modèles estimés se trouvent à la figure 3.10.

Deux types de bruits additifs en entrée/sortie du système sont considérés.

Bruits blancs en entrée/sortie

Les bruits perturbant les signaux d'entrée/sortie sont définis comme des bruits blancs gaus-

FIG. 3.8 – Partie des signaux d'entrée/sortie bruités et non-bruités



(a) Signal d'entrée

(b) Signal de sortie

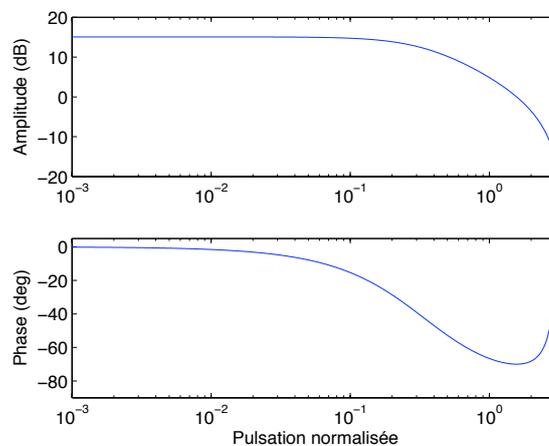
siens, dont les variances sont ajustées de sorte que le RSB soit égal à 5 dB en entrée et en sortie.

Bruit blanc en entrée et coloré en sortie

Les bruits sont définis par :

$$\tilde{u}(t_k) = e_{\tilde{u}}(t_k) \quad \tilde{y}(t_k) = \frac{1 + 0.7q^{-1}}{1 - 0.7q^{-1}} e_{\tilde{y}}(t_k) \quad (3.61)$$

où q^{-1} est l'opérateur retard et $e_{\tilde{u}}(t_k)$, $e_{\tilde{y}}(t_k)$ des bruits blancs gaussiens. À nouveau, les variances des bruits blancs sont ajustées de manière à obtenir un RSB égal à 5 dB en entrée et en sortie. Le diagramme de Bode de $H_{\tilde{y}}$ est représenté à la [figure 3.9](#).

FIG. 3.9 – Diagramme de Bode de $H_{\tilde{y}}$ 

Discussion

TAB. 3.3 – Comparaison des méthodes eivsvf, focls et focils (bruits blancs, $N = 1000$)

	λ	M	$a_1 = 2$	$a_0 = 1$	$b_1 = 1$	$b_0 = -1$	EQMR	#it.
eivsvf	–	–	2.135 ± 0.186	0.998 ± 0.127	1.012 ± 0.088	-1.091 ± 0.098	10.8%	–
focls	3	110	2.005 ± 0.171	1.030 ± 0.157	0.990 ± 0.082	-0.978 ± 0.100	9.5%	–
		160	1.983 ± 0.175	1.026 ± 0.142	0.983 ± 0.080	-0.981 ± 0.096	9.2%	–
focils	–	110	2.025 ± 0.168	1.039 ± 0.131	0.997 ± 0.079	-0.967 ± 0.088	8.9%	3
		160	2.012 ± 0.168	1.031 ± 0.133	0.997 ± 0.080	-0.975 ± 0.095	8.9%	2.9

TAB. 3.4 – Comparaison des méthodes eivsvf, focls et focils (bruits blancs, $N = 5000$)

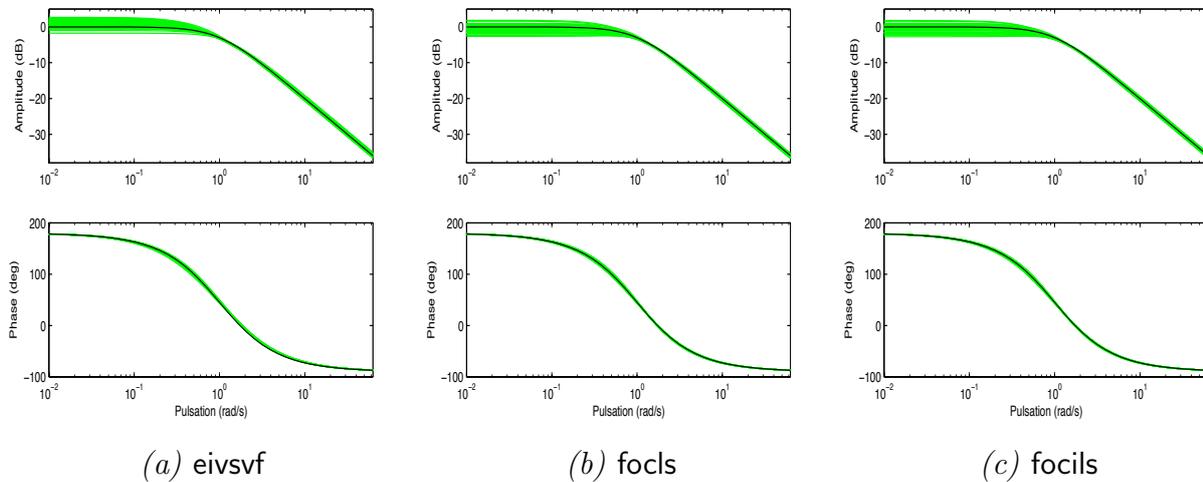
	λ	M	$a_1 = 2$	$a_0 = 1$	$b_1 = 1$	$b_0 = -1$	EQMR	#it.
eivsvf	–	–	2.131 ± 0.077	0.978 ± 0.071	1.021 ± 0.041	-1.098 ± 0.046	7.3%	–
focls	3	110	1.997 ± 0.072	1.035 ± 0.082	0.985 ± 0.036	-0.968 ± 0.049	4.7%	–
		160	1.987 ± 0.071	1.018 ± 0.071	0.985 ± 0.034	-0.977 ± 0.045	4.2%	–
focils	–	110	2.015 ± 0.072	1.037 ± 0.075	0.995 ± 0.035	-0.971 ± 0.046	4.6%	2.6
		160	2.011 ± 0.070	1.019 ± 0.072	0.997 ± 0.035	-0.974 ± 0.047	4.2%	2.3

TAB. 3.5 – Comparaison des méthodes eivsvf, focls et focils (bruit blanc en entrée et coloré en sortie, $N = 5000$)

	λ	M	$a_1 = 2$	$a_0 = 1$	$b_1 = 1$	$b_0 = -1$	EQMR	#it.
eivsvf	–	–	2.129 ± 0.135	1.014 ± 0.129	1.342 ± 0.088	-1.062 ± 0.088	15.3%	–
focls	3	110	1.970 ± 0.138	1.043 ± 0.129	0.972 ± 0.070	-0.961 ± 0.089	8.2%	–
		160	1.955 ± 0.134	1.029 ± 0.111	0.971 ± 0.067	-0.971 ± 0.084	7.6%	–
focils	–	110	1.987 ± 0.137	1.047 ± 0.117	0.980 ± 0.067	-0.953 ± 0.083	7.8%	2.7
		160	1.980 ± 0.134	1.028 ± 0.115	0.983 ± 0.067	-0.971 ± 0.085	7.5%	2.6

Lorsque les bruits sont tous deux blancs, les trois méthodes fournissent de bons résultats. Les estimées obtenues par les algorithmes **focls** et **focils** ont un biais sur les estimées inférieur à celui sur les estimées de la méthode **eivsvf**, tandis que les variances sont comparables. L'EQMR est par conséquent légèrement meilleure pour les méthodes fondées sur les SOS. Cette différence devient plus franche lorsque la longueur du jeu de données augmente ($N = 5000$) : les estimées des cumulants s'améliorent, et par conséquent les estimateurs **focls** et

FIG. 3.10 – Diagrammes de Bode du vrai système (‘-’) et des modèles estimés (bruits blancs $N = 5000$, $M = 160$, $\text{SNR} = 5$ dB)



focls deviennent plus précis.

Lorsque le bruit en sortie est coloré, les performances de la méthode eivsvf se détériorent logiquement. Elle suppose en effet que les bruits en entrée/sortie sont blancs : l’hypothèse $(H_o^{\hat{u}}, H_o^{\hat{y}}) \in \mathcal{H}^*$ n’est donc plus vérifiée. Les algorithmes focls et focils fournissent quant à eux toujours de bons résultats, montrant ainsi leur robustesse face à des bruits gaussiens, qu’ils soient blancs ou colorés.

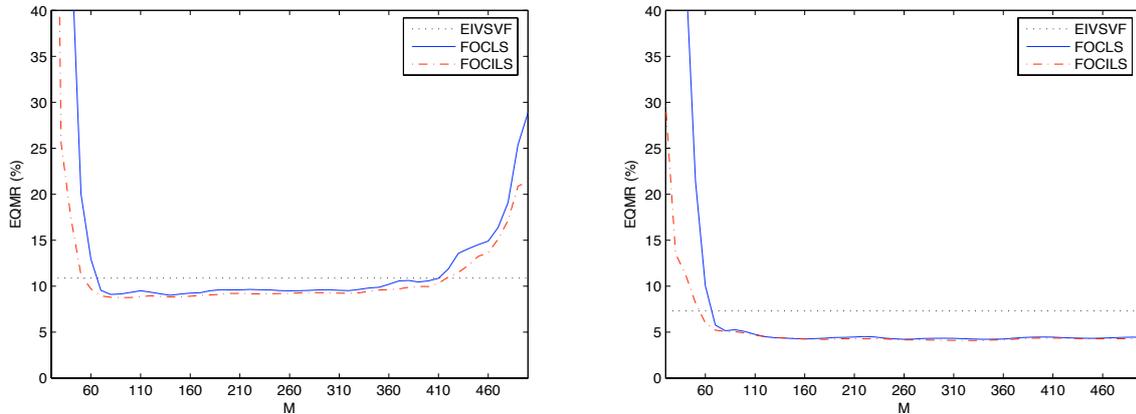
3.4.2.2 Influence de l’horizon de calcul M

Pour étudier l’influence de l’hyper-paramètre M sur la qualité des estimées, les algorithmes focls et focils ont été appliqués pour différentes valeurs de M . L’EQMR sur les paramètres estimés à l’aide des trois méthodes eivsvf, focls et focils est tracée en fonction de M pour $N = 1000$ et pour $N = 5000$ à la [figure 3.11](#).

On constate que les deux algorithmes ne sont pas très sensibles à cet hyper-paramètre, tant qu’il n’est pas choisi trop petit, ou trop proche de la longueur du jeu de données. En effet, l’accroissement brutal de l’EQMR à la [figure 3.11](#)-(a) s’explique par la valeur de M relativement à celle de N : les cumulants sont estimés avec trop peu de données pour être fiables, ce qui a pour conséquence d’engendrer de grandes erreurs sur les estimées du vecteur des paramètres.

3.5 Conclusion

Ce chapitre a traité de l’identification de modèles à temps continu avec des erreurs en les variables. Peu de travaux s’intéressant à ce problème ont été effectués, et ce domaine de

FIG. 3.11 – EQMR en fonction de l'horizon de calcul M (a) Bruits blancs, $N = 1000$ (b) Bruits blancs, $N = 5000$

recherche est encore très largement ouvert.

Après avoir motivé l'intérêt particulier des approches directes pour l'identification de modèles à temps continu, nous sommes dans un premier temps revenus succinctement sur les quelques méthodes existantes, toutes fondées sur l'utilisation de statistiques d'ordre deux. L'étape de pré-traitement de données, utilisée pour éviter d'avoir à calculer les dérivées des signaux d'entrée/sortie bruités, complique bien souvent le problème d'identification, et la plupart des méthodes développées supposent en conséquence blancs les bruits en entrée/sortie du système. Cette hypothèse permet en effet, outre d'éviter quelques cas de non-identifiabilité, de simplifier le développement des algorithmes.

Les approches fondées sur les statistiques d'ordre supérieur (SOS) sont donc intéressantes d'une part pour leur « insensibilité » aux bruits, indépendamment de leur structure, et d'autre part pour s'affranchir des problèmes d'identifiabilité. Si plusieurs méthodes utilisant les SOS ont été développées dans le cadre de l'identification de modèles EIV à temps discret, aucune ne l'a en revanche été pour le cas des modèles à temps continu. Nous avons donc proposé des algorithmes fondés sur l'utilisation de cumulants d'ordre trois et quatre, et avons exposé les spécificités liées à leur mise en œuvre. Ces estimateurs permettent en particulier de s'affranchir des hypothèses structurelles sur les bruits additifs en entrée/sortie du système. Ils peuvent par conséquent être appliqués, sans changement, lorsque ceux-ci sont tous deux colorés (et même mutuellement corrélés), cas de figure qu'aucune des méthodes précédemment développées ne pouvait traiter.

Une étude s'appuyant sur des simulations numériques a permis d'illustrer différents aspects des méthodes proposées, *i.e.* la comparaison de leurs performances avec celles d'une méthode fondée sur les statistiques d'ordre deux, ainsi que l'influence des hyper-paramètres à choisir par l'utilisateur.

Perspectives

Comme nous l'avons déjà précisé, l'identification de modèles EIV à temps continu est encore très ouvert. Les perspectives sont donc nombreuses : par exemple, étudier quelles méthodes à temps discret présentées au [chapitre 2](#) pourraient être transformées pour l'identification directe de modèles EIV à temps continu ; ou encore quelles méthodes d'identification directe de modèles à temps continu « classiques » peuvent être étendues au cas des modèles EIV. Par ailleurs, nous donnons ci-après quelques perspectives ouvertes par le travail présenté dans ce chapitre :

- l'estimation et le choix des lignes de cumulants – et même le choix des cumulants – sont autant de problèmes intéressants. Les études réalisées à ce sujet font généralement intervenir des variables aléatoires n'ayant pas de lien spécifique. Existe-t-il des choix particuliers judicieux, prenant en compte la structure du modèle choisi ?
- il serait intéressant d'étudier le comportement des méthodes fondées sur les statistiques d'ordre supérieur lorsque les hypothèses distributionnelles ne sont pas exactement vérifiées. Par exemple, supposons que la distribution du signal d'entrée non-bruité s'approche d'une distribution gaussienne. Est-il possible d'utiliser les valeurs du critère de gaussiannité pour spécifier à partir de quel moment ces méthodes fournissent des résultats satisfaisants ?
- les méthodes de compensation du biais des estimateurs de la variable instrumentale et des moindres carrés mentionnés à la [section 3.2.1](#) utilisent un pré-traitement fondé sur un filtre linéaire intégral. Est-il possible de développer ces méthodes en utilisant un pré-traitement fondé sur un filtre linéaire avec un choix automatique et optimal des hyper-paramètres ? L'idée serait alors de développer une méthode du type `srvc` pour le modèle EIV.

Chapitre 4

Conclusion et perspectives

Ce mémoire a traité de l'identification de modèles lorsque les signaux d'entrée/sortie sont tous deux perturbés par un bruit additif. La présentation s'est focalisée sur les méthodes d'estimation paramétrique dans le domaine temporel.

Le chapitre d'introduction a tout d'abord permis de motiver l'intérêt porté aux modèles avec des erreurs en les variables et de mettre en évidence les difficultés inhérentes à leur identification. Le second chapitre de la thèse a traité de l'identification de modèles à temps discret. Dans la première partie de ce chapitre, l'attention s'est portée sur les méthodes utilisant les statistiques d'ordre deux; les méthodes d'identification ayant recours aux statistiques d'ordre supérieur ont fait l'objet de la seconde partie de ce chapitre. Enfin, le troisième chapitre a traité de l'identification de modèles à temps continu avec des erreurs en les variables. Nous sommes dans un premier temps revenus succinctement sur les rares méthodes existantes, avant de proposer des algorithmes fondés sur l'utilisation de cumulants d'ordre trois et quatre.

Nous avons vu que les méthodes s'appuyant sur les statistiques d'ordre deux fournissent une estimation convergente du vecteur des paramètres sous des hypothèses structurelles sur les bruits additifs en entrée/sortie (et sur le signal d'entrée non bruité pour les méthodes de l'erreur de prédiction et du maximum de vraisemblance). Ces hypothèses permettent de simplifier le problème, et la plupart des méthodes considèrent le cas le plus simple, à savoir des bruits blancs en entrée/sortie du système. L'estimation de leur variance permet alors d'obtenir des estimées non biaisées du vecteur des paramètres. De nombreuses approches ont été développées pour traiter ce cas. Lorsque l'un des bruits additifs du modèle est coloré, la situation se complique. Il est alors possible d'estimer la fonction d'autocorrélation du bruit coloré (pour quelques délais, suivant l'ordre du modèle); cette opération est cependant délicate, surtout si le nombre de délais pour lesquels la fonction d'autocorrélation doit être estimée devient important. Nous avons également proposé et développé des méthodes fondées sur la technique de la variable instrumentale. Plusieurs méthodes peuvent ainsi prendre en compte le cas où un des bruits additifs est coloré; elles sont pour la plupart des extensions

des méthodes développées pour le cas de bruits blancs en entrée/sortie (comme récemment la méthode de Frisch, par exemple). Le cas de figure où les bruits sont tous deux colorés est en revanche encore relativement ouvert, puisque jusqu'à présent seules les méthodes de l'erreur de prédiction, du maximum de vraisemblance et des moindres carrés séparables permettent théoriquement de le traiter¹. Il faut toutefois noter que la convergence des estimateurs fondés sur les statistiques d'ordre deux est assurée uniquement si les hypothèses structurelles sont vérifiées. Une difficulté supplémentaire est donc de modéliser de manière appropriée les bruits additifs (et, le cas échéant, le signal d'entrée non bruité).

Les méthodes fondées sur les statistiques d'ordre supérieur permettent de s'affranchir de ces problèmes de structure. En effet, puisque les cumulants d'ordre supérieur sont « insensibles » à certaines classes de bruits (ayant une distribution symétrique pour l'ordre trois, gaussienne pour l'ordre quatre), ils offrent la possibilité de développer des estimateurs convergents sans avoir à modéliser les bruits additifs en entrée/sortie du système. Cela implique également qu'à l'aide de méthodes fondées sur les statistiques d'ordre supérieur, il est possible de traiter, à l'aide d'une même méthode, des situations allant du cas le plus simple (bruits blancs en entrée/sortie) au cas général de bruits colorés en entrée et en sortie (et même mutuellement corrélés). Évidemment, ces avantages ne viennent pas seuls, et il faut poser des hypothèses distributionnelles sur les bruits et le signal d'entrée non bruité. Si les hypothèses sur les bruits sont assez douces (distribution symétrique ou gaussienne), celles sur le signal d'entrée non bruité sont en revanche plutôt restrictives (distribution asymétrique ou non gaussienne). Il convient par ailleurs de noter que, de manière générale, pour obtenir une précision comparable, l'estimation des cumulants d'ordre supérieur nécessite un plus grand nombre de données que l'estimation des corrélations. Par l'intermédiaire de simulations numériques, nous avons toutefois pu constater qu'il n'est pas nécessaire que la longueur du jeu de données soit démesurément grande.

Les estimateurs fondés sur les statistiques d'ordre deux et ceux fondés sur les statistiques d'ordre supérieur ont ainsi leurs avantages et inconvénients respectifs, et il serait insensé de privilégier les uns ou les autres. Lors du choix entre les différentes méthodes, les conditions expérimentales sont alors déterminantes : longueur du jeu de données, vraisemblance des hypothèses structurelles/distributionnelles, *etc.*

Les principales contributions du travail effectué lors de la thèse sont les suivantes :

- une présentation unifiée de différentes méthodes de type compensation du biais de l'estimateur des moindres carrés ;
- la proposition d'estimateurs de type variable instrumentale pour l'identification de modèles EIV à temps discret ;
- la proposition d'un estimateur des moindres carrés fondé sur les cumulants d'ordre supérieur,

¹Précisons encore une fois que nous nous sommes focalisés sur les méthodes d'estimation paramétrique dans le domaine temporel. Ce bilan est donc restreint à ce domaine.

- pour le modèle EIV à temps discret ;
- le calcul de la matrice de covariance asymptotique d'un estimateur fondé sur les cumulants d'ordre supérieur, pour le modèle EIV à temps discret ;
 - la proposition et le développement de plusieurs estimateurs originaux de modèles EIV à temps continu, fondés sur les cumulants d'ordre supérieur ;
 - une analyse comparative des différents algorithmes présentés à l'aide de simulations de Monte Carlo ;
 - l'ajout des algorithmes développés pour l'identification de modèles EIV à temps continu au sein de la boîte à outils logicielle CONTSID.

Perspectives

Le travail présenté dans ce mémoire a permis d'ouvrir de nombreuses perspectives. Concernant l'identification de modèle EIV à partir des statistiques d'ordre deux, l'extension à des cas plus généraux des méthodes ayant uniquement été développées pour le cas de bruits blancs en entrée et en sortie du système est une première perspective, à laquelle peuvent s'ajouter quelques autres, plus spécifiques aux différentes méthodes présentées, proposées en fin des chapitres 2 et 3. L'identification de modèles EIV à partir de statistiques d'ordre supérieur soulève également nombre de questions ; celles-ci semblent toutefois moins liées la procédure d'identification proprement dite qu'aux aspects pratiques de l'utilisation de cumulants.

De manière générale, les travaux suivants nous semblent être des perspectives directes du travail effectué lors de la thèse :

- il serait intéressant d'évaluer la robustesse des approches présentées vis-à-vis de leurs hypothèses structurelles et distributionnelles ;
- de nombreuses méthodes présentées utilisent des hyper-paramètres devant être choisis par l'utilisateur. Est-il possible d'en faire des méthodes *non supervisées*, c'est-à-dire sans hyper-paramètres ou avec un choix un choix automatique et optimal de ceux-ci ?
- il est également possible d'utiliser les statistiques d'ordre supérieur pour estimer l'ordre du système (voir [CC94, Tug95, TL99] par exemple). Il serait donc intéressant de développer des algorithmes qui estiment conjointement l'ordre et les paramètres du modèle ;
- dans ce mémoire, nous nous sommes exclusivement intéressés aux modèles *linéaires* avec des erreurs en les variables. Or, les systèmes non linéaires engendrent en général des signaux non gaussiens, et l'utilisation des statistiques d'ordre supérieur pour une description complète de ces systèmes est donc nécessaire. L'identification de modèles non linéaires avec des erreurs en les variables est donc une perspective directe du travail présenté dans ce mémoire.

Enfin, notons que les études menées et les connaissances théoriques acquises durant ces trois années ne m'ont pas laissé le loisir de m'intéresser plus avant aux possibles applications. L'utilisation des méthodes étudiées et des algorithmes développés sur des jeux de données

réelles est néanmoins un objectif de tout premier plan.

Annexe A

Statistiques d'ordre supérieur : définitions et propriétés

Ce qui suit provient en grande partie de [LAC97].

A.1 Cas des variables aléatoires réelles scalaires

Soit x une variable aléatoire à valeurs réelles. $F_x(u)$, la fonction de répartition de x , est la probabilité d'apparition de la variable aléatoire dans le segment $] -\infty, u]$. Si x admet une densité de probabilité $p_x(u)$, alors $dF_x(u) = p_x(u)du$. La densité de probabilité $p_x(u)$ est positive et a pour somme l'unité.

Les moments généralisés de x sont définis pour toute application réelle g par :

$$E\{g(x)\} = \int_{-\infty}^{+\infty} g(u)p_x(u)du \quad (\text{A.1})$$

On utilise souvent des fonctions polynomiales pour g , conduisant aux moments « classiques » de différents ordres, tels que la moyenne ou la variance. En utilisant des fonctions exponentielles, on associe aux variables aléatoires des fonctions caractéristiques.

DÉFINITION A.1.1 (PREMIÈRE FONCTION CARACTÉRISTIQUE)

La première fonction caractéristique de x est définie par :

$$\Phi_x(v) = E\{e^{jvx}\} \quad (\text{A.2})$$

où j désigne la racine carrée de -1 . ■

Lorsque la variable aléatoire x admet une densité de probabilité $p_x(u)$, la première fonction caractéristique Φ_x est sa transformée de Fourier :

$$\Phi_x(v) = \int_{-\infty}^{+\infty} e^{jvu}p_x(u)du \quad (\text{A.3})$$

Dans ce cas, on retrouve la densité de probabilité à partir de la première fonction caractéristique par la transformation de Fourier inverse :

$$p_x(u) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-jvu} \Phi_x(v) dv \quad (\text{A.4})$$

La fonction caractéristique est continue et vaut 1 à l'origine. Elle est donc non nulle dans un voisinage de l'origine, sur lequel on pourra définir son logarithme népérien.

DÉFINITION A.1.2 (SECONDE FONCTION CARACTÉRISTIQUE)

La seconde fonction caractéristique de x est définie comme le logarithme népérien de la première fonction caractéristique de x :

$$\Psi_x(v) = \ln(\Phi_x(v)) \quad (\text{A.5})$$

■

Lorsqu'ils existent, on note $\mu_r(x)$ les moments d'ordre r de x :

$$\mu_r(x) = \text{E}\{x^r\} \quad (\text{A.6})$$

et $\mu'_r(x)$ les moments centrés :

$$\mu'_r(x) = \text{E}\{(x - \mu_1(x))^r\} \quad (\text{A.7})$$

Les fonctions caractéristiques décrivent complètement la variable aléatoire à laquelle elles sont associées. Si, dans l'équation (A.2), on développe e^{jvx} au voisinage de l'origine, et que l'on identifie ses coefficients avec ceux du développement de Taylor de la première fonction caractéristique, on obtient les moments :

$$\mu_r(x) = (-j)^r \left. \frac{d^r \Phi_x(v)}{dv^r} \right|_{v=0} = \text{E}\{x^r\} \quad (\text{A.8})$$

De manière similaire les dérivées de la seconde fonction caractéristique, prises à l'origine, définissent les *cumulants*.

DÉFINITION A.1.3 (CUMULANTS)

Les cumulants de la variable aléatoire x sont donnés par les coefficients du développement de Taylor à l'origine de la seconde fonction caractéristique. Par exemple à l'ordre r :

$$\text{Cum}\{x, x, \dots, x\} = (-j)^r \left. \frac{d^r \Psi_x(v)}{dv^r} \right|_{v=0} \quad (\text{A.9})$$

■

Les cumulants d'ordre r peuvent être calculés à partir des moments d'ordre inférieur ou égal à r . Pour les ordres 1 à 4, ces relations sont :

$$\text{Cum}\{x\} = \mu_1(x) \quad (\text{A.10a})$$

$$\text{Cum}\{x, x\} = \mu_2'(x) = \mu_2(x) - \mu_1(x)^2 \quad (\text{A.10b})$$

$$\text{Cum}\{x, x, x\} = \mu_3'(x) = \mu_3(x) - 3\mu_2(x)\mu_1(x) + 2\mu_1(x)^3 \quad (\text{A.10c})$$

$$\text{Cum}\{x, x, x, x\} = \mu_4(x) - 4\mu_3(x)\mu_1(x) - 3\mu_2(x)^2 + 12\mu_2(x)\mu_1(x)^2 - 6\mu_1(x)^4 \quad (\text{A.10d})$$

Dans le cas de variables aléatoires centrées, les expressions se simplifient en :

$$\text{Cum}\{x\} = 0 \quad (\text{A.11a})$$

$$\text{Cum}\{x, x\} = \mu_2(x) \quad (\text{A.11b})$$

$$\text{Cum}\{x, x, x\} = \mu_3(x) \quad (\text{A.11c})$$

$$\text{Cum}\{x, x, x, x\} = \mu_4(x) - 3\mu_2(x)^2 \quad (\text{A.11d})$$

Un cas particulier important : lorsque la variable x est gaussienne, sa seconde fonction caractéristique est :

$$\Psi_x(v) = j\mu_1(x)v - \frac{1}{2}\mu_2(x)v^2 \quad (\text{A.12})$$

Ses cumulants d'ordre supérieur à 2 sont par conséquent tous nuls. Les variables gaussiennes sont donc entièrement décrites par leurs propriétés au second ordre. Ceci explique pourquoi on se limite souvent au second ordre. Le théorème de la limite centrale peut en effet laisser penser que la plupart des signaux ont tendance à être gaussiens.

En centrant la variable aléatoire x et en la normalisant on obtient la variable standardisée :

$$\tilde{x} = \frac{x - \mu_1(x)}{\sqrt{\text{Cum}\{x, x\}}} \quad (\text{A.13})$$

de laquelle on déduit les cumulants standardisés, qui sont les cumulants de la variable aléatoire standardisée. Le cumulants standardisé d'ordre 3 est dénommé l'*asymétrie*. L'asymétrie est nulle si la densité de probabilité de x possède un axe de symétrie. Le cumulants standardisé d'ordre 4 est le facteur d'aplatissement ou *kurtosis*. Le signe du kurtosis caractérise le comportement à l'infini des densités de probabilité unimodales. Les densités de probabilité unimodales possédant un kurtosis négatif sont dites sous-gaussiennes car elles tendent vers 0 à l'infini plus rapidement que la loi gaussienne. Les densités de probabilité unimodales à kurtosis positif, dites sur-gaussiennes, tendent vers 0 à l'infini moins vite que la loi gaussienne.

A.2 Cas des vecteurs aléatoires

Soit $\mathbf{x}^T = (x_1, x_2, \dots, x_n)$ un vecteur aléatoire. De la même manière que dans le cas scalaire, on définit la première fonction caractéristique conjointe de ces n variables aléatoires par la

relation :

$$\Phi_{\mathbf{x}}(\mathbf{v}) = \mathbb{E}\left\{e^{j\mathbf{v}^T\mathbf{x}}\right\} \quad (\text{A.14})$$

où $\mathbf{v}^T = (v_1, v_2, \dots, v_n)$. Si les composantes du vecteur aléatoire \mathbf{x} admettent une densité de probabilité conjointe $p_{\mathbf{x}}(u)$, alors la fonction caractéristique de \mathbf{x} est donnée par la transformée de Fourier de cette densité :

$$\Phi_{\mathbf{x}}(\mathbf{v}) = \int_{\mathbb{R}^n} e^{j\mathbf{v}^T\mathbf{u}} p_{\mathbf{x}}(\mathbf{u}) d\mathbf{u} \quad (\text{A.15})$$

Comme précédemment, la seconde fonction caractéristique est $\Psi_{\mathbf{x}}(\mathbf{v}) = \log(\Phi_{\mathbf{x}}(\mathbf{v}))$. Les fonctions caractéristiques peuvent servir à générer les moments et les cumulants. Il est possible d'écrire des relations générales liant moments et cumulants de manière compacte grâce à la formule de Leonov et Shiriyayev.

DÉFINITION A.2.1 (FORMULE DE LEONOV ET SHIRYAYEV)

La relation générale liant les moments et les cumulants du vecteur aléatoire \mathbf{x} est (donnée ici à l'ordre r) :

$$\text{Cum}\{x_1, \dots, x_r\} = \sum (-1)^{k-1} (k-1)! \mathbb{E}\left\{\prod_{i \in v_1} x_i\right\} \mathbb{E}\left\{\prod_{i \in v_2} x_i\right\} \dots \mathbb{E}\left\{\prod_{i \in v_p} x_i\right\} \quad (\text{A.16})$$

où la sommation s'étend sur tous les ensembles $\{v_1, v_2, \dots, v_p \mid 1 \leq p \leq r\}$ formant une partition de $\{1, 2, \dots, r\}$, et où k est le nombre d'éléments composant cette partition. ■

Dans le cas d'un vecteur aléatoire centré

$$\text{Cum}\{x_1\} = \mathbb{E}\{x_1\} \quad (\text{A.17a})$$

$$\text{Cum}\{x_1, x_2\} = \mathbb{E}\{x_1 x_2\} \quad (\text{A.17b})$$

$$\text{Cum}\{x_1, x_2, x_3\} = \mathbb{E}\{x_1 x_2 x_3\} \quad (\text{A.17c})$$

$$\begin{aligned} \text{Cum}\{x_1, x_2, x_3, x_4\} = & \mathbb{E}\{x_1 x_2 x_3 x_4\} - \mathbb{E}\{x_1 x_2\} \mathbb{E}\{x_3 x_4\} \\ & - \mathbb{E}\{x_1 x_3\} \mathbb{E}\{x_2 x_4\} - \mathbb{E}\{x_1 x_4\} \mathbb{E}\{x_2 x_3\} \end{aligned} \quad (\text{A.17d})$$

Les cumulants de variables multidimensionnelles apparaîtront naturellement lorsque l'on s'intéressera aux multicorrélations des signaux aléatoires.

A.3 Propriétés des cumulants

On se restreint ici aux seules propriétés qui seront utilisées dans la suite. Il en existe un grand nombre que l'on peut trouver dans [Bri81, Bri01] par exemple.

PROPRIÉTÉ A.3.1 (MULTILINÉARITÉ)

Soient \mathbf{x} et \mathbf{y} deux vecteurs aléatoires liés par la relation linéaire $\mathbf{y} = \mathbf{A}\mathbf{x}$, où \mathbf{A} est une matrice quelconque. Alors les cumulants de \mathbf{y} sont des fonctions linéaires de chacune des composantes A_{ij} . Par exemple, on a

$$\text{Cum}\{y_i, y_j, y_k\} = \sum_{a,b,c} A_{ia}A_{jb}A_{kc} \text{Cum}\{x_a, x_b, x_c\} \quad (\text{A.18})$$

■

En particulier, quelles que soient les variables aléatoires x, y, z_1, \dots, z_n , réelles ou complexes, scalaires ou multidimensionnelles, indépendantes ou non, on a :

$$\text{Cum}\{x + y, z_1, \dots, z_n\} = \text{Cum}\{x, z_1, \dots, z_n\} + \text{Cum}\{y, z_1, \dots, z_n\} \quad (\text{A.19})$$

$$\text{Cum}\{\lambda x, z_1, \dots, z_n\} = \lambda \text{Cum}\{x, z_1, \dots, z_n\} \quad (\text{A.20})$$

PROPRIÉTÉ A.3.2 (INVARIANCE PAR TRANSLATION)

Les cumulants sont invariants par translation déterministe : si $\mathbf{y} = \mathbf{x} + \mathbf{t}$ où \mathbf{t} est déterministe, alors les cumulants d'ordre supérieur ou égal à 2 de \mathbf{y} sont identiques à ceux de \mathbf{x} . En effet, la translation provoque un déphasage de la première fonction caractéristique, qui implique que les secondes fonctions caractéristiques de \mathbf{x} et \mathbf{y} sont liées par $\Psi_{\mathbf{y}}(\mathbf{v}) = j\text{Re}[\mathbf{t}^T \mathbf{v}] + \Psi_{\mathbf{x}}(\mathbf{v})$. Cette relation montre que la translation n'affecte que le cumulants d'ordre 1, c'est-à-dire la moyenne. ■

La propriété d'invariance par translation n'est évidemment pas vérifiée par les moments.

A.3.1 Variables aléatoires indépendantes

L'indépendance statistique de variables aléatoires implique la factorisation des moments (via la factorisation des premières fonctions caractéristiques).

Soient \mathbf{x} et \mathbf{y} deux vecteurs aléatoires indépendants de dimensions respectives n et p , et soit $\mathbf{z}^T = [\mathbf{x}^T \ \mathbf{y}^T]$ la concaténation des deux vecteurs. Alors la première fonction caractéristique de \mathbf{z} est le produit des premières fonctions caractéristiques de \mathbf{x} et \mathbf{y} . Par conséquent, la seconde fonction caractéristique de \mathbf{z} s'écrit $\Psi_{\mathbf{z}}(\mathbf{u}, \mathbf{v}) = \Psi_{\mathbf{x}}(\mathbf{u}) + \Psi_{\mathbf{y}}(\mathbf{v})$. Les cumulants croisés entre \mathbf{x} et \mathbf{y} sont donc nuls. En effet on a de manière générale :

$$\text{Cum}\{x_{i_1}, \dots, x_{i_n}, y_{j_1}, \dots, y_{j_p}\} = (-j)^{(n+p)} \frac{\partial^{(n+p)} \Psi_{\mathbf{z}}(\mathbf{u}, \mathbf{v})}{\partial u_{i_1} \dots \partial u_{i_n} \partial v_{j_1} \dots \partial v_{j_p}} \Big|_{(\mathbf{u}, \mathbf{v})=\mathbf{0}} \quad (\text{A.21})$$

Puisque $\Psi_{\mathbf{z}}(\mathbf{u}, \mathbf{v}) = \Psi_{\mathbf{x}}(\mathbf{u}) + \Psi_{\mathbf{y}}(\mathbf{v})$, le cumulants précédent est nul dès que l'un des i_k et l'un des j_k sont non nuls simultanément. En d'autres termes, les cumulants croisés de \mathbf{x} et \mathbf{y} sont nuls.

Une propriété en découlant immédiatement est que si x et y sont des variables aléatoires indépendantes, alors quelle que soit z :

$$\text{Cum}\{x, y, z\} = 0 \quad (\text{A.22})$$

On en déduit une propriété fondamentale : la propriété d'additivité.

PROPRIÉTÉ A.3.3 (ADDITIVITÉ)

Le cumulants de la somme de deux vecteurs aléatoires indépendants \mathbf{x} et \mathbf{y} est égal à la somme de leurs cumulants :

$$\text{Cum}\{x_1 + y_1, \dots, x_n + y_n\} = \text{Cum}\{x_1, \dots, x_n\} + \text{Cum}\{y_1, \dots, y_n\} \quad (\text{A.23})$$

■

Cette propriété, bien connue à l'ordre 2 (variance), est évidemment fausse pour les moments d'ordre supérieurs.

Enfin, énonçons des propriétés moins générales, dans le sens où elles s'appliquent à des variables aléatoires ayant une distribution particulière :

PROPRIÉTÉ A.3.4 (V.A. DE DISTRIBUTION SYMÉTRIQUE OU GAUSSIENNE)

Les cumulants d'ordre trois d'une variable aléatoire dont la distribution possède un axe de symétrie sont tous nuls.

Les cumulants d'ordre quatre d'une variable aléatoire ayant une distribution gaussienne sont tous nuls.

■

A.4 Signaux aléatoires

Les propriétés statistiques d'ordre supérieur des signaux aléatoires stationnaires peuvent être décrites, comme le sont les propriétés d'ordre 2, dans le domaine du temps ou dans le domaine des fréquences. L'étude dans le domaine du temps conduit aux multicorrélations, l'étude dans le domaine des fréquences aux multispectres.

Les multicorrélations sont définies à partir des cumulants d'ordre supérieur à 2, de la même manière que les corrélations sont définies à partir des cumulants d'ordre 2.

DÉFINITION A.4.1 (MULTICORRÉLATION)

Soit x un signal aléatoire à valeurs réelles. La multicorrélation d'ordre r de x est le cumulants d'ordre r des valeurs (variables aléatoires) du signal aux instants t_0, \dots, t_{r-1} :

$$C_{r,x}(\mathbf{t}) = \text{Cum}\{x(t_0), x(t_1), \dots, x(t_{r-1})\} \quad (\text{A.24})$$

où $\mathbf{t} = (t_0, \dots, t_{r-1})$.

■

Dans le cas général, la multicorrélation d'ordre r est fonction d'un temps absolu t_0 et des $r-1$ temps t_i . Lorsque le signal x est stationnaire au sens strict, ses statistiques sont invariantes par changement de l'origine des temps. Un des temps peut être pris comme origine (nous choisirons t_0 ici). La multicorrélation d'ordre r n'est alors fonction que des $r-1$ écarts de temps $\tau_i = t_i - t_0$:

$$C_{r,x}(\boldsymbol{\tau}) = \text{Cum}\{x(t_k), x(t_{k+\tau_1}), \dots, x(t_{k+\tau_{r-1}})\} \quad (\text{A.25})$$

où $\boldsymbol{\tau} = (\tau_1, \dots, \tau_{r-1})$.

REMARQUE

Dans la suite, par un abus de langage usuel, nous appellerons les multicorrélations « cumulants ». De plus, comme ne seront considérés que les cumulants d'ordre trois et quatre (les cumulants d'ordre supérieur à quatre ne présentent que peu d'intérêt pratique), une notation plus intuitive est adoptée : $C_{xxx}(\boldsymbol{\tau})$ pour les cumulants d'ordre trois et $C_{xxxx}(\boldsymbol{\tau})$ pour ceux d'ordre quatre (le nombre de composantes du vecteur $\boldsymbol{\tau}$ étant déterminé par l'ordre du cumulant). ■

A.5 Aspects pratiques liés à l'utilisation de cumulants

L'utilisation de cumulants dans un problème pratique soulève de nombreuses questions, pour la plupart liées à leur estimation ou à la redondance d'information. Dans les paragraphes suivants, nous essayons de donner quelques voies pour y répondre. Néanmoins, force est de constater qu'il ne semble pas y avoir de réponse définitive à la plupart de ces questions.

A.5.1 Lignes de cumulants

Lorsque l'on utilise les cumulants d'un signal stationnaire x , on obtient des fonctions dépendant de plusieurs variables : deux pour les cumulants d'ordre trois, trois pour les cumulants d'ordre quatre. En général, l'ensemble de ces valeurs fournit trop d'informations [LAC97]. Il est possible de se limiter à certaines parties de l'espace des délais, ne dépendant que d'une seule variable : ces « parties » de cumulants sont appelées *lignes de cumulants*. Les cumulants d'ordre trois de x s'écrivent :

$$C_{xxx}(\boldsymbol{\tau}) = C_{xxx}(\tau_1, \tau_2) = \text{E}\{x(t_k)x(t_{k+\tau_1})x(t_{k+\tau_2})\} \quad (\text{A.26})$$

Une ligne de cumulant est obtenue en fixant par exemple τ_2 à une valeur particulière. D'autres ont mis en avant l'intérêt particulier des diagonales [GM89], obtenues en fixant $\tau_1 = \tau_2$, mais, malgré tout, il n'existe pas de règle pour choisir les lignes de cumulants à utiliser¹.

¹Il est cependant à noter qu'il existe généralement des symétries qui aident à se restreindre à certaines parties de l'espace des délais.

Un autre aspect est que, dans la pratique, les cumulants sont estimés à partir des N échantillons disponibles. Or plus le délai est grand, moins l'estimation est fiable, puisque réalisée à l'aide de moins de points. Il est donc préférable de se limiter à utiliser les estimées des cumulants n'ayant pas un délai trop grand. À nouveau, il n'y a pas de règle pour se limiter à un délai particulier.

A.5.2 Estimation des cumulants

Supposons que N échantillons d'un signal aléatoire centré, stationnaire et ergodique x sont disponibles. Comparativement à l'estimation des corrélations, l'estimation des cumulants d'ordre supérieur est connue pour nécessiter un plus grand nombre d'échantillons pour réduire la variance des estimées obtenues à partir de données réelles [Men91]. Nous nous intéressons ici à l'estimation des cumulants, et dans la suite sont présentées les estimées classiques des cumulants de troisième et de quatrième ordre. Le théorème ergodique [Bol96], justifie le remplacement des moyennes d'ensemble par des moyennes temporelles. Une étude de la convergence des cumulants *échantillonnés* vers les vrais cumulants peut être trouvée dans [RN65]. Les covariances sur les estimées des cumulants sont disponibles dans [LAC97, PF89, FP90].

A.5.2.1 Cumulants d'ordre trois

Ce cas est simple, puisqu'il suffit de remplacer l'espérance mathématique par son estimateur classique pour obtenir une estimée consistante et non biaisée des cumulants d'ordre trois de x :

$$\hat{C}_{xxx}(\boldsymbol{\tau}) = \frac{1}{N} \sum_{k=1}^N x(t_k)x(t_{k+\tau_1})x(t_{k+\tau_2}) \quad (\text{A.27})$$

A.5.2.2 Cumulants d'ordre quatre

De la même manière que pour les cumulants d'ordre trois, une estimée naturelle des cumulants d'ordre quatre de x est :

$$\begin{aligned} \hat{C}_{xxxx}(\boldsymbol{\tau}) &= \frac{1}{N} \sum_{k=1}^N x(t_k)x(t_{k+\tau_1})x(t_{k+\tau_2})x(t_{k+\tau_3}) \\ &\quad - \frac{1}{N^2} \sum_{k=1}^N x(t_k)x(t_{k+\tau_1}) \sum_{k=1}^N x(t_{k+\tau_2})x(t_{k+\tau_3}) \\ &\quad - \frac{1}{N^2} \sum_{k=1}^N x(t_k)x(t_{k+\tau_2}) \sum_{k=1}^N x(t_{k+\tau_1})x(t_{k+\tau_3}) \\ &\quad - \frac{1}{N^2} \sum_{k=1}^N x(t_k)x(t_{k+\tau_3}) \sum_{k=1}^N x(t_{k+\tau_1})x(t_{k+\tau_2}) \quad (\text{A.28}) \end{aligned}$$

Des calculs simples montrent que l'estimateur (A.28) est asymptotiquement non biaisé. En revanche, il est biaisé pour N fini. Par conséquent, lorsque le nombre d'échantillon N est grand, (A.28) fournit un estimateur approximativement non biaisé. Néanmoins, lorsque l'on ne dispose que d'un petit nombre d'échantillons, le biais ne peut plus être négligé. Pour ces situations, il devient crucial d'utiliser un estimateur non biaisé, appelé k -statistique [LAC97, Bri01, SO94] :

$$\begin{aligned} \hat{C}_{xxxx}(\boldsymbol{\tau}) = \frac{1}{(N-1)(N-2)(N-3)} & \left\{ N(N+1) \sum_{k=1}^N x(t_k)x(t_{k+\tau_1})x(t_{k+\tau_2})x(t_{k+\tau_3}) \right. \\ & - 3(N-1) \sum_{k=1}^N x(t_k)x(t_{k+\tau_1}) \sum_{k=1}^N x(t_{k+\tau_2})x(t_{k+\tau_3}) \\ & - 3(N-1) \sum_{k=1}^N x(t_k)x(t_{k+\tau_2}) \sum_{k=1}^N x(t_{k+\tau_1})x(t_{k+\tau_3}) \\ & \left. - 3(N-1) \sum_{k=1}^N x(t_k)x(t_{k+\tau_3}) \sum_{k=1}^N x(t_{k+\tau_1})x(t_{k+\tau_2}) \right\} \quad (\text{A.29}) \end{aligned}$$

Notons enfin qu'une analyse très détaillée des k -statistiques est donnée dans [McC87].

Annexe B

Annexe des chapitres 2 et 3

B.1 Borne de Cramér-Rao

La borne de Cramér-Rao permet de juger de la précision d'un estimateur non-biaisé. C'est en effet une borne inférieure pour la matrice de covariance du vecteur des paramètres :

$$\text{cov}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \geq \mathbf{J}^{-1} \quad (\text{B.1a})$$

$$\mathbf{J} = \bar{\mathbb{E}} \left\{ \left(\frac{\partial \log \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \left(\frac{\partial \log \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^{\text{T}} \right\} \quad (\text{B.1b})$$

où $\ell(\boldsymbol{\theta})$ est la fonction de vraisemblance et \mathbf{J} la matrice d'information de Fisher. Le résultat (B.1) est vrai pour tout estimateur non biaisé, et pour un jeu de données de longueur infinie. Lors du calcul de la valeur de cette borne pour le modèle EIV, un aspect essentiel est la modélisation du signal d'entrée non bruité. Si les données sont supposées gaussiennes, il existe au moins deux manières de calculer la borne de Cramér-Rao, voir [KSS00] et [Söd06] pour plus de détails.

B.2 Corrélation des résidus de la méthode des moindres carrés

On utilise ici majoritairement les deux relations suivantes :

$$\bar{\mathbb{E}} \{ \boldsymbol{\varphi}(t_k) (y(t_k) - \boldsymbol{\varphi}^{\text{T}}(t_k) \boldsymbol{\theta}_{\text{ls}}^*) \} = 0 \quad (\text{B.2})$$

$$y(t_k) = \boldsymbol{\varphi}^{\text{T}}(t_k) \boldsymbol{\theta} + \tilde{y}(t_k) - \tilde{\boldsymbol{\varphi}}^{\text{T}}(t_k) \boldsymbol{\theta} \quad (\text{B.3})$$

On a donc :

$$J(\boldsymbol{\theta}_{\text{ls}}^*) = \bar{\mathbb{E}}\left\{ (y(t_k) - \boldsymbol{\varphi}^T(t_k)\boldsymbol{\theta}_{\text{ls}}^*)^2 \right\} \quad (\text{B.4})$$

$$= \bar{\mathbb{E}}\left\{ y(t_k) (y(t_k) - \boldsymbol{\varphi}^T(t_k)\boldsymbol{\theta}_{\text{ls}}^*) \right\} \quad (\text{B.5})$$

$$= \bar{\mathbb{E}}\left\{ (\tilde{y}(t_k) - \tilde{\boldsymbol{\varphi}}^T(t_k)\boldsymbol{\theta}) (y(t_k) - \boldsymbol{\varphi}^T(t_k)\boldsymbol{\theta}_{\text{ls}}^*) \right\} \quad (\text{B.6})$$

$$= \bar{\mathbb{E}}\left\{ (\tilde{y}(t_k) - \tilde{\boldsymbol{\varphi}}^T(t_k)\boldsymbol{\theta}) (\boldsymbol{\varphi}^T(t_k)(\boldsymbol{\theta} - \boldsymbol{\theta}_{\text{ls}}^*) + \tilde{y}(t_k) - \tilde{\boldsymbol{\varphi}}^T(t_k)\boldsymbol{\theta}) \right\} \quad (\text{B.7})$$

$$= (\boldsymbol{\theta} - \boldsymbol{\theta}_{\text{ls}}^*)^T \bar{\mathbb{E}}\left\{ \boldsymbol{\varphi}(t_k) (\tilde{y}(t_k) - \tilde{\boldsymbol{\varphi}}^T(t_k)\boldsymbol{\theta}) \right\} + \bar{\mathbb{E}}\left\{ (\tilde{y}(t_k) - \tilde{\boldsymbol{\varphi}}^T(t_k)\boldsymbol{\theta})^2 \right\} \quad (\text{B.8})$$

$$= (\boldsymbol{\theta} - \boldsymbol{\theta}_{\text{ls}}^*)^T \bar{\mathbb{E}}\left\{ \tilde{\boldsymbol{\varphi}}(t_k) (\tilde{y}(t_k) - \tilde{\boldsymbol{\varphi}}^T(t_k)\boldsymbol{\theta}) \right\} + \bar{\mathbb{E}}\left\{ (\tilde{y}(t_k) - \tilde{\boldsymbol{\varphi}}^T(t_k)\boldsymbol{\theta})^2 \right\} \quad (\text{B.9})$$

$$= (\boldsymbol{\theta} - \boldsymbol{\theta}_{\text{ls}}^*)^T \left\{ \mathbf{r}_{\tilde{\boldsymbol{\varphi}}\tilde{y}} - \mathbf{R}_{\tilde{\boldsymbol{\varphi}}\tilde{\boldsymbol{\varphi}}}\boldsymbol{\theta} \right\} + \lambda_{e\tilde{y}} - 2\boldsymbol{\theta}^T \mathbf{r}_{\tilde{\boldsymbol{\varphi}}\tilde{y}} + \boldsymbol{\theta}^T \mathbf{R}_{\tilde{\boldsymbol{\varphi}}\tilde{\boldsymbol{\varphi}}}\boldsymbol{\theta} \quad (\text{B.10})$$

$$= \lambda_{e\tilde{y}} + \boldsymbol{\theta}_{\text{ls}}^{*\text{T}} (\mathbf{R}_{\tilde{\boldsymbol{\varphi}}\tilde{\boldsymbol{\varphi}}}\boldsymbol{\theta} - \mathbf{r}_{\tilde{\boldsymbol{\varphi}}\tilde{y}}) - \boldsymbol{\theta}^T \mathbf{r}_{\tilde{\boldsymbol{\varphi}}\tilde{y}} \quad (\text{B.11})$$

L'équation (B.11) donne le résultat pour le cas général de bruits colorés en entrée et en sortie du système. Lorsque ces bruits sont tous deux blancs, (B.11) se simplifie en :

$$J(\boldsymbol{\theta}_{\text{ls}}^*) = \lambda_{e\tilde{y}} + \boldsymbol{\theta}_{\text{ls}}^{*\text{T}} \mathbf{R}_{\tilde{\boldsymbol{\varphi}}\tilde{\boldsymbol{\varphi}}}\boldsymbol{\theta} \quad (\text{B.12})$$

$$= \lambda_{e\tilde{y}} \left(1 + \mathbf{a}_{\text{ls}}^{*\text{T}} \mathbf{a} \right) + \lambda_{e\tilde{u}} \mathbf{b}_{\text{ls}}^{*\text{T}} \mathbf{b} \quad (\text{B.13})$$

B.3 Preuve des propositions relatives à la méthode bels

Tout d'abord, une partie de démonstration commune.

La matrice de covariance du vecteur de régression augmenté $\tilde{\boldsymbol{\varphi}}(t_k)$ peut être décomposée en :

$$\mathbf{R}_{\tilde{\boldsymbol{\varphi}}\tilde{\boldsymbol{\varphi}}} = \begin{bmatrix} \mathbf{R}_{\boldsymbol{\varphi}\boldsymbol{\varphi}} & \mathbf{R}_{\boldsymbol{\varphi}\boldsymbol{\mu}} \\ \mathbf{R}_{\boldsymbol{\varphi}\boldsymbol{\mu}}^T & \mathbf{R}_{\boldsymbol{\mu}\boldsymbol{\mu}} \end{bmatrix} \quad (\text{B.14})$$

L'inversion de cette matrice nous donne :

$$\mathbf{R}_{\tilde{\boldsymbol{\varphi}}\tilde{\boldsymbol{\varphi}}}^{-1} = \begin{bmatrix} \mathbf{R}_{\boldsymbol{\varphi}\boldsymbol{\varphi}}^{-1} + \mathbf{R}_{\boldsymbol{\varphi}\boldsymbol{\mu}}^{-1} \boldsymbol{\Lambda}^{-1} \mathbf{R}_{\boldsymbol{\varphi}\boldsymbol{\mu}}^T \mathbf{R}_{\boldsymbol{\varphi}\boldsymbol{\varphi}}^{-1} & -\mathbf{R}_{\boldsymbol{\varphi}\boldsymbol{\mu}}^{-1} \mathbf{R}_{\boldsymbol{\varphi}\boldsymbol{\mu}} \boldsymbol{\Lambda}^{-1} \\ -\boldsymbol{\Lambda}^{-1} \mathbf{R}_{\boldsymbol{\varphi}\boldsymbol{\mu}}^T \mathbf{R}_{\boldsymbol{\varphi}\boldsymbol{\varphi}}^{-1} & \boldsymbol{\Lambda}^{-1} \end{bmatrix} \quad (\text{B.15})$$

$$\text{où } \boldsymbol{\Lambda} = \mathbf{R}_{\boldsymbol{\mu}\boldsymbol{\mu}} - \mathbf{R}_{\boldsymbol{\varphi}\boldsymbol{\mu}}^T \mathbf{R}_{\boldsymbol{\varphi}\boldsymbol{\varphi}}^{-1} \mathbf{R}_{\boldsymbol{\varphi}\boldsymbol{\mu}} \quad (\text{B.16})$$

On utilise alors cette expression dans des équations qui dépendent des hypothèses sur les bruits, spécifiques à chaque proposition.

B.3.1 Preuve de la proposition 2.2.2 (fbels)

On remplace $\mathbf{R}_{\tilde{\boldsymbol{\varphi}}\tilde{\boldsymbol{\varphi}}}^{-1}$ par son expression (B.15) dans les deux équations suivantes :

$$\begin{cases} \bar{\boldsymbol{\theta}}_{\text{ls}}^* = \begin{bmatrix} \boldsymbol{\theta}_{\text{ls}}^* \\ \boldsymbol{\alpha}_{\text{ls}}^* \end{bmatrix} = \bar{\boldsymbol{\theta}} - \mathbf{R}_{\tilde{\boldsymbol{\varphi}}\tilde{\boldsymbol{\varphi}}}^{-1} \begin{bmatrix} \lambda_{e\tilde{y}} \mathbf{I}_{n_a} & 0 & 0 \\ 0 & \lambda_{e\tilde{u}} \mathbf{I}_{n_b} & 0 \\ 0 & 0 & \lambda_{e\tilde{v}} \mathbf{I}_{n_{\text{bels}}} \end{bmatrix} \bar{\boldsymbol{\theta}} \\ \bar{\boldsymbol{\theta}}_{\text{ls}}^* = \mathbf{R}_{\tilde{\boldsymbol{\varphi}}\tilde{\boldsymbol{\varphi}}}^{-1} \bar{\mathbb{E}}\{ \tilde{\boldsymbol{\varphi}}(t_k) y(t_k) \} \end{cases}$$

On obtient respectivement :

$$\begin{cases} \boldsymbol{\alpha}_{\text{ls}}^* = \boldsymbol{\Lambda}^{-1} \mathbf{R}_{\varphi\mu}^T \mathbf{R}_{\varphi\varphi}^{-1} \begin{bmatrix} \lambda_{e\tilde{y}} \mathbf{I}_{n_a} & 0 \\ 0 & \lambda_{e\tilde{u}} \mathbf{I}_{n_b} \end{bmatrix} \boldsymbol{\theta} \\ \boldsymbol{\alpha}_{\text{ls}}^* = -\boldsymbol{\Lambda}^{-1} \mathbf{R}_{\varphi\mu}^T \mathbf{R}_{\varphi\varphi}^{-1} \bar{\mathbf{E}}\{\boldsymbol{\varphi}(t_k)y(t_k)\} + \boldsymbol{\Lambda}^{-1} \bar{\mathbf{E}}\{\boldsymbol{\mu}(t_k)y(t_k)\} \end{cases}$$

Ces deux dernières équations permettent alors d'écrire :

$$\mathbf{R}_{\varphi\mu}^T \mathbf{R}_{\varphi\varphi}^{-1} \begin{bmatrix} \lambda_{e\tilde{y}} \mathbf{a} \\ \lambda_{e\tilde{u}} \mathbf{b} \end{bmatrix} = \mathbf{r}_{\mu y} - \mathbf{R}_{\varphi\mu}^T \mathbf{R}_{\varphi\varphi}^{-1} \mathbf{r}_{\varphi y} \quad (\text{B.17})$$

$$\iff \mathbf{R}_{\varphi\mu}^T \mathbf{R}_{\varphi\varphi}^{-1} \begin{bmatrix} \lambda_{e\tilde{y}} \mathbf{a} \\ \lambda_{e\tilde{u}} \mathbf{b} \end{bmatrix} = \mathbf{r}_{\mu y} - \mathbf{R}_{\varphi\mu}^T \boldsymbol{\theta}_{\text{ls}}^* \quad (\text{B.18})$$

$$\iff \lambda_{e\tilde{y}} \mathbf{R}_{\varphi\mu}^T \boldsymbol{\Lambda}_{n_a} \mathbf{a} + \lambda_{e\tilde{u}} \mathbf{R}_{\varphi\mu}^T \boldsymbol{\Lambda}_{n_b} \mathbf{b} = \mathbf{r}_{\mu y} - \mathbf{R}_{\varphi\mu}^T \boldsymbol{\theta}_{\text{ls}}^* \quad (\text{B.19})$$

ce qui complète la preuve.

B.3.2 Preuve de la proposition 2.2.5 (bels2)

On remplace $\mathbf{R}_{\tilde{\varphi}\tilde{\varphi}}^{-1}$ par son expression (B.15) dans les deux équations suivantes :

$$\left\{ \begin{array}{l} \bar{\boldsymbol{\theta}}_{\text{ls}}^* = \begin{bmatrix} \boldsymbol{\theta}_{\text{ls}}^* \\ \boldsymbol{\beta}_{\text{ls}}^* \end{bmatrix} = \bar{\boldsymbol{\theta}} + \mathbf{R}_{\tilde{\varphi}\tilde{\varphi}}^{-1} \left\{ \begin{bmatrix} \mathbf{r}_{\varphi\tilde{y}} \\ 0 \end{bmatrix} - \begin{bmatrix} \mathbf{R}_{\tilde{\varphi}\tilde{\varphi}} & 0 \\ 0 & \lambda_{e\tilde{u}} \mathbf{I}_{n_{\text{bels}}} \end{bmatrix} \bar{\boldsymbol{\theta}} \right\} \\ \bar{\boldsymbol{\theta}}_{\text{ls}}^* = \mathbf{R}_{\tilde{\varphi}\tilde{\varphi}}^{-1} \mathbf{r}_{\tilde{\varphi}y} \end{array} \right.$$

On obtient respectivement :

$$\begin{cases} \boldsymbol{\beta}_{\text{ls}}^* = -\boldsymbol{\Lambda}^{-1} \mathbf{R}_{\varphi\mu}^T \mathbf{R}_{\varphi\varphi}^{-1} (\mathbf{r}_{\varphi\tilde{y}} - \mathbf{R}_{\tilde{\varphi}\tilde{\varphi}} \boldsymbol{\theta}) \\ \boldsymbol{\beta}_{\text{ls}}^* = -\boldsymbol{\Lambda}^{-1} \mathbf{R}_{\varphi\mu}^T \mathbf{R}_{\varphi\varphi}^{-1} \mathbf{r}_{\varphi y} + \boldsymbol{\Lambda}^{-1} \mathbf{r}_{\mu y} \end{cases}$$

D'où :

$$\mathbf{r}_{\mu y} - \mathbf{R}_{\varphi\mu}^T \mathbf{R}_{\varphi\varphi}^{-1} \mathbf{r}_{\varphi y} = \mathbf{R}_{\varphi\mu}^T \mathbf{R}_{\varphi\varphi}^{-1} (\mathbf{R}_{\tilde{\varphi}\tilde{\varphi}} \boldsymbol{\theta} - \mathbf{r}_{\varphi\tilde{y}}) \quad (\text{B.20})$$

$$\iff \mathbf{r}_{\mu y} - \mathbf{R}_{\varphi\mu}^T \boldsymbol{\theta}_{\text{ls}}^* = \mathbf{R}_{\varphi\mu}^T \mathbf{R}_{\varphi\varphi}^{-1} \mathbf{M}(\boldsymbol{\theta}) \boldsymbol{\rho}_{\tilde{u}, \tilde{y}} \quad (\text{B.21})$$

ce qui complète la preuve.

B.4 Preuve de la proposition 2.2.4 (pbcls)

On considère la corrélation entre les résidus des méthodes des moindres carrés et du prédicteur à un pas de la sortie :

$$\bar{\mathbb{E}}\{v(t_{k-1}, \boldsymbol{\theta}_{ls}^*)v(t_k, \boldsymbol{\theta}_{pbcls}^*)\} = \bar{\mathbb{E}}\{(y(t_{k-1}) - \boldsymbol{\varphi}^T(t_{k-1})\boldsymbol{\theta}_{ls}^*)v(t_k, \boldsymbol{\theta}_{pbcls}^*)\} \quad (\text{B.22})$$

$$= \bar{\mathbb{E}}\{(\boldsymbol{\varphi}^T(t_{k-1})(\boldsymbol{\theta} - \boldsymbol{\theta}_{ls}^*) + v(t_{k-1}, \boldsymbol{\theta}))v(t_k, \boldsymbol{\theta}_{pbcls}^*)\} \quad (\text{B.23})$$

$$= \bar{\mathbb{E}}\{v(t_{k-1}, \boldsymbol{\theta})v(t_k, \boldsymbol{\theta}_{pbcls}^*)\} \quad (\text{B.24})$$

$$= \bar{\mathbb{E}}\{(\tilde{y}(t_{k-1}) - \tilde{\boldsymbol{\varphi}}^T(t_{k-1})\boldsymbol{\theta})(\tilde{y}(t_k) - \tilde{\boldsymbol{\varphi}}^T(t_{k-1})\boldsymbol{\theta}_{pbcls}^*)\} \quad (\text{B.25})$$

$$= \boldsymbol{\theta}^T \mathbf{R}_{\tilde{\boldsymbol{\varphi}}\tilde{\boldsymbol{\varphi}}} \boldsymbol{\theta}_{pbcls}^* \quad (\text{B.26})$$

$$= \lambda_{e\bar{y}} \mathbf{a}^T \mathbf{a}_{pbcls}^* + \lambda_{e\bar{u}} \mathbf{b}^T \mathbf{b}_{pbcls}^* \quad (\text{B.27})$$

B.5 Corrélation des résidus de la méthode des moindres carrés filtrés

De même qu'à la section B.2, on utilise ici majoritairement les deux relations suivantes :

$$\bar{\mathbb{E}}\{\boldsymbol{\varphi}_f(t_k)(y_f(t_k) - \boldsymbol{\varphi}_f^T(t_k)\boldsymbol{\theta}_{fls})\} = 0 \quad (\text{B.28})$$

$$y_f(t_k) = \boldsymbol{\varphi}_f^T(t_k)\boldsymbol{\theta} + v_f(t_k, \boldsymbol{\theta}) \quad (\text{B.29})$$

Alors :

$$J(\boldsymbol{\theta}_{fls}^*) = \bar{\mathbb{E}}\left\{(y_f(t_k) - \boldsymbol{\varphi}_f^T(t_k)\boldsymbol{\theta}_{fls}^*)^2\right\} \quad (\text{B.30})$$

$$= \bar{\mathbb{E}}\{v_f(t_k, \boldsymbol{\theta})y_f(t_k)\} - \boldsymbol{\theta}_{fls}^{*,T} \bar{\mathbb{E}}\{\boldsymbol{\varphi}_f(t_k)v_f(t_k, \boldsymbol{\theta})\} \quad (\text{B.31})$$

Or on a d'une part :

$$\bar{\mathbb{E}}\{y_f(t_k)v_f(t_k, \boldsymbol{\theta})\} = \bar{\mathbb{E}}\left\{\left(\sum_{i=0}^{n_\alpha} \alpha_i y(t_{k-i})\right)\left(\sum_{j=0}^{n_\alpha} \alpha_j v(t_{k-j}, \boldsymbol{\theta})\right)\right\} \quad (\text{B.32})$$

$$= \sum_{i=0}^{n_\alpha} \sum_{j=0}^{n_\alpha} \alpha_i \alpha_j \bar{\mathbb{E}}\{\tilde{y}(t_{k-i})(\tilde{y}(t_{k-j}) - \tilde{\boldsymbol{\varphi}}^T(t_{k-j})\boldsymbol{\theta})\} \quad (\text{B.33})$$

$$= \sum_{i=0}^{n_\alpha} \sum_{j=0}^{n_\alpha} \alpha_i \alpha_j \bar{\mathbb{E}}\{\tilde{y}(t_{k-i})\tilde{y}(t_{k-j})\} - \sum_{i=0}^{n_\alpha} \sum_{j=0}^{n_\alpha} \alpha_i \alpha_j \bar{\mathbb{E}}\{\tilde{y}(t_{k-i})(-\tilde{\mathbf{y}}^T(t_{k-j}) \tilde{\mathbf{u}}^T(t_{k-j}))\boldsymbol{\theta}\} \quad (\text{B.34})$$

$$= \sum_{i=0}^{n_\alpha} \sum_{j=0}^{n_\alpha} \alpha_i \alpha_j \bar{\mathbb{E}}\{\tilde{y}(t_{k-i})\tilde{y}(t_{k-j})\} + \sum_{i=0}^{n_\alpha} \sum_{j=0}^{n_\alpha} \alpha_i \alpha_j \bar{\mathbb{E}}\{\tilde{y}(t_{k-i})\tilde{\mathbf{y}}^T(t_{k-j})\} \mathbf{a} \quad (\text{B.35})$$

Si $\tilde{y}(t_k)$ est un bruit blanc :

$$\bar{\mathbb{E}}\{v_f(t_k, \boldsymbol{\theta})y_f(t_k)\} = \lambda_{e\tilde{y}} \left(\sum_{i=0}^{n_\alpha} \alpha_i^2 + \mathbf{v}^T \mathbf{a} \right) \quad (\text{B.36})$$

où :

$$\mathbf{v}^T = \left[\sum_{i=0}^{n_\alpha-1} \alpha_i \alpha_{i+1} \quad \sum_{i=0}^{n_\alpha-2} \alpha_i \alpha_{i+2} \quad \dots \quad \sum_{i=0}^{n_\alpha-n_a} \alpha_i \alpha_{i+n_a} \right] \quad (\text{B.37})$$

D'autre part :

$$\bar{\mathbb{E}}\{\boldsymbol{\varphi}_f(t_k)v_f(t_k, \boldsymbol{\theta})\} = \sum_{i=0}^{n_\alpha} \sum_{j=0}^{n_\alpha} \alpha_i \alpha_j \bar{\mathbb{E}}\{\boldsymbol{\varphi}(t_{k-i})v(t_{k-j}, \boldsymbol{\theta})\} \quad (\text{B.38})$$

$$= \sum_{i=0}^{n_\alpha} \sum_{j=0}^{n_\alpha} \alpha_i \alpha_j \bar{\mathbb{E}}\{\tilde{\boldsymbol{\varphi}}(t_{k-i}) (\tilde{y}(t_{k-j}) - \tilde{\boldsymbol{\varphi}}^T(t_{k-j})\boldsymbol{\theta})\} \quad (\text{B.39})$$

$$= \sum_{i=0}^{n_\alpha} \sum_{j=0}^{n_\alpha} \alpha_i \alpha_j \bar{\mathbb{E}}\{\tilde{y}(t_{k-j})\tilde{\boldsymbol{\varphi}}(t_{k-i})\} \\ - \sum_{i=0}^{n_\alpha} \sum_{j=0}^{n_\alpha} \alpha_i \alpha_j \bar{\mathbb{E}}\{\tilde{\boldsymbol{\varphi}}(t_{k-i})\tilde{\boldsymbol{\varphi}}^T(t_{k-j})\} \boldsymbol{\theta} \quad (\text{B.40})$$

Si $\tilde{u}(t_k)$ et $\tilde{y}(t_k)$ sont des bruits blancs :

$$\bar{\mathbb{E}}\{\boldsymbol{\varphi}_f(t_k)v_f(t_k, \boldsymbol{\theta})\} = -\lambda_{e\tilde{y}} \begin{bmatrix} \mathbf{v} \\ \mathbf{0} \end{bmatrix} - \mathbf{D}\boldsymbol{\theta} \quad (\text{B.41})$$

où :

$$\mathbf{D} = \begin{bmatrix} \lambda_{e\tilde{y}} \mathbf{H}_{n_a} & \mathbf{0}_{n_a \times n_b} \\ \mathbf{0}_{n_b \times n_a} & \lambda_{e\tilde{u}} \mathbf{H}_{n_b} \end{bmatrix} \quad (\text{B.42})$$

$$\mathbf{H}_{n_a} = \sum_{j=0}^{n_a-1} \sum_{i=0}^{n_\alpha-j} \alpha_i \alpha_{i+j} \mathcal{I}_{n_a}(j) \quad (\text{B.43})$$

$$\mathbf{H}_{n_b} = \sum_{j=0}^{n_b-1} \sum_{i=0}^{n_\alpha-j} \alpha_i \alpha_{i+j} \mathcal{I}_{n_b}(j) \quad (\text{B.44})$$

Finalement :

$$J(\boldsymbol{\theta}_{\text{fls}}^*) = \bar{\mathbb{E}}\{v_f(t_k, \boldsymbol{\theta})y_f(t_k)\} - \boldsymbol{\theta}_{\text{fls}}^{*\text{T}} \bar{\mathbb{E}}\{\boldsymbol{\varphi}_f(t_k)v_f(t_k, \boldsymbol{\theta})\} \\ = \lambda_{e\tilde{y}} \left(\sum_{i=0}^{n_\alpha} \alpha_i^2 + \mathbf{v}^T \mathbf{a} \right) + \boldsymbol{\theta}_{\text{fls}}^{*\text{T}} \left[\lambda_{e\tilde{y}} \begin{bmatrix} \mathbf{v} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \lambda_{e\tilde{y}} \mathbf{H}_{n_a} \mathbf{a} \\ \lambda_{e\tilde{u}} \mathbf{H}_{n_b} \mathbf{b} \end{bmatrix} \right] \\ = \lambda_{e\tilde{y}} \left(\sum_{i=0}^{n_\alpha} \alpha_i^2 + \mathbf{v}^T (\mathbf{a} + \mathbf{a}_{\text{fls}}^*) + \mathbf{a}_{\text{fls}}^{*\text{T}} \mathbf{H}_{n_a} \mathbf{a} \right) + \lambda_{e\tilde{u}} \mathbf{b}_{\text{fls}}^{*\text{T}} \mathbf{H}_{n_b} \mathbf{b} \quad (\text{B.45})$$

B.6 Preuves des résultats du paragraphe 2.3.5

B.6.1 Preuve du lemme 2.3.1

$$\left(\mathbf{R} + \tilde{\mathbf{R}}\right)^{-1} = \left(\mathbf{R} + \tilde{\mathbf{R}}\right)^{-1} \left(\left(\mathbf{R} + \tilde{\mathbf{R}}\right) - \tilde{\mathbf{R}} \right) \mathbf{R}^{-1} \quad (\text{B.46})$$

$$= \mathbf{R}^{-1} - \left(\mathbf{R} + \tilde{\mathbf{R}}\right)^{-1} \tilde{\mathbf{R}} \mathbf{R}^{-1} \quad (\text{B.47})$$

$$= \mathbf{R}^{-1} - \left(\mathbf{R}^{-1} - \left(\mathbf{R} + \tilde{\mathbf{R}}\right)^{-1} \tilde{\mathbf{R}} \mathbf{R}^{-1} \right) \tilde{\mathbf{R}} \mathbf{R}^{-1} \quad (\text{B.48})$$

$$= \mathbf{R}^{-1} - \mathbf{R}^{-1} \tilde{\mathbf{R}} \mathbf{R}^{-1} + \left(\mathbf{R} + \tilde{\mathbf{R}}\right)^{-1} \tilde{\mathbf{R}} \mathbf{R}^{-1} \tilde{\mathbf{R}} \mathbf{R}^{-1} \quad (\text{B.49})$$

Puisque $\|\tilde{\mathbf{R}}\|$ est petite en comparaison avec $\|\mathbf{R}\|$, le dernier terme de (B.49) peut être négligé.

B.6.2 Preuve du théorème 2.3.1

On a :

$$\hat{\boldsymbol{\theta}}_{\text{tocls}} - \boldsymbol{\theta} \approx (\boldsymbol{\Phi} \boldsymbol{\Phi}^{\text{T}})^{-1} \boldsymbol{\Phi} \left(\hat{\mathbf{C}} - \hat{\boldsymbol{\Phi}}^{\text{T}} \boldsymbol{\theta} \right) \quad (\text{B.50})$$

$$\approx (\boldsymbol{\Phi} \boldsymbol{\Phi}^{\text{T}})^{-1} \boldsymbol{\Phi} \mathbf{w} \quad (\text{B.51})$$

Notons que $\boldsymbol{\Phi}$ est une matrice déterministe, tandis que \mathbf{w} est un terme stochastique de moyenne nulle et qui tend vers 0 lorsque N tend vers l'infini. Pour prouver le théorème, il suffit d'appliquer le lemme 9A.1 donné à la page 266 de l'ouvrage [Lju87], à :

$$Z_N = \mathbf{w} = \sum_{k=1}^N x_N(t_k) \quad (\text{B.52})$$

$$\text{où } x_N(t_k) = u^2(t_k) \left[\begin{pmatrix} y(t_k) \\ \vdots \\ y(t_{k+M-1}) \end{pmatrix} - \begin{pmatrix} \boldsymbol{\varphi}^{\text{T}}(t_k) \\ \vdots \\ \boldsymbol{\varphi}^{\text{T}}(t_{k+M-1}) \end{pmatrix} \boldsymbol{\theta} \right] \quad (\text{B.53})$$

Les termes $x_N(t_k)$ sont clairement de moyenne nulle, et les hypothèses (9A.4) et (9A.5) découlent de **Ht1**. D'après le lemme 9A.1, on a alors :

$$\mathbf{w} \xrightarrow{\text{dist}} \mathcal{N}(0, \mathbf{Q}) \quad (\text{B.54})$$

$$\mathbf{Q} = \lim_{N \rightarrow \infty} N \bar{\mathbf{E}} \{ \mathbf{w} \mathbf{w}^{\text{T}} \} \quad (\text{B.55})$$

Le théorème suit maintenant de (B.51), (B.54), (B.55).

B.6.3 Preuve du lemme 2.3.2

Le $j^{\text{ième}}$ élément du vecteur \mathbf{w} est donné par ($0 \leq j \leq M$) :

$$\mathbf{w}_{j+1} = \hat{C}_{uyu}(j, 0) - \hat{\boldsymbol{\varphi}}_{cum}^T(j, 0)\boldsymbol{\theta} = \frac{1}{N} \sum_t u^2(t_k)v(t_{k+j}, \boldsymbol{\theta}) \quad (\text{B.56})$$

où $v(t_k, \boldsymbol{\theta})$ est le bruit composite habituel :

$$v(t_{k+j}, \boldsymbol{\theta}) = A(q^{-1})\tilde{y}(t_{k+j}) - B(q^{-1})\tilde{u}(t_{k+j}) \quad (\text{B.57})$$

Pour éviter toute lourdeur, on notera $v(t_{k+j}, \boldsymbol{\theta}) = v(t_{k+j})$ dans la suite. D'après les hypothèses **H5**, **He2** et **Hb6**, on a :

$$\begin{aligned} \bar{\mathbb{E}}\{\mathbf{w}_{j_1+1}\mathbf{w}_{j_2+1}\} &= \bar{\mathbb{E}}\left\{\frac{1}{N^2} \sum_{k=1}^N \sum_{l=1}^N u^2(t_k)u^2(t_l)v(t_{k+j_1})v(t_{l+j_2})\right\} \\ &= \frac{1}{N^2} \sum_{k=1}^N \sum_{l=1}^N \left\{ \bar{\mathbb{E}}\{u_0^2(t_k)u_0^2(t_l)\} \bar{\mathbb{E}}\{v(t_{k+j_1})v(t_{l+j_2})\} \right. \\ &\quad + \bar{\mathbb{E}}\{u_0^2(t_k)\} \bar{\mathbb{E}}\{\tilde{u}^2(t_l)v(t_{k+j_1})v(t_{l+j_2})\} \\ &\quad + 4\bar{\mathbb{E}}\{u_0(t_k)u_0(t_l)\} \bar{\mathbb{E}}\{\tilde{u}(t_k)\tilde{u}(t_l)v(t_{k+j_1})v(t_{l+j_2})\} \\ &\quad + \bar{\mathbb{E}}\{u_0^2(t_l)\} \bar{\mathbb{E}}\{\tilde{u}^2(t_k)v(t_{k+j_1})v(t_{l+j_2})\} \\ &\quad \left. + \bar{\mathbb{E}}\{\tilde{u}^2(t_k)\tilde{u}^2(t_l)v(t_{k+j_1})v(t_{l+j_2})\} \right\} \quad (\text{B.58}) \end{aligned}$$

Posons $\tau = k - l$. Alors :

$$\begin{aligned} \bar{\mathbb{E}}\{\mathbf{w}_{j_1+1}\mathbf{w}_{j_2+1}\} &= \sum_{\tau=-N}^N \frac{N - |\tau|}{N^2} \left\{ \mu_{u_0^4}(0, \tau, \tau) \bar{\mathbb{E}}\{v(t_{k+\tau+j_1})v(t_{k+j_2})\} \right. \\ &\quad + \mu_{u_0^2}(0) \bar{\mathbb{E}}\{\tilde{u}^2(t_k)v(t_{k+\tau+j_1})v(t_{k+j_2})\} + 4\mu_{u_0^2}(\tau) \bar{\mathbb{E}}\{\tilde{u}(t_{k+\tau})\tilde{u}(t_k)v(t_{k+\tau+j_1})v(t_{k+j_2})\} \\ &\quad \left. + \mu_{u_0^2}(0) \bar{\mathbb{E}}\{\tilde{u}^2(t_{k+\tau})v(t_{k+\tau+j_1})v(t_{k+j_2})\} + \bar{\mathbb{E}}\{\tilde{u}^2(t_{k+\tau})\tilde{u}^2(t_k)v(t_{k+\tau+j_1})v(t_{k+j_2})\} \right\} \quad (\text{B.59}) \end{aligned}$$

où l'on a utilisé la notation (2.181). Or :

$$\begin{aligned} \bar{\mathbb{E}}\{v(t_{k+\tau+j_1})v(t_{k+j_2})\} &= \sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_a} a_{l_1}a_{l_2} \bar{\mathbb{E}}\{\tilde{y}(t_{k+\tau+j_1-l_1})\tilde{y}(t_{k+j_2-l_2})\} \\ &\quad - \sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_b} a_{l_1}b_{l_2} \bar{\mathbb{E}}\{\tilde{y}(t_{k+\tau+j_1-l_1})\tilde{u}(t_{k+j_2-l_2})\} - \sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_b} a_{l_1}b_{l_2} \bar{\mathbb{E}}\{\tilde{u}(t_{k+\tau+j_1-l_2})\tilde{y}(t_{k+j_2-l_1})\} \\ &\quad + \sum_{l_1=0}^{n_b} \sum_{l_2=0}^{n_b} b_{l_1}b_{l_2} \bar{\mathbb{E}}\{\tilde{u}(t_{k+\tau+j_1-l_1})\tilde{u}(t_{k+j_2-l_2})\} \quad (\text{B.60}) \end{aligned}$$

Donc :

$$\begin{aligned} \bar{E}\{v(t_{k+\tau+j_1})v(t_{k+j_2})\} &= \sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_a} a_{l_1} a_{l_2} \mu_{\tilde{y}^2}(\tau + j_1 - l_1 - j_2 + l_2) \\ &\quad - \sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_b} a_{l_1} b_{l_2} \left[\mu_{\tilde{u}\tilde{y}}(\tau + j_1 - l_1 - j_2 + l_2) + \mu_{\tilde{u}\tilde{y}}(-\tau - j_1 - l_1 + j_2 + l_2) \right] \\ &\quad + \sum_{l_1=0}^{n_b} \sum_{l_2=0}^{n_b} b_{l_1} b_{l_2} \mu_{\tilde{u}^2}(\tau + j_1 - l_1 - j_2 + l_2) \quad (\text{B.61}) \end{aligned}$$

Si l'on suit les mêmes étapes pour les quatre autres espérances dans (B.59) on obtient :

$$\begin{aligned} \bar{E}\{\tilde{u}(t_{k+\tau})\tilde{u}(t_k)v(t_{k+\tau+j_1})v(t_{k+j_2})\} &= \sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_a} a_{l_1} a_{l_2} \mu_{\tilde{u}^2\tilde{y}^2}(\tau, \tau + j_1 - l_1, j_2 - l_2) \\ &\quad - \sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_b} a_{l_1} b_{l_2} \left[\mu_{\tilde{u}^3\tilde{y}}(\tau, j_2 - l_2, \tau + j_1 - l_1) + \mu_{\tilde{u}^3\tilde{y}}(\tau, \tau + j_1 - l_2, j_2 - l_1) \right] \\ &\quad + \sum_{l_1=0}^{n_b} \sum_{l_2=0}^{n_b} b_{l_1} b_{l_2} \mu_{\tilde{u}^4}(\tau, \tau + j_1 - l_1, j_2 - l_2) \quad (\text{B.62}) \end{aligned}$$

$$\begin{aligned} \bar{E}\{\tilde{u}^2(t_k)v(t_{k+\tau+j_1})v(t_{k+j_2})\} &= \sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_a} a_{l_1} a_{l_2} \mu_{\tilde{u}^2\tilde{y}^2}(0, \tau + j_1 - l_1, j_2 - l_2) \\ &\quad - \sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_b} a_{l_1} b_{l_2} \left[\mu_{\tilde{u}^3\tilde{y}}(0, j_2 - l_2, \tau + j_1 - l_1) + \mu_{\tilde{u}^3\tilde{y}}(0, \tau + j_1 - l_2, j_2 - l_1) \right] \\ &\quad + \sum_{l_1=0}^{n_b} \sum_{l_2=0}^{n_b} b_{l_1} b_{l_2} \mu_{\tilde{u}^4}(0, \tau + j_1 - l_1, j_2 - l_2) \quad (\text{B.63}) \end{aligned}$$

$$\begin{aligned} \bar{E}\{\tilde{u}^2(t_{k+\tau})v(t_{k+\tau+j_1})v(t_{k+j_2})\} &= \sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_a} a_{l_1} a_{l_2} \mu_{\tilde{u}^2\tilde{y}^2}(0, j_1 - l_1, j_2 - l_2 - \tau) \\ &\quad - \sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_b} a_{l_1} b_{l_2} \left[\mu_{\tilde{u}^3\tilde{y}}(0, j_2 - l_2 - \tau, j_1 - l_1) + \mu_{\tilde{u}^3\tilde{y}}(0, j_1 - l_2, j_2 - l_1 - \tau) \right] \\ &\quad + \sum_{l_1=0}^{n_b} \sum_{l_2=0}^{n_b} b_{l_1} b_{l_2} \mu_{\tilde{u}^4}(0, j_1 - l_1, j_2 - l_2 - \tau) \quad (\text{B.64}) \end{aligned}$$

$$\begin{aligned} \bar{E}\{\tilde{u}^2(t_{k+\tau})\tilde{u}^2(t_k)v(t_{k+\tau+j_1})v(t_{k+j_2})\} &= \sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_a} a_{l_1} a_{l_2} \mu_{\tilde{u}^4\tilde{y}^2}(0, \tau, \tau, \tau + j_1 - l_1, j_2 - l_2) \\ &\quad - \sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_b} a_{l_1} b_{l_2} \left[\mu_{\tilde{u}^5\tilde{y}}(0, \tau, \tau, j_2 - l_2, \tau + j_1 - l_1) + \mu_{\tilde{u}^5\tilde{y}}(0, \tau, \tau, \tau + j_1 - l_2, j_2 - l_1) \right] \\ &\quad + \sum_{l_1=0}^{n_b} \sum_{l_2=0}^{n_b} b_{l_1} b_{l_2} \mu_{\tilde{u}^6}(0, \tau, \tau, \tau + j_1 - l_1, j_2 - l_2) \quad (\text{B.65}) \end{aligned}$$

Donc finalement :

$$\begin{aligned}
\bar{\mathbb{E}}\{\mathbf{w}_{j_1+1}\mathbf{w}_{j_2+1}\} &= \sum_{\tau=-N}^N \frac{N-|\tau|}{N^2} \left\{ \right. \\
&\sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_a} a_{l_1} a_{l_2} \left[\mu_{u_o^4}(0, \tau, \tau) \mu_{\tilde{y}^2}(\tau + j_1 - l_1 - j_2 + l_2) + 4\mu_{u_o^2}(\tau) \mu_{\tilde{u}^2 \tilde{y}^2}(\tau, \tau + j_1 - l_1, j_2 - l_2) \right. \\
&+ \mu_{u_o^2}(0) \left(\mu_{\tilde{u}^2 \tilde{y}^2}(0, \tau + j_1 - l_1, j_2 - l_2) + \mu_{\tilde{u}^2 \tilde{y}^2}(0, j_1 - l_1, j_2 - l_2 - \tau) \right) \\
&+ \left. \mu_{\tilde{u}^4 \tilde{y}^2}(0, \tau, \tau, \tau + j_1 - l_1, j_2 - l_2) \right] \\
&- \sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_b} a_{l_1} b_{l_2} \left[\mu_{u_o^4}(0, \tau, \tau) \left(\mu_{\tilde{u} \tilde{y}}(\tau + j_1 - l_1 - j_2 + l_2) + \mu_{\tilde{u} \tilde{y}}(-\tau - j_1 - l_1 + j_2 + l_2) \right) \right. \\
&+ 4\mu_{u_o^2}(\tau) \left(\mu_{\tilde{u}^3 \tilde{y}}(\tau, j_2 - l_2, \tau + j_1 - l_1) + \mu_{\tilde{u}^3 \tilde{y}}(\tau, \tau + j_1 - l_2, j_2 - l_1) \right) \\
&+ \mu_{u_o^2}(0) \left(\mu_{\tilde{u}^3 \tilde{y}}(0, j_2 - l_2, \tau + j_1 - l_1) + \mu_{\tilde{u}^3 \tilde{y}}(0, \tau + j_1 - l_2, j_2 - l_1) \right. \\
&+ \left. \mu_{\tilde{u}^3 \tilde{y}}(0, j_2 - l_2 - \tau, j_1 - l_1) + \mu_{\tilde{u}^3 \tilde{y}}(0, j_1 - l_2, j_2 - l_1 - \tau) \right) \\
&+ \left. \mu_{\tilde{u}^5 \tilde{y}}(0, \tau, \tau, j_2 - l_2, \tau + j_1 - l_1) + \mu_{\tilde{u}^5 \tilde{y}}(0, \tau, \tau, \tau + j_1 - l_2, j_2 - l_1) \right] \\
&+ \sum_{l_1=0}^{n_b} \sum_{l_2=0}^{n_b} b_{l_1} b_{l_2} \left[\mu_{u_o^4}(0, \tau, \tau) \mu_{\tilde{u}^2}(\tau + j_1 - l_1 - j_2 + l_2) \right. \\
&+ 4\mu_{u_o^2}(\tau) \mu_{\tilde{u}^4}(\tau, \tau + j_1 - l_1, j_2 - l_2) + \mu_{u_o^2}(0) \left(\mu_{\tilde{u}^4}(0, \tau + j_1 - l_1, j_2 - l_2) \right. \\
&+ \left. \mu_{\tilde{u}^4}(0, j_1 - l_1, j_2 - l_2 - \tau) \right) + \left. \mu_{\tilde{u}^6}(0, \tau, \tau, \tau + j_1 - l_1, j_2 - l_2) \right] \left. \right\} \quad (\text{B.66})
\end{aligned}$$

Ce qui prouve le lemme.

B.6.4 Preuve du lemme 2.3.3

D'après les hypothèses du lemme 2.3.3, on a :

$$\begin{aligned}
\mu_{\tilde{u}^4}(\boldsymbol{\tau}) &= \bar{\mathbb{E}}\{\tilde{u}(t_k) \tilde{u}(t_{k+\tau_1}) \tilde{u}(t_{k+\tau_2}) \tilde{u}(t_{k+\tau_3})\} \\
&= \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \sum_{j_3=0}^{\infty} \sum_{j_4=0}^{\infty} h_{j_1}^{\tilde{u}} h_{j_2}^{\tilde{u}} h_{j_3}^{\tilde{u}} h_{j_4}^{\tilde{u}} \bar{\mathbb{E}}\{e^{\tilde{u}}(t_{k-j_1}) e^{\tilde{u}}(t_{k+\tau_1-j_2}) e^{\tilde{u}}(t_{k+\tau_2-j_3}) e^{\tilde{u}}(t_{k+\tau_3-j_4})\} \quad (\text{B.67})
\end{aligned}$$

Comme $e^{\tilde{u}}(t_k)$ est un bruit blanc, l'espérance dans (B.67) est non-nulle uniquement lorsque les arguments sont égaux deux à deux, ou tous égaux. Subséquemment :

$$\begin{aligned}
&\bar{\mathbb{E}}\{e^{\tilde{u}}(t_{k-j_1}) e^{\tilde{u}}(t_{k+\tau_1-j_2}) e^{\tilde{u}}(t_{k+\tau_2-j_3}) e^{\tilde{u}}(t_{k+\tau_3-j_4})\} \\
&= \mu_{e^{\tilde{u}^2}}^2 \left(\delta_{j_1, j_2 - \tau_1} \delta_{j_3 - \tau_2, j_4 - \tau_3} + \delta_{j_1, j_3 - \tau_2} \delta_{j_2 - \tau_1, j_4 - \tau_3} + \delta_{j_1, j_4 - \tau_3} \delta_{j_2 - \tau_1, j_3 - \tau_2} \right) \\
&\quad + \left(\mu_{e^{\tilde{u}^4}} - 3\mu_{e^{\tilde{u}^2}}^2 \right) \delta_{j_1, j_2 - \tau_1} \delta_{j_2 - \tau_1, j_3 - \tau_2} \delta_{j_3 - \tau_2, j_4 - \tau_3} \quad (\text{B.68})
\end{aligned}$$

où $\delta_{\cdot,\cdot}$ est le symbole de Kronecker. Si l'on insère cette dernière expression dans (B.67) il vient :

$$\begin{aligned} \mu_{\tilde{u}^4}(\boldsymbol{\tau}) &= \mu_{e^{\tilde{u}^2}}^2 \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \left(h_{j_1}^{\tilde{u}} h_{j_1+\tau_1}^{\tilde{u}} h_{j_2}^{\tilde{u}} h_{j_2+\tau_2-\tau_3}^{\tilde{u}} + h_{j_1}^{\tilde{u}} h_{j_1+\tau_2}^{\tilde{u}} h_{j_2}^{\tilde{u}} h_{j_2+\tau_1-\tau_3}^{\tilde{u}} + h_{j_1}^{\tilde{u}} h_{j_1+\tau_3}^{\tilde{u}} h_{j_2}^{\tilde{u}} h_{j_2+\tau_2-\tau_1}^{\tilde{u}} \right) \\ &\quad + (\mu_{e^{\tilde{u}^4}} - 3\mu_{e^{\tilde{u}^2}}^2) \sum_{j_1=0}^{\infty} h_{j_1}^{\tilde{u}} h_{j_1+\tau_1}^{\tilde{u}} h_{j_1+\tau_2}^{\tilde{u}} h_{j_1+\tau_3}^{\tilde{u}} \quad (\text{B.69}) \end{aligned}$$

Pour le moment d'ordre six de \tilde{u} le résultat se complique quelque peu :

$$\begin{aligned} \mu_{\tilde{u}^6}(0, \tau, \tau, \tau_1, \tau_2) &= \bar{\mathbb{E}}\{\tilde{u}(t_k)^2 \tilde{u}(t_{k+\tau})^2 \tilde{u}(t_{k+\tau_1}) \tilde{u}(t_{k+\tau_2})\} \\ &= \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} \sum_{i_3=0}^{\infty} \sum_{i_4=0}^{\infty} \sum_{i_5=0}^{\infty} \sum_{i_6=0}^{\infty} h_{i_1}^{\tilde{u}} h_{i_2}^{\tilde{u}} h_{i_3}^{\tilde{u}} h_{i_4}^{\tilde{u}} h_{i_5}^{\tilde{u}} h_{i_6}^{\tilde{u}} \\ &\quad \bar{\mathbb{E}}\{e^{\tilde{u}(t_{k-i_1})} e^{\tilde{u}(t_{k-i_2})} e^{\tilde{u}(t_{k+\tau-i_3})} e^{\tilde{u}(t_{k+\tau-i_4})} e^{\tilde{u}(t_{k+\tau_1-i_5})} e^{\tilde{u}(t_{k+\tau_2-i_6})}\} \quad (\text{B.70}) \end{aligned}$$

Cette dernière espérance dans (B.70) est non-nulle lorsque :

1. il y a trois groupes d'arguments égaux, chacun étant composé de deux arguments (quinze combinaisons – voir le tableau B.1) ;
2. il y a deux groupes d'arguments égaux, un groupe étant composé de deux arguments, l'autre de quatre arguments (quinze combinaisons – voir le tableau B.1) ;
3. tous les arguments sont égaux (une combinaison).

Le moment d'ordre six de \tilde{u} peut donc s'écrire :

$$\begin{aligned} \mu_{\tilde{u}^6}(0, \tau, \tau, \tau_1, \tau_2) &= \mu_{e^{\tilde{u}^2}}^3 \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \sum_{j_3=0}^{\infty} \sum_{\tau_a, \tau_b, \tau_c} h_{j_1}^{\tilde{u}} h_{j_1+\tau_a}^{\tilde{u}} h_{j_2}^{\tilde{u}} h_{j_2+\tau_b}^{\tilde{u}} h_{j_3}^{\tilde{u}} h_{j_3+\tau_c}^{\tilde{u}} \\ &\quad + (\mu_{e^{\tilde{u}^2}} \mu_{e^{\tilde{u}^4}} - 3\mu_{e^{\tilde{u}^2}}^3) \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \sum_{\tau_d, \tau_e, \tau_f, \tau_g} h_{j_1}^{\tilde{u}} h_{j_1+\tau_d}^{\tilde{u}} h_{j_2}^{\tilde{u}} h_{j_2+\tau_e}^{\tilde{u}} h_{j_2+\tau_f}^{\tilde{u}} h_{j_2+\tau_g}^{\tilde{u}} \\ &\quad + (\mu_{e^{\tilde{u}^6}} - 15\mu_{e^{\tilde{u}^2}} \mu_{e^{\tilde{u}^4}} + 30\mu_{e^{\tilde{u}^2}}^3) \sum_{j_1=0}^{\infty} (h_{j_1}^{\tilde{u}})^2 (h_{j_1+\tau}^{\tilde{u}})^2 h_{j_1+\tau_1}^{\tilde{u}} h_{j_1+\tau_2}^{\tilde{u}} \quad (\text{B.71}) \end{aligned}$$

où τ_a, \dots, τ_g prennent des valeurs explicitées dans le tableau B.1, dans lesquels la colonne 'coef.' représente le coefficient multiplicatif du terme considéré.

B.7 Preuve des propositions 2.3.1 et 3.3.1 (tocls, tocils, focls, focils)

La démonstration est effectuée ici pour le modèle EIV à temps continu. Dans le cas du modèle à temps discret, la démonstration est similaire.

TAB. B.1 – Valeurs prises par $\tau_a, \tau_b, \tau_c, \tau_d, \tau_e$ et τ_f dans (B.71)

coef.	τ_a	τ_b	τ_c
1	0	0	$\tau_1 - \tau_2$
2	0	$\tau - \tau_1$	$\tau - \tau_2$
2	τ	τ	$\tau_1 - \tau_2$
2	τ	τ_1	$\tau - \tau_2$
2	τ	τ_2	$\tau - \tau_1$
2	τ_1	τ	$\tau - \tau_2$
2	τ_2	τ	$\tau - \tau_1$
2	τ_1	τ_2	0

coef.	τ_d	τ_e	τ_f	τ_g
1	0	0	$\tau - \tau_1$	$\tau - \tau_2$
2	τ	τ	τ_1	τ_2
2	τ_1	τ	τ	τ_2
2	τ_2	τ	τ	τ_1
2	τ	τ	τ_1	τ_2
1	0	0	τ_1	τ_2
1	τ_1	0	τ	τ_2
1	τ_2	0	τ	τ_1
1	$\tau - \tau_1$	0	τ	τ_2
1	$\tau - \tau_2$	0	τ	τ_1
1	$\tau_1 - \tau_2$	0	τ	τ

Soit $g(t)$ la réponse impulsionnelle du système $G(p, \boldsymbol{\theta})$. À l'aide de la multilinéarité des cumulants, une formule pour leur filtrage peut être trouvée [LAC97, NP93]. Considérons par exemple les cumulants d'ordre trois (croisés) entre les signaux d'entrée/sortie mesurés u et y . Alors :

$$C_{uyu}(\boldsymbol{\tau}) = C_{u_0y_0u_0}(\boldsymbol{\tau}) \quad (\text{B.72})$$

$$= \bar{\mathbb{E}}\{u_o(t)y_o(t + \tau_1)u_o(t + \tau_2)\} \quad (\text{B.73})$$

$$= \bar{\mathbb{E}}\left\{u_o(t) \left(\int_{\mathbb{R}} g(\tau)u_o(t + \tau_1 - \tau)d\tau \right) u_o(t + \tau_2)\right\} \quad (\text{B.74})$$

$$= \int_{\mathbb{R}} g(\tau)\bar{\mathbb{E}}\{u_o(t)u_o(t + \tau_1 - \tau)u_o(t + \tau_2)\} d\tau \quad (\text{B.75})$$

$$= \int_{\mathbb{R}} g(\tau)C_{u_0u_0u_0}(\tau_1 - \tau, \tau_2)d\tau \quad (\text{B.76})$$

$$= G(p, \boldsymbol{\theta})C_{u_0u_0u_0}(\boldsymbol{\tau}) \quad (\text{B.77})$$

$$= G(p, \boldsymbol{\theta})C_{uuu}(\boldsymbol{\tau}) \quad (\text{B.78})$$

ce qui complète la preuve.

Notons par ailleurs que puisque l'intégration dans (B.76) est faite par rapport au premier argument du cumulants, l'opérateur p représente $\frac{\partial}{\partial \tau_1}$ dans la notation compacte (B.77).

Annexe C

Ajouts à la boîte à outils Contsid

Ci-après sont donnés les programmes – écrits pour le logiciel Matlab – des algorithmes développés pour l’estimation paramétrique de modèles EIV à temps continu. Ces estimations sont fondées sur les statistiques d’ordre supérieur (voir la [section 3.3](#) du [chapitre 3](#)). On pourra trouver les méthodes s’appuyant sur les cumulants d’ordre trois (algorithmes `tocls` et `tocils`) et les cumulants d’ordre quatre (algorithmes `focls` et `focils`).

C.1 Algorithme `tocls`

```
function [model] = tocls(data,nn,M,lambda,tol);
% TOCLS -- Computes the continuous-time SISO EIV model parameters
% from regular sampled I/O data by the Least Squares method
% using Third-order cumulants (input of the system must
% have a skewed distribution, the noises a symmetrical distribution).
%
% Usage :
% [model] = tocls(data,nn,M,lambda,tol);
%
% Outputs:
% model : returns the estimated model in an IDPOLY object format
%         along with structure information.
%          $y(t) = [B(s)/F(s)]u(t) + e(t)$ 
%         For the exact format of M see also help IDPOLY.
%
% Inputs:
% data : The estimation data in IDDATA object format.
%        Data are must be equally sampled. See help IDDATA.
% nn    : nn = [nb nf nk] where
%         nb : number of parameters to be estimated for the numerator
%         nf : number of parameters to be estimated for the denominator
%         nk : delay of the model (integer number of sampling period Ts)
% M     : user parameter. See (Thil, Garnier and Gilson, 2008), Automatica (44:3).
%
% Optional inputs :
% lambda : cut-off frequency of the SVF (default value: 3)
% tol    : tolerance of the stop criterion (default value: 1e-4)
%
% author : Stéphane Thil
% date   : February 2007
```

```

% name      : tocls.m
%
% CRAN - Centre de Recherche en Automatique de Nancy

%*** Preliminary calculations ***
timescaling=0;

if nargin<4,error('Not enough input arguments!'),return,end
if nargin<5,tol = 0.0001; end;
if data.ny~=1,error('For SISO systems ny must be one !'),end;
if data.nu~=1,error('For SISO systems nu must be one !'),end;
if data.nu==0,error('No input data '),end;

% Beginning of the algorithm
% -----

Ts = data.Ts;
N = data.N;
u = data.u;
y = data.y;
nf = nn(1);
nb = nn(2);
nk = nn(3);
tau2 = 0;

datay = iddata([],data.y,Ts,'Intersample','foh');
datau = iddata([],data.u,Ts,'Intersample','foh');

% Differentiation of the input/output signals by SVF
ysvf = svf(datay,nf,lambda);
usvf = svf(datau,nf,lambda);

% Estimation of the third-order cumulants
c_uuu = zeros(M,nf+1);
c_uyu = zeros(M,nf+1);
for tau1=0:M-1
    w = N-max(tau1,tau2);
    for der=1:nf+1
        c_uuu(tau1+1,der) = u(1:w)'*(usvf(1+tau1:w+tau1,der).*u(1+tau2:w+tau2))/w;
        c_uyu(tau1+1,der) = u(1:w)'*(ysvf(1+tau1:w+tau1,der).*u(1+tau2:w+tau2))/w;
    end
end

% Estimation using least squares
i = 1;
phi = [-c_uyu(:,2:end) c_uuu(:,end-nb:end)];
Y = c_uyu(:,1);
param = phi\Y;
param = [param(nf+1:end,end)' param(1:nf,end)'];

% End of the algorithm
% -----

% Final model construction
model = par2poly(data,0,nb+1,nf,0,param,'LSHOS',zeros(length(param)),0,0,0);

```

C.2 Algorithmes tocils

```

function [model] = tocils(data,nn,M,maxiter,lambda,tol);
% TOCILS -- Computes the continuous-time SISO EIV model parameters
% from regular sampled I/O data by the Iterative Least
% Squares method using Third-order cumulants (input of the system must
% have a skewed distribution, the noises a symmetrical distribution).
%
% Usage :
%   [model,nbIter] = tocils(data,nn,M,maxiter,lambda,tol);
%
% Outputs:
%   model : returns the estimated model in an IDPOLY object format
%           along with structure information.
%            $y(t) = [B(s)/F(s)]u(t) + e(t)$ 
%           For the exact format of M see also help IDPOLY.
%
% Inputs:
%   data : The estimation data in IDDATA object format.
%           Data are must be equally sampled. See help IDDATA.
%   nn   : nn = [nb nf nk] where
%           nb : number of parameters to be estimated for the numerator
%           nf : number of parameters to be estimated for the denominator
%           nk : delay of the model (integer number of sampling period Ts)
%   M    : user parameter. See (Thil, Garnier and Gilson, 2008) Automatica (44:3).
%
% Optional inputs :
%   maxiter : maximum of loops for the iterations (default value: 20)
%   lambda  : cut-off frequency of the SVF (default value: 3)
%   tol     : tolerance of the stop criterion (default value: 1e-4)
%
% author : Stéphane Thil
% date   : February 2007
% name   : tocils.m
%
% CRAN - Centre de Recherche en Automatique de Nancy

%*** Preliminary calculations ***
timescaling=0;

if nargin<3,error('Not enough input arguments!'),return,end
if nargin<4,maxiter = 20; end;
if nargin<5,lambda = 3; end;
if nargin<6,tol = 0.0001; end;
if data.ny~=1,error('For SISO systems ny must be one !'),end;
if data.nu~=1,error('For SISO systems nu must be one !'),end;
if data.nu==0,error('No input data '),end;

% Beginning of the algorithm
% -----

Ts = data.Ts;
N = data.N;
u = data.u;
y = data.y;
nf = nn(1);
nb = nn(2);
nk = nn(3);
tau2 = 0;

```

```

datay = iddata([],data.y,Ts,'Intersample','foh');
datau = iddata([],data.u,Ts,'Intersample','foh');

% Differentiation of the input/output signals by SVF
ysvf = svf(datay,nf,lambda);
usvf = svf(datau,nf,lambda);

% Estimation of the third-order cumulants
c_uuu = zeros(M,nf+1);
c_uyu = zeros(M,nf+1);
for tau1=0:M-1
    w = N-max(tau1,tau2);
    for der=1:nf+1
        c_uuu(tau1+1,der) = u(1:w)'*(usvf(1+tau1:w+tau1,der).*u(1+tau2:w+tau2))/w;
        c_uyu(tau1+1,der) = u(1:w)'*(ysvf(1+tau1:w+tau1,der).*u(1+tau2:w+tau2))/w;
    end
end

% First estimation using least squares
i = 1;
phi = [-c_uyu(:,2:end) c_uuu(:,end-nb:end)];
Y = c_uyu(:,1);
theta_tocils(:,i) = (phi'*phi)\phi'*Y;

% Model stabilisation
den = mkstable([1 theta_tocils(1:nf,i)']);

% Differentiation and filtering by estimated 1/A(p) using SVF
ysvf = svf(datay,den);
usvf = svf(datau,den);

for tau1=0:M-1
    w = N-max(tau1,tau2);
    for der=1:nf+1
        c_uuu(tau1+1,der) = u(1:w)'*(usvf(1+tau1:w+tau1,der).*u(1+tau2:w+tau2))/w;
        c_uyu(tau1+1,der) = u(1:w)'*(ysvf(1+tau1:w+tau1,der).*u(1+tau2:w+tau2))/w;
    end
end

% Second estimation by using the Least-squares-based SVF method
phi = [-c_uyu(:,2:end) c_uuu(:,end-nb:end)];
Y = c_uyu(:,1);
% old version    theta_tocils(:,i+1)=(phi'*phi)\phi'*Y;
theta_tocils(:,i+1)=phi\Y;

% Iteration
while ( norm(theta_tocils(:,i+1)-theta_tocils(:,i))/norm(theta_tocils(:,i)) > tol )&&(i < maxiter)

    i = i+1;
    % Differentiation and filtering by estimated 1/A(p) using SVF
    den = mkstable([1 theta_tocils(1:nf,i)']);
    ysvf = svf(datay,[1 theta_tocils(1:nf,i)']);
    usvf = svf(datau,[1 theta_tocils(1:nf,i)']);
    % Calculation of the third-order cumulants
    for tau1=0:M-1
        w = N-max(tau1,tau2);
        for der=1:nf+1
            c_uuu(tau1+1,der) = u(1:w)'*(usvf(1+tau1:w+tau1,der).*u(1+tau2:w+tau2))/w;

```

```

        c_uyu(tau1+1,der) = u(1:w)'*(ysvf(1+tau1:w+tau1,der).*u(1+tau2:w+tau2))/w;
    end
end
% Estimation by using the least squares
phi = [-c_uyu(:,2:end) c_uuu(:,end-nb:end)];
Y = c_uyu(:,1);
theta_tocils(:,i+1)=phi\Y;

end
param = [theta_tocils(nf+1:end,end)' theta_tocils(1:nf,end)'];

% End of the algorithm
% -----

% Final model construction
model = par2poly(data,0,nb+1,nf,0,param,'TOCILS',zeros(length(param)),0,0,0);
model.es.Iterations=i;

```

C.3 Algorithmes focls

```

function [model] = focls(data,nn,M,lambda,tol);
% FOCLS -- Computes the continuous-time SISO EIV model parameters
% from regular sampled I/O data by the Least Squares method
% using fourth-order cumulants (input of the system must
% have a non-Gaussian distribution, the noises a Gaussian distribution).
%
% Usage:
% [model] = focls(data,nn,M,lambda,tol);
%
% Outputs:
% model : returns the estimated model in an IDPOLY object format
% along with structure information.
% y(t) = [B(s)/F(s)]u(t) + e(t)
% For the exact format of M see also help IDPOLY.
%
% Inputs:
% data : The estimation data in IDDATA object format.
% Data are must be equally sampled. See help IDDATA.
% nn : nn = [nb nf nk] where
% nb : number of parameters to be estimated for the numerator
% nf : number of parameters to be estimated for the denominator
% nk : delay of the model (integer number of sampling period Ts)
% M : user parameter. See (Thil, Garnier and Gilson, 2008), Automatica (44:3).
%
% Optional inputs :
% lambda : cut-off frequency of the SVF (default value: 3)
% tol : tolerance of the stop criterion (default value: 1e-4)
%
% author : Stéphane Thil
% date : February 2007
% name : focls.m
%
% CRAN - Centre de Recherche en Automatique de Nancy

%*** Preliminary calculations ***
timescaling=0;

```

```

if nargin<4,error('Not enough input arguments!'),return,end
if nargin<5,tol = 0.0001; end;
if data.ny~=1,error('For SISO systems ny must be one !'),end;
if data.nu~=1,error('For SISO systems nu must be one !'),end;
if data.nu==0,error('No input data '),end;

% Beginning of the algorithm
% -----

Ts = data.Ts;
N = data.N;
y = data.y;
u = data.u;
nf = nn(1);
nb = nn(2);
nk = nn(3);
tau2 = 0;
tau3 = 0;

datay = iddata([],data.y,Ts,'Intersample','foh');
datau = iddata([],data.u,Ts,'Intersample','foh');

% Differentiation of the input/output signals by SVF
ysvf = svf(datay,nf,lambda);
usvf = svf(datau,nf,lambda);

% Fourth-order cumulants estimation with k-statistics
c_uuuu = zeros(M,nf+1);
c_uyuu = zeros(M,nf+1);
for tau1=0:M-1
    minn = 1-min(min(0,tau2-tau3),tau3-tau1),tau1-tau2);
    maxx = N-max(max(tau1,tau2),tau3);
    w = maxx-minn+1;
    for der=1:na+1
        c_uuuu(tau1+1,der) =
            u(1:w)'*(usvf(1+tau1:w+tau1,der).*u(1+tau2:w+tau2).*u(1+tau3:w+tau3))*(w+2)/(w*(w-1)) ...
            -( u(1:w)'*usvf(1+tau1:w+tau1,der) * (u(1+tau2:w+tau2)'*u(1+tau3:w+tau3)) )*3/(w*(w-1)) ...
            -( u(1:w)'*u(1+tau2:w+tau2) * (usvf(1+tau1:w+tau1,der)'*u(1+tau3:w+tau3)) )*3/(w*(w-1)) ...
            -( u(1:w)'*u(1+tau3:w+tau3) * (usvf(1+tau1:w+tau1,der)'*u(1+tau2:w+tau2)) )*3/(w*(w-1));
        c_uyuu(tau1+1,der) =
            u(1:w)'*(ysvf(1+tau1:w+tau1,der).*u(1+tau2:w+tau2).*u(1+tau3:w+tau3))*(w+2)/(w*(w-1)) ...
            -( u(1:w)'*ysvf(1+tau1:w+tau1,der) * (u(1+tau2:w+tau2)'*u(1+tau3:w+tau3)) )*3/(w*(w-1)) ...
            -( u(1:w)'*u(1+tau2:w+tau2) * (ysvf(1+tau1:w+tau1,der)'*u(1+tau3:w+tau3)) )*3/(w*(w-1)) ...
            -( u(1:w)'*u(1+tau3:w+tau3) * (ysvf(1+tau1:w+tau1,der)'*u(1+tau2:w+tau2)) )*3/(w*(w-1));
    end
end

% Estimation using least-squares
phi = [-c_uyuu(:,2:end) c_uuuu(:,end-nb:end)];
c_y = c_uyuu(:,1);
param = phi\c_y;
param = [param(nf+1:end,end)' param(1:nf,end)'];

% End of the algorithm
% -----

% Final model construction
model = par2poly(data,0,nb+1,nf,0,param,'FOCLS',zeros(length(param)),0,0,0);

```

C.4 Algorithmes focils

```

function [model] = focils(data,nn,M,maxiter,lambda,tol);
% FOCILS -- Computes the continuous-time SISO EIV model parameters
% from regular sampled I/O data by the Iterative Least
% Squares method using fourth-order cumulants (input of the system must
% have a non-Gaussian distribution, the noises a Gaussian distribution).
%
% Usage :
%   [model,nbIter] = tocils(data,nn,M,maxiter,lambda,tol);
%
% Outputs:
%   model : returns the estimated model in an IDPOLY object format
%           along with structure information.
%            $y(t) = [B(s)/F(s)]u(t) + e(t)$ 
%           For the exact format of M see also help IDPOLY.
%
% Inputs:
%   data : The estimation data in IDDATA object format.
%           Data are must be equally sampled. See help IDDATA.
%   nn   : nn = [nb nf nk] where
%           nb : number of parameters to be estimated for the numerator
%           nf : number of parameters to be estimated for the denominator
%           nk : delay of the model (integer number of sampling period Ts)
%   M    : user parameter. See (Thil, Garnier and Gilson, 2008) Automatica (44:3).
%
% Optional inputs :
%   maxiter : maximum of loops for the iterations (default value: 20)
%   lambda  : cut-off frequency of the SVF (default value: 3)
%   tol     : tolerance of the stop criterion (default value: 1e-4)
%
% author   : Stéphane Thil
% date     : February 2007
% name     : focils.m
%
% CRAN - Centre de Recherche en Automatique de Nancy

%*** Preliminary calculations ***
timescaling=0;

if nargin<3,error('Not enough input arguments!'),return,end
if nargin<4,maxiter = 20; end;
if nargin<5,lambda = 3; end;
if nargin<6,tol = 0.0001; end;
if data.ny~=1,error('For SISO systems ny must be one !'),end;
if data.nu~=1,error('For SISO systems nu must be one !'),end;
if data.nu==0,error('No input data '),end;

% Beginning of the algorithm
% -----

Ts = data.Ts;
N = data.N;
y = data.y;
u = data.u;
nf = nn(1);
nb = nn(2);
nk = nn(3);

```

```

tau2 = 0;
tau3 = 0;

datay = iddata([],data.y,Ts,'Intersample','foh');
datau = iddata([],data.u,Ts,'Intersample','foh');

% Differentiation of the input/output signals by SVF
ysvf = svf(datay,nf,lambda);
usvf = svf(datau,nf,lambda);

% Fourth-order cumulants estimation with k-statistics
c_uuuu = zeros(M,nf+1);
c_uyuu = zeros(M,nf+1);
for tau1=0:M-1
    minn = 1-min(min(min(0,tau2-tau3),tau3-tau1),tau1-tau2);
    maxx = N-max(max(tau1,tau2),tau3);
    w = maxx-minn+1;
    for der=1:na+1
        c_uuuu(tau1+1,der) =
            u(1:w)'*(usvf(1+tau1:w+tau1,der).*u(1+tau2:w+tau2).*u(1+tau3:w+tau3))*(w+2)/(w*(w-1)) ...
            -(u(1:w)'*usvf(1+tau1:w+tau1,der)) * (u(1+tau2:w+tau2)'*u(1+tau3:w+tau3)) )*3/(w*(w-1)) ...
            -(u(1:w)'*u(1+tau2:w+tau2)) * (usvf(1+tau1:w+tau1,der)'*u(1+tau3:w+tau3)) )*3/(w*(w-1)) ...
            -(u(1:w)'*u(1+tau3:w+tau3)) * (usvf(1+tau1:w+tau1,der)'*u(1+tau2:w+tau2)) )*3/(w*(w-1));
        c_uyuu(tau1+1,der) =
            u(1:w)'*(ysvf(1+tau1:w+tau1,der).*u(1+tau2:w+tau2).*u(1+tau3:w+tau3))*(w+2)/(w*(w-1)) ...
            -(u(1:w)'*ysvf(1+tau1:w+tau1,der)) * (u(1+tau2:w+tau2)'*u(1+tau3:w+tau3)) )*3/(w*(w-1)) ...
            -(u(1:w)'*u(1+tau2:w+tau2)) * (ysvf(1+tau1:w+tau1,der)'*u(1+tau3:w+tau3)) )*3/(w*(w-1)) ...
            -(u(1:w)'*u(1+tau3:w+tau3)) * (ysvf(1+tau1:w+tau1,der)'*u(1+tau2:w+tau2)) )*3/(w*(w-1));
    end
end

% First estimation using least squares
i = 1;
phi = [-c_uyuu(:,2:end) c_uuuu(:,end-nb:end)];
c_y = c_uyuu(:,1);
% old version theta_focils(:,i) = (phi'*phi)\phi'*c_y;
theta_focils(:,i) = phi\c_y;

% Model stabilisation
den = mkstable([1 theta_focils(1:nf,i)']);

% Differentiation and filtering by estimated 1/A(p) using SVF
ysvf = svf(datay,den);
usvf = svf(datau,den);

% Fourth-order cumulants estimation with k-statistics
for tau1=0:M-1
    minn = 1-min(min(min(0,tau2-tau3),tau3-tau1),tau1-tau2);
    maxx = N-max(max(tau1,tau2),tau3);
    w = maxx-minn+1;
    for der=1:na+1
        c_uuuu(tau1+1,der) =
            u(1:w)'*(usvf(1+tau1:w+tau1,der).*u(1+tau2:w+tau2).*u(1+tau3:w+tau3))*(w+2)/(w*(w-1)) ...
            -(u(1:w)'*usvf(1+tau1:w+tau1,der)) * (u(1+tau2:w+tau2)'*u(1+tau3:w+tau3)) )*3/(w*(w-1)) ...
            -(u(1:w)'*u(1+tau2:w+tau2)) * (usvf(1+tau1:w+tau1,der)'*u(1+tau3:w+tau3)) )*3/(w*(w-1)) ...
            -(u(1:w)'*u(1+tau3:w+tau3)) * (usvf(1+tau1:w+tau1,der)'*u(1+tau2:w+tau2)) )*3/(w*(w-1));
        c_uyuu(tau1+1,der) =
            u(1:w)'*(ysvf(1+tau1:w+tau1,der).*u(1+tau2:w+tau2).*u(1+tau3:w+tau3))*(w+2)/(w*(w-1)) ...
            -(u(1:w)'*ysvf(1+tau1:w+tau1,der)) * (u(1+tau2:w+tau2)'*u(1+tau3:w+tau3)) )*3/(w*(w-1)) ...

```

```

        -( (u(1:w)'*u(1+tau2:w+tau2)) * (ysvf(1+tau1:w+tau1,der)'*u(1+tau3:w+tau3)) )*3/(w*(w-1)) ...
        -( (u(1:w)'*u(1+tau3:w+tau3)) * (ysvf(1+tau1:w+tau1,der)'*u(1+tau2:w+tau2)) )*3/(w*(w-1));
    end
end

% Second estimation by using the Least-squares-based SVF method
phi = [-c_uyuu(:,2:end) c_uuuu(:,end-nb:end)];
c_y = c_uyuu(:,1);
% old version theta_focils(:,i) = (phi'*phi)\phi'*c_y;
theta_focils(:,i) = phi\c_y;

% Iteration
while ( norm(theta_focils(:,i+1)-theta_focils(:,i))/norm(theta_focils(:,i)) > tol )&&(i < maxiter)

    i = i+1;
    den = mksttable([1 theta_focils(1:nf,i)']);
    % Differentiation and filtering by estimated 1/A(p) using SVF
    ysvf = svf(datay,den);
    usvf = svf(datau,den);
    % Fourth-order cumulants estimation with k-statistics
    for tau1=0:M-1
        minn = 1-min(min(0,tau2-tau3),tau3-tau1),tau1-tau2);
        maxx = N-max(max(tau1,tau2),tau3);
        w = maxx-minn+1;
        for der=1:na+1
            c_uuuu(tau1+1,der) =
                u(1:w)'*(usvf(1+tau1:w+tau1,der).*u(1+tau2:w+tau2).*u(1+tau3:w+tau3))*(w+2)/(w*(w-1)) ...
                -( (u(1:w)'*usvf(1+tau1:w+tau1,der)) * (u(1+tau2:w+tau2)'*u(1+tau3:w+tau3)) )*3/(w*(w-1)) ...
                -( (u(1:w)'*u(1+tau2:w+tau2)) * (usvf(1+tau1:w+tau1,der)'*u(1+tau3:w+tau3)) )*3/(w*(w-1)) ...
                -( (u(1:w)'*u(1+tau3:w+tau3)) * (usvf(1+tau1:w+tau1,der)'*u(1+tau2:w+tau2)) )*3/(w*(w-1));
            c_uyuu(tau1+1,der) =
                u(1:w)'*(ysvf(1+tau1:w+tau1,der).*u(1+tau2:w+tau2).*u(1+tau3:w+tau3))*(w+2)/(w*(w-1)) ...
                -( (u(1:w)'*ysvf(1+tau1:w+tau1,der)) * (u(1+tau2:w+tau2)'*u(1+tau3:w+tau3)) )*3/(w*(w-1)) ...
                -( (u(1:w)'*u(1+tau2:w+tau2)) * (ysvf(1+tau1:w+tau1,der)'*u(1+tau3:w+tau3)) )*3/(w*(w-1)) ...
                -( (u(1:w)'*u(1+tau3:w+tau3)) * (ysvf(1+tau1:w+tau1,der)'*u(1+tau2:w+tau2)) )*3/(w*(w-1));
        end
    end
    % Estimation by using the least squares
    phi = [-c_uyuu(:,2:end) c_uuuu(:,end-nb:end)];
    c_y = c_uyuu(:,1);
    theta_focils(:,i) = phi\c_y;

end

param = [theta_focils(nf+1:end,end)' theta_focils(1:nf,end)'];

% End of the algorithm
% -----

% Final model construction
model = par2poly(data,0,nb+1,nf,0,param,'FOCILS',zeros(length(param)),0,0,0);
model.es.Iterations=i;

```


Bibliographie

- [AD84] B.D.O. ANDERSON et M. DEISTLER. Identifiability in dynamic errors-in-variables models. *Journal of Time Series Analysis*, 5(1) :1–13, 1984.
- [AD87] B.D.O. ANDERSON et M. DEISTLER. Dynamic errors-in-variables systems with three variables. *Automatica*, 23(5) :611–616, septembre 1987.
- [Adc77] R.J. ADCOCK. Note on the methods of least squares. *The Analyst*, 4(6) :183–184, 1877.
- [Adc78] R.J. ADCOCK. A problem in least squares. *The Analyst*, 5(2) :53–54, 1878.
- [AE97] J.M.M. ANDERSON et W. EDMONDSON. System identification with noisy input/output data using cumulant-based Steiglitz-McBride algorithm. *IEEE Transactions on Circuits and Systems II : Analog and Digital Signal Processing*, 44(5) :1021–1024, février 1997.
- [AG96] J.M.M. ANDERSON et G.B. GIANNAKIS. Noisy input/output system identification using cumulants and the Steiglitz-McBride algorithm. *IEEE Transactions on Signal Processing*, 44(4) :1021–1024, avril 1996.
- [AG06] J.C. AGÜERO et G.C. GOODWIN. Identifiability of errors-in-variables dynamic systems. Dans *14th IFAC Symposium on System Identification*, pages 196–201, Newcastle, Australia, mars 2006.
- [AG08] J.C. AGÜERO et G.C. GOODWIN. Identifiability of errors-in-variables dynamic systems. *Automatica*, 44(2), février 2008.
- [ÅHS84] K.J. ÅSTRÖM, P. HAGANDER, et J. STERNBY. Zeros of sampled systems. *Automatica*, 20(1) :31–38, janvier 1984.
- [Aka66] H. AKAIKE. On the use of non-Gaussian process in the identification of a linear dynamic system. *Annals of the Institute of Statistical Mathematics*, 18(3) :269–276, 1966.
- [AM79] B.D.O. ANDERSON et J.B. MOORE. *Optimal Filtering*. Prentice Hall, Englewood Cliffs, 1979.
- [And85] B.D.O. ANDERSON. Identification of scalar errors-in-variables models with dynamics. *Automatica*, 21 :709–716, novembre 1985.

- [BCS97] S. BEGHELLI, P. CASTALDI, et U. SOVERINI. A frequential approach for errors-in-variables models. Dans *European Control Conference (ECC'97)*, Brussels, Belgium, juillet 1997.
- [BGS90] S. BEGHELLI, R.P. GUIDORZI, et U. SOVERINI. The Frisch scheme in dynamic system identification. *Automatica*, 26(1) :171–176, janvier 1990.
- [Bol96] L. BOLTZMANN. *Vorlesungen über Gastheorie*. J.A. Barth, Leipzig, 1896.
- [BP90] P. BONDON et B. PICINBONO. De la blancheur et de ses transformations. *Traitement du signal*, 7(5) :385–395, 1990.
- [Bri81] D.R. BRILLINGER. *Time Series, Data Analysis and Theory*. Holden Day, San Fransisco, 1981.
- [Bri01] D.R. BRILLINGER. *Time Series, Data Analysis and Theory*. Society for Industrial and Applied Mathematics, Philadelphia, 2001.
- [CC94] J.-M. CHEN et B.-S. CHEN. A higher-order correlation method for model-order and parameter estimation. *Automatica*, 30(8) :1339–1344, août 1994.
- [CS96a] P. CASTALDI et U. SOVERINI. Identification of dynamic errors-in-variables models. *Automatica*, 32(4) :631–636, avril 1996.
- [CS96b] M. CERDEVALL et P. STOICA. System identification from noisy measurements by using instrumental variables and subspace fitting. *Circuits, Systems and Signal Processing*, 15(2) :275–290, mars 1996.
- [CV97] C.T. CHOU et M. VERHAEGEN. Subspace algorithms for the identification of multivariable dynamic errors-in-variables models. *Automatica*, 33(10) :1857–1869, octobre 1997.
- [CVJ99] C.T. CHOU, M. VERHAEGEN, et R. JOHANSSON. Continuous-time identification of SISO systems using Laguerre functions. *IEEE Transactions on Signal Processing*, 47(2) :349–362, février 1999.
- [DA89] M. DEISTLER et B.D.O. ANDERSON. Linear dynamic errors-in-variables models. Some structure theory. *Journal of Econometrics*, 41 :39–63, mai 1989.
- [Dei86] M. DEISTLER. Linear errors-in-variables models. Dans S. BITTANTI, éditeur, *Time Series and Linear Systems*, volume 86 de *Lecture Notes in Control and Information Sciences*, pages 39–63. Springer-Verlag, Berlin, 1986.
- [DG94] A. DELOPOULOS et G.B. GIANNAKIS. Consistent identification of stochastic linear systems with noisy input-output data. *Automatica*, 30(8) :1271–1294, août 1994.
- [DGS03] R. DIVERSI, R. GUIDORZI, et U. SOVERINI. A new criterion in EIV identification and filtering applications. Dans *13th IFAC Symposium on System Identification*, pages 1993–1998, Rotterdam, The Netherlands, août 2003.

- [dM01] M. DE MATHELIN. Panorama des algorithmes récursifs d'estimation paramétrique. Dans R. LOZANO et D. TAOUTAOU, éditeurs, *Identification et Commande Adaptative*, pages 101–160, Paris, 2001. Hermes Science.
- [EHS06] M. EKMAN, M. HONG, et T. SÖDERSTRÖM. A separable nonlinear least squares approach for identification of linear systems with errors in variables. Dans *14th IFAC Symposium on System Identification*, Newcastle, Australia, mars 2006.
- [Ekm05] M. EKMAN. Identification of linear systems with errors in variables using separable nonlinear least squares. Dans *16th IFAC World Congress on Automatic Control*, pages 178–183, Prague, Czech Republic, juillet 2005.
- [FGM99] U. FORSSELL, F. GUSTAFSSON, et T. MCKELVEY. Time-domain identification of dynamic errors-in-variables systems using periodic excitation signals. Dans *14th IFAC World Congress on Automatic Control*, Beijing, China, juillet 1999.
- [FN85] K.V. FERNANDO et H. NICHOLSON. Identification of linear systems with input and output noise : the Koopmans-Levin method. *IEE Proceedings, Part D*, 132(1) :30–36, 1985.
- [FP90] B. FRIEDLANDER et B. PORAT. Asymptotically optimal estimation of MA and ARMA parameters of non-Gaussian processes from high-order moments. *IEEE Transactions on Automatic Control*, 35(1) :27–35, janvier 1990.
- [Fri34] R. FRISCH. Statistical confluence analysis by means of complete regression systems. Technical Report 5, University of Oslo, Economics Institute, Oslo, Norway, 1934.
- [Gea42] R.C. GEARY. Inherent relations between random variables. *Proceedings of the Royal Irish Academy*, A47 :63–76, 1942.
- [Gia90] G.B. GIANNAKIS. On the identifiability of non-gaussian ARMA models using cumulants. *IEEE Transactions on Automatic Control*, 35(1) :18–26, janvier 1990.
- [GL80] G.H. GOLUB et C.F. Van LOAN. An analysis of the total least squares problem. *SIAM Journal on Numerical Analysis*, 17(6) :883–893, décembre 1980.
- [GM89] G.B. GIANNAKIS et J.M. MENDEL. Identification of non-minimum phase systems using higher-order statistics. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(3) :360–377, mars 1989.
- [GMR03] H. GARNIER, M. MENSLER, et A. RICHARD. Continuous-time model identification from sampled data. Implementation issues and performance evaluation. *International Journal of Control*, 76(13) :1337–1357, septembre 2003.
- [GPS95] P. GUILLAUME, R. PINTELON, et J. SCHOUKENS. Robust parametric transfer function estimation using complex logarithmic frequency response data. *IEEE Transactions on Automatic Control*, 40(7) :1180–1190, juillet 1995.

- [GVdH01] M. GILSON et P. VAN DEN HOF. On the relation between a bias-eliminated least-squares (BELS) and an IV estimator in closed-loop identification. *Automatica*, 37(10) :1593–1600, octobre 2001.
- [GVdH05] M. GILSON et P. VAN DEN HOF. Instrumental variable methods for closed-loop system identification. *Automatica*, 41(2) :241–249, février 2005.
- [GYTG07] H. GARNIER, P.C. YOUNG, S. THIL, et M. GILSON. Data-based continuous-time modelling of environmental systems. Dans *STIC & environnement*, Lyon, France, novembre 2007.
- [HG02] E. HUSELSTEIN et H. GARNIER. An approach to continuous-time model identification from non-uniformly sampled data. Dans *41st IEEE Conference on Decision and Control*, Las Vegas, Nevada (USA), décembre 2002.
- [HSZ06a] M. HONG, T. SÖDERSTRÖM, et W.X. ZHENG. Accuracy analysis of bias-eliminating least squares for errors-in-variables identification. Dans *14th IFAC Symposium on System Identification*, pages 190–195, Newcastle, Australia, mars 2006.
- [HSZ06b] M. HONG, T. SÖDERSTRÖM, et W.X. ZHENG. A simplified form of the bias-eliminating least squares method for errors-in-variables identification. *Technical report 2006-040*, Department of Information Technology, Uppsala University, 2006.
- [HSZ07] M. HONG, T. SÖDERSTRÖM, et W.X. ZHENG. Accuracy analysis of bias-eliminating least squares estimates for errors-in-variables identification. *Automatica*, 43(9) :1590–1596, septembre 2007.
- [IKYW05] M. IKENOUE, S. KANAE, Z.-J. YANG, et K. WADA. Identification of noisy input-output system using bias-compensated least squares method. Dans *16th IFAC World Congress on Automatic Control*, Prague, Czech Republic, juillet 2005.
- [IS94] Y. INOUE et Y. SUGA. Identification of linear systems with noisy input using input-output cumulants. *International Journal of Control*, 59(5) :1231–1253, mai 1994.
- [IT91] Y. INOUE et H. TSUCHIYA. Identification of linear systems using input-output cumulants. *International Journal of Control*, 53(6) :1431–1448, juin 1991.
- [JIJW01] L.-J. JIA, M. IKENOUE, C.-Z. JIN, et K. WADA. On bias compensated least squares method for noisy input-output system identification. Dans *40th IEEE Conference on Decision and Control*, pages 3332–3337, Orlando, Florida, USA, décembre 2001.
- [JJYW02] L.-J. JIA, C.-Z. JIN, Z.-J. YANG, et K. WADA. On the relationship between BELS and IV methods. Dans *41st IEEE Conference on Decision and Control*, pages 3440–3445, Las Vegas, Nevada, USA, décembre 2002.

- [Joh93] R. JOHANSSON. *System Modeling and Identification*. Prentice Hall, Englewood Cliffs, 1993.
- [Koo36] T.C. KOOPMANS. *Linear regression analysis of economic time series*. Thèse de doctorat, University of Leiden, 1936.
- [KSS00] E. KARLSSON, T. SÖDERSTRÖM, et P. STOICA. The Cramer-Rao lower bound for noisy input-output systems. *Signal Processing*, 80(11) :2421–2447, novembre 2000.
- [LAC97] J.-L. LACOUME, P.-O. AMBLARD, et P. COMON. *Statistiques d'ordre Supérieur pour le Traitement du Signal*. Masson, Paris, 1997.
- [Lev64] M. LEVIN. Estimation of a system pulse transfer function in the presence of noise. *IEEE Transactions on Automatic Control*, 9(3) :229–235, juillet 1964.
- [Lju87] L. LJUNG. *System Identification : Theory for the User*. Prentice Hall, Englewood Cliffs, 1987.
- [Lju99] L. LJUNG. *System Identification : Theory for the User*. Prentice Hall, Englewood Cliffs, 2nd édition, 1999.
- [Lju03] L. LJUNG. Initialisation aspects for subspace and output–error identification methods. Dans *European Control Conference (ECC'03)*, Cambridge, UK, 2003.
- [Mah07] K. MAHATA. An improved bias-compensation approach for errors-in-variables model identification. *Automatica*, 43(8) :1339–1354, août 2007.
- [MBR05] S. MOUSSAOUI, D. BRIE, et A. RICHARD. Regularization aspects in continuous-time model identification. *Automatica*, 41(2) :197–208, février 2005.
- [McC87] P. MCCULLAGH. *Tensor Methods in Statistics*. Monographs on Statistics and Applied Probability. Chapman and Hall, London, 1987.
- [Men91] J.M. MENDEL. Tutorial on high-order statistics (spectra) in signal processing and system theory : theoretical results and some applications. *Proceedings of the IEEE*, 79(3) :278–305, mars 1991.
- [Men99] M. MENSLER. *Analyse et étude comparative de méthodes d'identification des systèmes à représentation continue. Développement d'une boîte à outils logicielle*. Thèse de doctorat, Université Henri Poincaré, Nancy 1, 1999.
- [MG05] K. MAHATA et H. GARNIER. Direct identification of continuous-time errors-in-variables models. Dans *16th IFAC World Congress on Automatic Control*, Prague, Czech Republic, juillet 2005.
- [MG06] K. MAHATA et H. GARNIER. Identification of continuous-time errors-in-variables models. *Automatica*, 49(9) :1470–1490, septembre 2006.
- [MKVH06] I. MARKOVSKY, A. KUKUSH, et S. VAN HUFFEL. On errors-in-variables estimation with unknown noise variance ratio. Dans *14th IFAC Symposium on System Identification*, Newcastle, Australia, mars 2006.

- [NP88] C.L. NIKIAS et R. PAN. Time-delay estimation in unknown Gaussian spatially correlated noise. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 36 :1706–1714, novembre 1988.
- [NP93] C.L. NIKIAS et A.P. PETROPULU. *Higher-order Spectra Analysis*. Signal Processing Series. Prentice Hall, Englewood Cliffs, 1993.
- [PF89] B. PORAT et B. FRIEDLANDER. Performance analysis of parameter estimation algorithms based on high-order moments. *International Journal of Adaptive Control and Signal Processing*, 3 :191–229, septembre 1989.
- [PS01] R. PINTELON et J. SCHOUKENS. *System Identification : A Frequency Domain Approach*. IEEE Press, New york, 2001.
- [PW01] L. PRONZATO et E. WALTER. Eliminating suboptimal local minimizers in non-linear estimation. *Technometrics*, 43(4) :434–442, 2001.
- [PW02] L. PRONZATO et E. WALTER. Comment éviter les optimiseurs locaux parasites en estimation paramétrique non linéaire. *Journal Européen des Systèmes Automatisés*, 36(3) :351–362, 2002.
- [Rei41] O. REIERSØL. Confluence analysis by means of lag moments and other methods of confluence analysis. *Econometrica*, 9(1) :1–24, janvier 1941.
- [Rei50] O. REIERSØL. Identifiability of a linear relation between variables that are subject to error. *Econometrica*, 18(4) :375–389, octobre 1950.
- [RN65] M. ROSENBLATT et J.W. Van NESS. Estimation of the bispectrum. *Annals of Mathematical Statistics*, pages 420–436, 1965.
- [SCE95] P. STOICA, M. CERDEVALL, et A. ERIKSSON. Combined instrumental variable and subspace fitting approach to estimation of noisy input-output systems. *IEEE Transactions on Signal Processing*, 43(10) :2386–2397, octobre 1995.
- [SD98] W. SCHERRER et M. DEISTLER. A structure theory for linear errors-in-variables models. *SIAM Journal on Control and Optimization*, 36(6) :2148–2175, novembre 1998.
- [SH05] T. SÖDERSTRÖM et M. HONG. Identification of dynamic errors-in-variables systems with periodic data. Dans *16th IFAC World Congress on Automatic Control*, Prague, Czech Republic, juillet 2005.
- [SHZ05] T. SÖDERSTRÖM, M. HONG, et W.X. ZHENG. Convergence properties of bias-eliminating algorithms for errors-in-variable identification. *International Journal of Adaptive Control and Signal Processing*, 19 :703–722, 2005.
- [SLMM06] T. SÖDERSTRÖM, E. LARSSON, K. MAHATA, et M. MOSSBERG. Using continuous-time modeling for errors-in-variables identification. Dans *14th IFAC Symposium on System Identification*, pages 428–433, Newcastle, Australia, mars 2006.

- [SM65] K. STEIGLITZ et L.E. MCBRIDE. A technique for the identification of linear systems. *IEEE Transactions on Automatic Control*, 43(10) :2386–2397, octobre 1965.
- [SM02] T. SÖDERSTRÖM et K. MAHATA. On instrumental variable and total least squares approaches for identification of noisy systems. *International Journal of Control*, 75(6) :381–389, avril 2002.
- [SMS03] T. SÖDERSTRÖM, K. MAHATA, et U. SOVERINI. Identification of dynamic errors-in-variables models : approaches based on two-dimensional ARMA modeling of the data. *Automatica*, 39(5) :929–935, mai 2003.
- [SN87] P. STOICA et A. NEHORAI. On the uniqueness of prediction error models for systems with noisy input-output data. *Automatica*, 23(4) :541–543, juillet 1987.
- [SO94] A. STUART et J.K. ORD. *Kendall's Advanced Theory of Statistics*, volume 1. Edward Arnold, London, 6 édition, 1994.
- [Söd81] T. SÖDERSTRÖM. Identification of stochastic linear systems in presence of input noise. *Automatica*, 17(5) :713–725, septembre 1981.
- [Söd02] T. SÖDERSTRÖM. *Discrete-time Stochastic Systems – Estimation and Control*. Springer-Verlag, London, 2nd édition, 2002.
- [Söd03] T. SÖDERSTRÖM. Why are errors-in-variables problems often tricky ? Dans *European Control Conference (ECC'03)*, University of Cambridge, UK, septembre 2003.
- [Söd05] T. SÖDERSTRÖM. Accuracy analysis of the Frisch estimates for identifying errors-in-variables systems. Dans *44th IEEE Conference on Decision and Control*, Seville, Spain, décembre 2005.
- [Söd06] T. SÖDERSTRÖM. On computing the Cramer-Rao bound and covariance matrices for PEM estimates in linear state space models. Dans *14th IFAC Symposium on System Identification*, pages 600–605, Newcastle, Australia, mars 2006.
- [Söd07] T. SÖDERSTRÖM. Errors-in-variables methods in system identification. *Automatica*, 43(6) :939–958, juin 2007.
- [Söd08] T. SÖDERSTRÖM. Extending the Frisch scheme for errors-in-variables identification to correlated output noise. *International Journal of Adaptive Control and Signal Processing*, 22(1) :55–73, février 2008.
- [Sol86] V. SOLO. Identifiability of time series models with errors in variables. *Journal of Applied Probability*, 23 A :63–71, juillet 1986.
- [SP91] J. SCHOUKENS et R. PINTELON. *Identification of Linear Systems : a Practical Guideline to Accurate Modeling*. Pergamon, Oxford, 1991.

- [SPVG97] J. SCHOUKENS, R. PINTELON, G. VANDERSTEEN, et P. GUILLAUME. Frequency-domain system identification using non-parametric noise models estimated from a small number of data sets. *Automatica*, 30(6) :1073–1086, juillet 1997.
- [SS81a] T. SÖDERSTRÖM et P. STOICA. Comparison of some instrumental variable methods – consistency and accuracy aspects. *Automatica*, 17(1) :101–115, janvier 1981.
- [SS81b] P. STOICA et T. SÖDERSTRÖM. The Steiglitz-McBride identification algorithm revisited – convergence analysis and accuracy aspects. *IEEE Transactions on Automatic Control*, 26(3) :712–717, juin 1981.
- [SS83] T. SÖDERSTRÖM et P. STOICA. *Instrumental Variable Methods for System Identification*. Springer, Berlin, 1983.
- [SS89] T. SÖDERSTRÖM et P. STOICA. *System identification*. Series in Systems and Control Engineering. Prentice Hall, New York, 1989.
- [SYW91] S. SAGARA, Z.-J. YANG, et K. WADA. Identification of continuous systems from noisy sampled input-output data. Dans *IFAC Symposium on System Identification and Parameter Estimation*, pages 1922–1627, Budapest, Hungary, 1991.
- [SZS99] T. SÖDERSTRÖM, W.X. ZHENG, et P. STOICA. Comments on "On a least-squares-based algorithm for identification of stochastic linear systems". *IEEE Transactions on Signal Processing*, 47(5) :1395–1396, 1999.
- [TGG05] S. THIL, M. GILSON, et H. GARNIER. Méthodes de compensation de biais pour l'identification de modèles erreurs en les variables. Dans *JDMACS 2005*, Lyon, France, septembre 2005.
- [TGG08a] S. THIL, H. GARNIER, et M. GILSON. Third-order cumulants based methods for continuous-time errors-in-variables model identification. *Automatica*, 44(3) :647–658, mars 2008.
- [TGG08b] S. THIL, M. GILSON, et H. GARNIER. On instrumental variable methods for errors-in-variables model identification. Dans *17th IFAC World Congress*, Seoul, Korea, 2008.
- [TGGM07] S. THIL, H. GARNIER, M. GILSON, et K. MAHATA. Continuous-time model identification from noisy input/output measurements using fourth-order cumulants. Dans *46th IEEE Conference on Decision and Control (CDC'2007)*, New Orleans, LA, USA, décembre 2007.
- [Tho05] M. THOMASSIN. *Estimation de retard dans des conditions expérimentales passives. Application à l'identification d'un bief de rivière*. Thèse de doctorat, Université Henri Poincaré, Nancy 1, 2005.
- [THS⁺08] S. THIL, M. HONG, T. SÖDERTSTRÖM, M. GILSON, et H. GARNIER. Statistical analysis of a third-order cumulants based algorithm for discrete-time errors-in-variables identification. Dans *17th IFAC World Congress*, Seoul, Korea, 2008.

- [TL99] J.K. TUGNAIT et E. LIU. Model validation and order selection for linear model fitting using third- and fourth-order cumulants. *IEEE Transactions on Signal Processing*, 47(9) :2433–2443, septembre 1999.
- [Tug91] J.K. TUGNAIT. On time delay estimation with unknown spatially correlated Gaussian noise using fourth order cumulants and cross cumulants. *IEEE Transactions on Signal Processing*, pages 1258–1267, juin 1991.
- [Tug92] J.K. TUGNAIT. Stochastic system identification with noisy input using cumulant statistics. *IEEE Transactions on Automatic Control*, 37(4) :476–485, avril 1992.
- [Tug94] J.K. TUGNAIT. Linear model validation and order selection using higher order statistics. *IEEE Transactions on Signal Processing*, 42(7) :1728–1736, juillet 1994.
- [Tug95] J.K. TUGNAIT. An improved test for linear model validation and order selection using higher order statistics. *IEEE Signal Processing Letters*, 2(6) :123–125, juin 1995.
- [TY95] J.K. TUGNAIT et Y. YE. Stochastic system identification with noisy input-output measurements using polyspectra. *IEEE Transactions on Automatic Control*, 40(4) :670–683, avril 1995.
- [UR90] H. UNBEHAUEN et G.P. RAO. Continuous-time approaches to system identification - a survey. *Automatica*, 26(1) :23–35, janvier 1990.
- [VdH96] P. VAN DEN HOF. System identification. *Technical Report*, 1996.
- [VH02] S. VAN HUFFEL. *Total least squares and errors-in-variables modeling*. Kluwer Academic Publishers, Dordrecht, 2002.
- [WFS98] K.-P. WONG, D. FENG, et W.-C. SIU. Generalized linear least squares algorithm for non-uniformly sampled biomedical system identification with possible repeated eigenvalues. *Computer methods and programs in biomedicine*, 57(3) :167–177, novembre 1998.
- [WP94] E. WALTER et L. PRONZATO. *Identification de Modèles Paramétriques à partir de Données Expérimentales*. Masson, Paris, 1994.
- [YGG08] P.C. YOUNG, H. GARNIER, et M. GILSON. Refined instrumental variable identification of continuous-time hybrid Box-Jenkins models. Dans H. GARNIER et L. WANG, éditeurs, *Identification of Continuous-time Models from Sampled Data*. Springer-Verlag, London, 2008.
- [YJ80] P.C. YOUNG et A.J. JAKEMAN. Refined instrumental variable methods of time-series analysis : Part III, extensions. *International Journal of Control*, 31(4) :741–764, avril 1980.
- [You64] P.C. YOUNG. In flight dynamic checkout - a discussion. *IEEE Transactions on Aerospace*, 2 :1106–1111, 1964.

- [You70] P.C. YOUNG. An instrumental variable method for real-time identification of a noisy process. *Automatica*, 6(2) :271–287, mars 1970.
- [You81] P.C. YOUNG. Parameter estimation for continuous-time models - a survey. *Automatica*, 17(1) :23–39, janvier 1981.
- [You84] P.C. YOUNG. *Recursive Estimation and Time Series Analysis*. Springer-Verlag, Berlin, 1984.
- [YSW93] Z.-J. YANG, S. SAGARA, et K. WADA. Identification of continuous-time systems from sampled input-output data using bias eliminating techniques. *Control Theory and Advanced Technology*, 9(1) :53–75, mars 1993.
- [Zhe98] W.X. ZHENG. Transfer function estimation from noisy input and output data. *International Journal of Adaptive Control and Signal Processing*, 12(4) :365–380, juin 1998.
- [Zhe99] W.X. ZHENG. On least-squares identification of stochastic linear systems with noisy input-output data. *International Journal of Adaptive Control and Signal Processing*, 13(3) :131–143, mai 1999.
- [Zhe00] W.X. ZHENG. Unbiased identification of stochastic linear systems from noisy input and output measurements. Dans *39th IEEE Conference on Decision and Control*, volume 3, pages 2710–2715, Sydney, Australia, décembre 2000.
- [Zhe02] W.X. ZHENG. A bias correction method for identification of linear dynamic errors-in-variables models. *IEEE Transactions on Automatic Control*, 47(7) :1142–1147, juillet 2002.