



Réalisation et identification de systèmes bilinéaires homogènes

Gibran Etcheverry

► To cite this version:

Gibran Etcheverry. Réalisation et identification de systèmes bilinéaires homogènes. Automatique / Robotique. Université Paul Sabatier - Toulouse III, 2006. Français. NNT : . tel-00130071

HAL Id: tel-00130071

<https://theses.hal.science/tel-00130071>

Submitted on 9 Feb 2007

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

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

**Laboratoire d'Analyse et d'Architecture des
Systèmes du CNRS**

Ecole Doctorale Systèmes

Spécialité : Systèmes Automatiques

Par Gibran ETCHEVERRY DOGER

**Réalisation et Identification de Systèmes Bilinéaires
Homogènes**

Directeur de Thèse : André MONIN

Mourir..., mourir pour se réveiller au lever du soleil. Grandir, croire dans l'indubitable, dans l'infinitude de la mort et nos pensées, car la renaissance est éternelle.

Renaître dans la peau du temps, maîtriser la lumière dans l'espace et rouler en jouant avec l'âme qui jamais existera. Une goutte d'eau égarée dans l'univers sans couleur ni odeur, pleine d'oubli et mystères, d'échos sourds et regards profonds, de regards sans couleur.

Est-ce ça la vie, ou simplement le jeu de l'illusion qu'en courant dans mes veines m'invente et me fait croire..., croire que demain j'existerai pour regarder vers l'horizon en rêvant de la frontière entre la chimère et l'éternel?

Non, non, je ne suis que l'illusion de l'existence devenue une image encrée dans l'espace par la lumière, me créant ainsi moi, l'être sans regard et sans couleur, celui qui déambule un instant et s'en va, piégé dans un éclat inattendu et errant par là... être de matière une seconde parce qu'une loi absurde l'a voulu, moi, celui-là c'est moi !

GED

Remerciements

Je remercie M. André MONIN de m'avoir accueilli au sein du groupe RST (MRS) au LAAS et de m'avoir proposé ce sujet de thèse, ainsi que de sa patience et du dernier coup de pouce. Je remercie aussi M. Gérard FAVIER et M. Mohammed M'SAAD d'avoir bien voulu être rapporteurs pour la soutenance de cette thèse. Je remercie M. Jean-Louis CALVET d'avoir accepté d'être président du jury.

Je remercie le Conseil National de Science et Technologie au Mexique (CONACyT) pour son exceptionnel soutien financier, lequel m'a permis de réaliser ce travail.

Je remercie la famille Félix de Nîmes : Marie, Edgard, et Mireille, sans l'aide de qui je n'aurais pas été si bien accueilli à mon arrivée en France et encore moins en mesure de trouver un logement. Je remercie Guilhèm ZAMORE pour avoir eu la patience de lire mon manuscrit et de m'apporter les réflexions nécessaires pour l'améliorer, ainsi que nos discussions sur le cours du monde : peut être un jour le présent sera le futur caché d'une meilleure promesse.

Je remercie Citlalih et Sergio pour son écoute et sa fraternité, Ali pour ses conseils et soutien, et Joan pour son soutien dans le partage de la complexe situation humaine-académique. Je remercie Efrain, Maggie, et Gustavo, d'avoir partagé l'amitié pendant ce temps et les discussions sur les diverses pierres angulaires de la philosophie, la cuisine, le Mexique, et la France.

Je remercie Héctor, Diana, Neftali, Judith, Changuis, Hélène, Israel, et Mathilde, pour les rencontres ici et là bas, dans la joie et la réflexion, ainsi que dans la pénurie et la douleur dans laquelle vous êtes arrivés pour faire la lumière. Je remercie aussi Victor, Alfredo, Omar, Rabi, Benjy, Ricardo, et Angel, pour ses lettres et son soutien.

Je remercie mon père, parce que ses réflexions dans la vie, une fois enregistrés en étant petit, aujourd'hui font surface pour mieux repérer les nuances dans le monde, si riche et si complexe, si simple et si caché sous les couvertures humaines.

A ma mère..., nous nous verrons bientôt pour dialoguer ensemble avec la terre, l'entendre trembler, et rire ainsi de tout et de tous comme dans l'enfance de nos rêves ; celle qui me ramène dans tes bras au cours de la nuit, celle dans laquelle tu étais toujours là pour secourir mes chutes, celle dans laquelle on construisait simplement le temps en le scellant d'espoir.

De l'autre coté du monde il y a une diversité d'idées et d'espaces multiformes dans la rencontre de l'autre. Les possibilités cachées ne dépendent que d'un esprit sage et d'un cœur solide comme la pierre à force de tant sentir... sentir envers d'autres mondes...

... América

LISTE DES FIGURES	III
LISTE DES TABLEAUX.....	III
INTRODUCTION.....	1
1. ETAT DE L'ART DE L'IDENTIFICATION DES SYSTEMES NON LINEAIRES.....	3
1.1 REFERENCES CLE	3
1.2 TECHNIQUE D'IDENTIFICATION PAR LA METHODE DES SOUS-ESPACES	5
1.2.1 Réalisation d'un système linéaire.....	5
1.2.2 Identification	7
1.2.3 Cas bilinéaire général.....	9
1.2.4 Cas de modèles Hammerstein et Wiener	10
1.3 REPRESENTATION ENTREE – SORTIE.....	13
1.3.1 Cas linéaire	13
1.3.2 Cas non linéaire	16
2. STRUCTURES POLYNOMIALES	21
2.1 INTRODUCTION	21
2.2 DEFINITION D'UNE SERIE DE VOLTERRA	22
2.3 REALISATION D'UNE SERIE DE VOLTERRA A NOYAUX SEPARABLES PAR UN SYSTEME BILINEAIRE	24
2.3.1 Série de degré infini	24
2.3.2 Série homogène (de degré fini)	25
2.4 TRANSFORMES EN Z ET FONCTIONS DE TRANSFERT MULTIVARIABLES	27
2.5 FORMES CANONIQUES DES SYSTEMES BILINEAIRES HOMOGENES	30
2.5.1 Forme observateur	30
2.5.2 Forme contrôleur	33
2.5.3 Forme hybride.....	35
2.5.4 Forme linéaire en les paramètres	36
2.6 FORMES ENTREE/SORTIES MULTI-DEGRE.....	39
2.6.1 Cas général.....	39
2.6.2 Forme linéaire en les paramètres	43
3. REALISATION ET IDENTIFICATION PAR L'APPROCHE HEREDITAIRE	45
3.1 INTRODUCTION	45
3.2 REALISATION DES SYSTEMES LINEAIRES.....	46
3.2.1 Principe	46
3.2.2 Algorithme.....	48
3.3 IDENTIFICATION DES SYSTEMES LINEAIRES	52
3.3.1 Principe.....	52
3.3.2 Algorithme.....	53
3.4 REALISATION DES SYSTEMES BILINEAIRES HOMOGENES	54
3.4.1 Principe général.....	54
3.4.2 Cas Quadratique	56
3.5 IDENTIFICATION DES SYSTEMES BILINEAIRES HOMOGENES.....	63
3.5.1 Principe général.....	63
3.5.2 Algorithme héréditaire	65
4. EXEMPLES DE MISE EN OEUVRE EN IDENTIFICATION HEREDITAIRE.....	67
4.1 EXEMPLE LINEAIRE.....	67
4.2 EXEMPLE QUADRATIQUE	69
4.2.1 Système scalaire quadratique.....	69
4.2.2 Approximation d'une Série de Volterra de Degré Infini	71
CONCLUSION.....	76
BIBLIOGRAPHIE.....	77

Liste des figures

2.1. Modèle de Hammerstein (a) et Wiener (b)	10
2.2. Modèle de Hammerstein-Wiener	11
4.1 Comparaisons des performances des algorithmes	67
4.2 Estimations des paramètres	68
4.3 Paramètre a_1^2	69
4.4 Paramètre a_1^1	70
4.5 Paramètre b_1	71
4.6 Approximation de la sortie sans bruit du modèle non linéaire par (4.4)	
Identification/validation (50% des échantillons pour chaque opération) .	73
4.7 Approximation de la sortie bruitée du modèle non linéaire avec les méthodes	
héréditaire et des sous-espaces. Identification/validation (50% des échantillons	
pour chaque opération)	74

Liste des Tableaux

4.1 Identification du modèle non linéaire	74
---	----

Introduction

L'utilisation des séries de Volterra pour l'analyse de systèmes non linéaires, et ses principes théoriques, ont été amplement étudiés par le passé [Sche80]. Les séries de Volterra se retrouvent dans l'analyse des distorsions non linéaires, plus particulièrement dans le fonctionnement des filtres employées pour la réception des signaux [Cher98]. Ainsi, on peut trouver son application dans l'ingénierie biomédicale pour l'identification des noyaux de Volterra de systèmes non linéaires viscoélastiques, et obtenir de cette façon, la structure du système [Zhang98].

Les séries de Volterra peuvent être vues comme une série de Taylor dynamique, laquelle présente dans chaque terme non linéaire une convolution multidimensionnelle. Sa partie linéaire correspond à la réponse impulsionnelle classique d'un système dynamique linéaire.

Sa représentation générale est donnée par :

$$y_t = \sum_{k=1}^{\infty} \sum_{\tau_1=-\infty}^t \sum_{\tau_2=-\infty}^t \cdots \sum_{\tau_k=-\infty}^t W_k(t, \tau_1, \dots, \tau_k) u_{\tau_1} u_{\tau_2} \cdots u_{\tau_k}$$

où u_t , y_t sont respectivement l'entrée et la sortie du système dynamique, et les W_k sont les noyaux de chaque terme de degré k dans la série.

Citons quelques exemples d'application récents. La modélisation des produits d'intermodulation par des séries de Volterra dans des amplificateurs de haute puissance pour des applications mobiles, et de transmission satellitaire, ont été développées dans [Voule01, LeGal01]. Ils sont également utilisés pour la transmission en bande passante de signaux modulés en QAM et PSK par un système s'avérant non linéaire [Cheng01]. Par ailleurs, une nouvelle approche pour l'identification des noyaux de Volterra a été définie dans le but de modéliser le comportement de circuits non linéaires et de systèmes opérationnels à haute fréquence [Wang03].

Ceci montre l'intérêt réel que les séries de Volterra représentent aujourd'hui dans l'analyse des systèmes non linéaires. Ainsi, leur application - pour modéliser, calculer, et expliquer le comportement de certains dispositifs de transmission et de réception des signaux - est importante. En conséquence, la capacité à approcher ces comportements non linéaires, repérés dans les signaux, trouve ses principes dans le traitement de ces signaux, et dans l'analyse de ces séries.

On peut remarquer que dans la littérature, la plupart des travaux concernent les séries de Volterra à horizon fini, c'est-à-dire définies par des équations du type :

$$y_t = \sum_{k=1}^d \sum_{\tau_1=t-h}^t \sum_{\tau_2=t-h}^t \cdots \sum_{\tau_k=t-h}^t W_k(t - \tau_1, \dots, t - \tau_k) u_{\tau_1} u_{\tau_2} \cdots u_{\tau_k}$$

où d désigne le degré de la fonctionnelle de Volterra et h l'horizon du passé. En effet, dans ce cas, le nombre de paramètres, représentés ici par les noyaux, est fini. Par ailleurs, ceux-ci

apparaissent linéairement dans l'expression de la sortie ce qui facilite leur détermination par une technique des moindres carrés par exemple. En revanche, pour décrire de manière suffisamment riche de tels systèmes, il s'avère que le nombre de paramètres nécessaire à la description des phénomènes physiques, comportant naturellement une mémoire infinie à travers une boucle de rétroaction, est souvent prohibitif.

Dans ce mémoire, nous travaillons sur les séries de Volterra à horizon infini mais dont les noyaux ont la particularité d'être séparables. On montre en effet que de telles séries peuvent être réalisées par des systèmes bilinéaires dits homogènes. Cela signifie qu'avec un nombre de paramètres raisonnable, on est capable de représenter des comportements beaucoup plus riches que ceux décrits par les séries de Volterra décrites plus haut.

L'objectif poursuivi est de développer des méthodes capables de résoudre les problèmes de réalisation et d'identification des systèmes non linéaires en les approximant par des séries de Volterra de degré infini à noyaux séparables.

Le problème de réalisation, tel que nous le posons, considère une approche stochastique du système à représenter. En effet, le problème consiste à rechercher un système bilinéaire convenablement paramétré dont le comportement entrée/sortie est aussi proche que possible de celui d'un système dont on connaît les intercorrélations entrée /sortie quand l'entrée est un bruit blanc. Dans le cas linéaire, la série des noyaux de Volterra correspond à la réponse impulsionnelle du système.

Le problème de l'identification des systèmes que nous considérons également dans ce mémoire est posé de manière classique : on cherche le système bilinéaire convenablement paramétré qui explique au mieux un comportement entrée /sortie observé, c'est-à-dire dont on connaît une ou plusieurs trajectoires entrée/sortie.

Le chapitre 1 présente l'état de l'art dans l'identification de systèmes non linéaires, et offre un parcours, que nous espérons à jour, des diverses techniques utilisées jusqu'à aujourd'hui. Ceci en restant autour des approches paramétriques de base analytique, et en s'attardant sur le cas de la technique populaire des sous espaces.

Le chapitre 2 s'attache à décrire les formes particulières des systèmes bilinéaires homogènes et d'en donner diverses représentations (état, entrée/sortie multi-degré) en leurs associant les formes canoniques, originales à notre connaissance. Ce préalable est en effet nécessaire à la résolution des problèmes de réalisation et d'identification.

Le chapitre 3 développe et décrit les algorithmes héréditaires proprement dits de réalisation et d'identification des systèmes dynamiques pour le cas des systèmes linéaires et bilinéaires homogènes.

Enfin, le chapitre 4 présente quelques résultats d'identification obtenus en simulation pour des systèmes quadratiques académiques.

1. Etat de l'Art de l'Identification des Systèmes Non Linéaires

1.1 Références clé

Cette section a été construite de façon régressive dans le temps pour permettre au lecteur d'appréhender le sujet traité en partant du présent. Cette approche permet ainsi de répondre à deux questions : « Où en sommes-nous de l'identification non linéaire ? » et « Quels sont les repères trouvés par le passé pour retracer le chemin du passé ? ». L'identification des systèmes non linéaires est un sujet très largement étudié, on peut dénombrer près de 550 publications dans les deux dernières années contenant le mot clé « Identification de Systèmes » dans ses résumés [Lju05].

Malgré cette importante quantité d'information, des éditions spéciales et des comptes rendus ont été réalisés par la communauté travaillant sur l'identification des systèmes au fil du temps. Par exemple l'édition spéciale la plus récente a été publiée en 2005 et on y trouve, parmi ses titres, l'application de la technique d'identification des sous espaces aux modèles de Hammerstein et Wiener. Ceux-ci sont des modèles dits « orientés en bloc » et représentent des sous classes de modèle de Volterra. On y trouve également d'autres approches plus récentes pour l'identification des systèmes non linéaires comme l'utilisation des « splines », lesquels sont des polynômes construits par morceaux, et des « machines à support vectoriel » [IEEE05].

Plus tôt, en 2001 et 1995, des analyses sur l'état de l'art dans l'identification des systèmes non linéaires ont montré les diverses possibilités d'approche de ces modèles par l'approche « boîte noire ». Cette approche revient à étudier la structure du système sans information *a priori*, en utilisant, ces dernières années, des réseaux de neurones, des représentations à base d'ondelettes, et des modèles flous. On y trouve référence à la régression non paramétrique employée par la communauté en statistique depuis plusieurs années, ainsi que les fondations mathématiques des nouvelles approches [Autom95, Lju01]. L'édition spéciale citée de 1995 a présentée dans plusieurs de ses publications la technique des sous espaces et l'unification théorique des ses différents algorithmes connus à la date de sa publication. Cette édition a également présenté un compte rendu de son application pour l'identification des systèmes linéaires.

Une édition spéciale a été dédiée en 1992 à l'identification des systèmes pour le contrôle robuste, incluant des aspects validation de modèle, contrôle de procédés industriels et utilisation d'un schéma récursif dans la tâche d'identification-contrôle [IEEE92]. En 1985, deux publications sur les modèles paramétriques de type entrée-sortie pour des systèmes non linéaires ont développé un modèle dérivé NARMAX (Nonlinear Autoregressive Moving Average with eXogenous input), lequel a été comparé à d'autres représentations comme celles de type polynomial (de Volterra) et affine dans l'état [Ln85a, Ln85b].

Un compte rendu effectué en 1980 a été centré sur les expansions en fonctionnelles de Wiener et Volterra, ainsi que sur l'identification de systèmes bilinéaires généraux et des structures en bloc liées aux modèles de Hammerstein déjà mentionnés. On y trouve beaucoup de références montrant, par son nombre et diversité, l'intérêt de la recherche menée dans le

domaine de l'identification des systèmes non linéaires à cette époque. Pour les systèmes linéaires, la théorie d'identification était déjà largement aboutie [Bill80, Lju83].

Finalement pour compléter cette partie de références clés, on ne pouvait pas ne pas mentionner le papier publié en 2001 par Giannakis et Serpedin. Ceux-ci y ont répertorié une grande liste de références sur l'identification des systèmes non linéaires et ses applications [Gian01].

En considérant notre intérêt pour l'identification des systèmes non linéaires vis-à-vis du traitement du signal, on a choisi dans ce chapitre de décrire l'état de l'art par deux approches essentielles : celles d'espace d'état et celles d'entrée-sortie. La méthode des sous espaces est la représentante principale de la première approche.

Des travaux récents montrent que les méthodes des sous espaces et d'erreur de prédiction peuvent être complémentaires pour achever une bonne performance dans l'identification des systèmes non linéaires. L'algorithme des sous espaces permettant une bonne initialisation pour l'estimation des paramètres, cette estimation peut alors être affinée par l'algorithme d'erreur de prédiction [Verd02].

1.2 Technique d'identification par la méthode des sous-espaces

1.2.1 Réalisation d'un système linéaire

Rappelons que le concept de réalisation d'un système linéaire consiste à déterminer la réalisation d'état d'un système linéaire sous la forme

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t \\ y_t &= Cx_t + Du_t \end{aligned} \quad (1.1)$$

(où $x_t \in \mathfrak{R}^n$, $A \in \mathfrak{R}^{n \times n}$, $B \in \mathfrak{R}^{n \times m}$, $C \in \mathfrak{R}^{p \times n}$, et $D \in \mathfrak{R}^{p \times m}$) dont on connaît la réponse impulsionnelle $\mathbf{h}_\tau \in \mathfrak{R}^{p \times m}$ (*paramètres de Markov*). Celle-ci est définie par :

$$y_t = \sum_{\tau=0}^{\infty} \mathbf{h}_\tau u_{t-\tau} \quad (1.2)$$

où $u_t \in \mathfrak{R}^{m \times 1}$ est l'entrée et $y_t \in \mathfrak{R}^{p \times 1}$ la sortie. Si on applique une entrée impulsionnelle à l'équation (1.1), en considérant l'état initial nul ($x_0=0$), on obtient :

$$\mathbf{h}_t = \begin{cases} 0 & t < 0 \\ D & t = 0 \\ CA^{t-1}B & t > 0 \end{cases} \quad (1.3)$$

Il est clair que la détermination de la matrice D est donnée directement par la valeur de h_0 .

Considérons à présent la matrice de Hankel d'ordre $(n+1) \times (n+1)$ définie par :

$$\mathbf{H} = \begin{bmatrix} h_1 & h_2 & \cdots & h_{n+1} \\ h_2 & h_3 & \cdots & h_{n+2} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n+1} & h_{n+2} & \cdots & h_{2n+1} \end{bmatrix} \quad \mathbf{U} = \begin{bmatrix} u_1 & u_2 & \cdots & u_N \\ u_2 & u_3 & \cdots & u_{N+1} \\ \vdots & \vdots & \ddots & \vdots \\ u_\alpha & u_{\alpha+1} & \cdots & u_{N+\alpha-1} \end{bmatrix}$$

On montre alors aisément que $\mathbf{H}$ se factorise selon le produit suivant :

$$\mathbf{H} = \Gamma_{n+1} \Omega_{n+1}$$

avec

$$\Gamma_n = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \quad (1.4)$$

et

$$\Omega_n = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} \quad (1.5)$$

Ces matrices, nommées respectivement matrices d'observabilité et de contrôlabilité, sont de rang plein si, respectivement, le système est observable et contrôlable, condition nécessaire et suffisante pour que la réalisation définie par (1.1) soit minimale. Dans ce cas, la matrice $\mathbf{H}$ est naturellement de rang plein égal à la dimension du système.

Par construction, il est aisé de vérifier que cette factorisation possède la propriété d'invariance suivante :

$$\Gamma_{2:n+1} = \Gamma_{1:n} A \quad (1.6)$$

où $\Gamma_{i:j}$ désigne la sous matrice de Γ_{n+1} définie par les p (due à la sortie) blocs de lignes i à j .

On obtient alors une réalisation minimale du système de la manière suivante :

1. Calcul de la suite de matrices de Hankel de dimensions $n \geq 1$ et calcul du rang de celles-ci. La première valeur r pour laquelle la matrice devient singulière fournit la dimension du système, c'est-à-dire $n = r-1$
2. Calcul d'une factorisation de la matrice de Hankel $\mathbf{H} = \Gamma_{n+1} \Omega_{n+1}$ telle que Γ_{n+1} et Ω_{n+1} soient de rang plein
3. Les premiers blocs de la matrice Γ_{n+1} donnent la matrice C tandis que les premiers blocs de Ω_{n+1} donnent la matrice B
4. La matrice A est alors déterminée par la pseudo inverse suivante (*en général Γ_n n'est carrée*) :

$$A = (\Gamma_{1:n})^{-1} \Gamma_{2:n+1} \quad (1.7)$$

Il est bien connu que le résultat obtenu n'est pas unique, tout changement de variable inversible sur l'état x_t donnant la même représentation entrée/sortie du système. A noter que la factorisation de la matrice de Hankel définie dans la première étape de l'algorithme peut s'obtenir à partir d'une décomposition de celle-ci en ses valeurs singulières. En effet, si on calcule la décomposition en valeurs singulières de la matrice $\mathbf{H}$ sous la forme $\mathbf{H} = QSV$ où $S \in \mathbb{R}^{n \times n}$ est la matrice diagonale des valeurs singulières de $\mathbf{H}$ et où $Q \in \mathbb{R}^{m \times n}$, $V \in \mathbb{R}^{n \times n}$ sont des matrices orthogonales, il suffit de définir $\Gamma_{n+1} = US^{1/2}$ et $\Omega_{n+1} = S^{1/2}V$ pour définir la factorisation désirée.

1.2.2 Identification

1.2.2.1 Principe de base

En général, on ne dispose pas aisément en pratique de la réponse impulsionnelle d'un système réel. Outre la difficulté pratique de définir une entrée impulsionnelle réaliste, elle donne lieu naturellement à des saturations des actionneurs. Pour contourner cette difficulté, une approche stochastique permet à partir de la notion de bruit blanc d'obtenir la réponse impulsionnelle de la manière suivante : si à la place d'une impulsion on applique au système un bruit blanc, l'intercorrélation entrée/sortie donne directement la réponse impulsionnelle. En effet, considérant la définition (1.2) avec u_t bruit blanc, il est clair que $E[y_t u_{t-\tau}^T] = h_\tau$. On peut alors imaginer diverses techniques statistiques pour obtenir cette espérance mathématique à partir de relevés expérimentaux.

Par ailleurs, les mesures faites sur un système sont toujours imparfaites et soumises à des bruits perturbateurs. En pratique, la matrice de Hankel est toujours de rang plein. Diverses approches permettent de contourner cette difficulté. Une approche possible consiste par exemple à considérer une matrice $\hat{\mathbf{H}}$ surdimensionnée, c'est-à-dire de dimension largement supérieure à la dimension de l'état, à décomposer celle-ci en valeurs singulières et à ne conserver de cette décomposition que les valeurs singulières supérieures à un certain seuil laissé à la discrétion de l'utilisateur.

Une seconde approche, plus directe, s'appuie sur une représentation de la transformation de **trajectoires** entrée/sortie qui sont de fait les données du problème. Pour ce faire, on construit les matrices de Hankel suivantes :

$$\mathbf{U} = \begin{bmatrix} u_1 & u_2 & \cdots & u_N \\ u_2 & u_3 & \cdots & u_{N+1} \\ \vdots & \vdots & \ddots & \vdots \\ u_\alpha & u_{\alpha+1} & \cdots & u_{N+\alpha-1} \end{bmatrix}$$

$$\mathbf{Y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_N \\ y_2 & y_3 & \cdots & y_{N+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_\alpha & y_{\alpha+1} & \cdots & y_{N+\alpha-1} \end{bmatrix}$$

et la matrice

$$\mathbf{X} = [x_1, \cdots, x_N]$$

où $\alpha > n$, la dimension présumée du système, afin d'assurer que ces matrices soient de rang plein. On montre aisément alors que ces matrices sont liées par la relation suivante :

$$\mathbf{Y} = \Gamma_\alpha \mathbf{X} + \Phi_\alpha \mathbf{U} \quad (1.8)$$

où

$$\Phi_\alpha = \begin{bmatrix} D & 0 & \dots & \dots & 0 \\ CB & D & & & 0 \\ CAB & CB & D & & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ C(A)^{(\alpha-2)}B & \dots & \dots & CB & D \end{bmatrix} \quad (1.9)$$

Pour éliminer l'influence du terme Φ_α inconnu, on réalise la projection hilbertienne de $\mathbf{Y}$ sur l'espace linéaire engendré par $\mathbf{U}$. Cette projection est définie par :

$$P[\mathbf{Y}|\mathbf{U}] = \mathbf{Y}\mathbf{U}^T (\mathbf{U}\mathbf{U}^T)^{-1} \mathbf{U} \quad (1.10)$$

Par définition de la projection, on a l'orthogonalité suivante :

$$(\mathbf{Y} - P[\mathbf{Y}|\mathbf{U}])\mathbf{U}^T = 0 \quad (1.11)$$

Si on note $\Pi = \mathbf{I} - \mathbf{U}^T (\mathbf{U}\mathbf{U}^T)^{-1} \mathbf{U}$, l'orthogonalité du sous espace de projection se traduit par $\mathbf{U}\Pi = \Pi\mathbf{U}^T = 0$. En exploitant cette orthogonalité dans l'expression de $\mathbf{Y}$ (1.8), on obtient :

$$\mathbf{Y}\Pi = \Gamma_\alpha \mathbf{X}\Pi \quad (1.12)$$

Une décomposition en valeurs singulières de $\mathbf{Y}\Pi$ donne alors Γ_α . En pratique, α étant largement plus grand que la dimension présumée du système, on ne retient pour l'estimation de Γ_α que les n valeurs singulières les plus significatives, et ainsi obtenir la dimension du système. Une fois de plus, le choix du seuil qui élimine les faibles valeurs singulières est laissé à la discrétion de l'utilisateur. Ainsi, $\mathbf{Y}\Pi$ est décomposée en la somme suivante :

$$\mathbf{Y}\Pi = \mathcal{Q}_s S_s V_s^T + \mathcal{Q}_b S_b V_b^T \quad (1.13)$$

où le premier terme représente la partie « signal » et le deuxième la partie « bruit ». A noter que ces sous espaces sont orthogonaux et qu'en particulier $\mathcal{Q}_s \mathcal{Q}_b^T = 0$. Donc, l'estimée de la matrice d'observabilité est donnée par :

$$\Gamma_\alpha = \mathcal{Q}_s \quad (1.14)$$

Cette matrice étant ainsi déterminée, elle permet par les équations (1.4) et (1.6-1.7) d'estimer les matrices A et C . La matrice Φ_n est alors déterminée en projetant l'équation (1.8) sur l'espace des bruits, soit en multipliant à gauche par $\mathcal{Q}_b^T$ et à droite par $\mathbf{U}^T (\mathbf{U}\mathbf{U}^T)^{-1}$:

$$\mathcal{Q}_b^T \mathbf{Y}\mathbf{U}^T (\mathbf{U}\mathbf{U}^T)^{-1} = \mathcal{Q}_b^T \Gamma_\alpha \mathbf{X}\mathbf{U}^T (\mathbf{U}\mathbf{U}^T)^{-1} + \mathcal{Q}_b^T \Phi_n \mathbf{U}\mathbf{U}^T (\mathbf{U}\mathbf{U}^T)^{-1}$$

et d'après (1.14) :

$$\mathcal{Q}_b^T \mathbf{Y}\mathbf{U}^T (\mathbf{U}\mathbf{U}^T)^{-1} = \mathcal{Q}_b^T \mathcal{Q}_s \mathbf{X}\mathbf{U}^T (\mathbf{U}\mathbf{U}^T)^{-1} + \mathcal{Q}_b^T \Phi_n \mathbf{I}$$

ce qui fait possible de trouver, à cause de l'orthogonalité des espaces signal-bruit (donc $Q_b^T Q_s = 0$), la relation suivante :

$$Q_b^T Y U^T (U U^T)^{-1} = Q_b^T \Phi_n \quad (1.15)$$

d'où nous pouvons extraire les matrices B et D en regardant l'expression (1.9).

1.2.2.2 Extensions

La technique de sous espace est apparue à la fin des années quatre-vingt et a été largement développée dans la décennie suivante. Les pionniers à l'origine de cette méthode sont Larimore : algorithme **CVA** (Canonical Variate Analysis) [Lar90]; Verhaegen : algorithme **MOESP** (Multivariable Output-Error State Space) [Verh94]; et Viberg, Van Overschee et De Moor : algorithme **N4SID** (Numerical algorithms for Subspace State Space System Identification) [Vib93, Van94]. Les différences entre ces algorithmes sont basées sur deux aspects :

1. Le sous espace caractéristique obtenu à partir des données d'entrée-sortie.
2. L'obtention en premier lieu à partir de l'ordre n du système soit :
 - a. De l'estimation de deux séquences consécutives de l'état pour déterminer les matrices du système en résolvant un système d'équations par moindres carrés.
 - b. Des matrices A , C à partir de la matrice étendue d'observabilité pour après déterminer les matrices restantes du système.

Plusieurs extensions à des systèmes plus riches, tels que de systèmes ayant des bruits non mesuré en entrée.

Les méthodes de sous espaces ont été comparées à l'algorithme dit « traditionnel » d'erreur de prédiction (PEM, Prediction Error Method), ainsi qu'avec la technique des variables instrumentales (IVM, Instrumental Variables Method). Ces comparaisons ont permis d'identifier sa performance. Deux types d'entrée ont été utilisés pour cet effet : un bruit blanc et une somme de signaux sinusoïdaux. [Vib95]. Le résultat de cette comparaison a montré que l'algorithme IV-4SID est compétitif par rapport à celui d'erreur de prédiction pour le premier type d'entrée et moins performante pour une entrée moins riche en fréquence comme celle de somme des sinusoïdes. Une autre comparaison entre la méthode N4SID et celle de PEM a été menée aussi par le passé, mais cette fois-ci en utilisant un ensemble de données de procédés industriels [Fav00].

1.2.3 Cas bilinéaire général

Il y a eu quelques tentatives d'extension au cas de l'identification de systèmes bilinéaires généraux de la forme :

$$\begin{aligned} x_{t+1} &= Ax_t + Nu_t \otimes x_t + Bu_t + w_t \\ y_t &= Cx_t + Du_t + v_t \end{aligned} \quad (1.16)$$

où la matrice $N = [N_1 N_2 \cdots N_m] \in \mathbb{R}^{n \times nm}$, $N_i \in \mathbb{R}^{n \times n}$, $i = 1, \dots, m$ caractérise la bilinéarité du système. Le produit de Kronecker de deux vecteurs quelconque $a \in \mathbb{R}^p$ et $b \in \mathbb{R}^q$ est défini comme $a \otimes b = [a_1 b^T \cdots a_p b^T]^T \in \mathbb{R}^{pq}$.

Afin de généraliser cette technique aux systèmes bilinéaires, une contrainte sur l'entrée du système comme étant un bruit blanc indépendant du bruit de mesure v_k et du bruit du processus w_k [Fav99] a été ajoutée. De cette façon la théorie de sous espaces pour les systèmes linéaires a pu être étendue au cas bilinéaire avec les matrices de covariance définies par :

$$\mathbf{E} \left[\begin{pmatrix} u_p \\ w_p \\ v_p \end{pmatrix} \begin{pmatrix} u_q^T & w_q^T & v_q^T \end{pmatrix} \right] = \begin{pmatrix} I_m & 0 & 0 \\ 0 & Q & S \\ 0 & S^T & R \end{pmatrix} \delta_{pq} \geq 0 \quad (1.17)$$

Il a été souligné, dans la référence ici concernée, que l'utilisation du bruit blanc comme entrée pour identifier le système peut sembler restrictive. Mais la richesse de son spectre permet d'exciter suffisamment un plus large nombre des modes du système, ce qui est important pour la tâche d'identification. Plus tard cette restriction a été surmontée en utilisant la technique de sous espaces à « quatre blocs » [Chen00].

Les dimensions de matrices du système bilinéaire à identifier étant très large, une technique d'optimisation du modèle bilinéaire a été proposée par l'emploi de la technique dite des moindres carrés séparables. Pour résoudre ce problème, une estimation initiale des matrices A , C , et N du système bilinéaire (1.16) par la méthode de sous espaces est réalisée, pour ensuite l'améliorer de façon numérique optimisation non linéaire. Finalement, on obtient les matrices B et D comme solution d'un problème linéaire de moindres carrés [Verd99], [Verd01].

1.2.4 Cas de modèles Hammerstein et Wiener

Ce type de modèles sont nommés « orientés en bloc », du fait de leur structure particulière. Celle-ci consiste en une interconnexion de systèmes linéaires invariants dans le temps et de non linéarités statiques sans mémoire. Le modèle nommé de Hammerstein utilise une non linéarité statique suivie d'un système linéaire dynamique, tandis que le modèle de Wiener inverse l'ordre de ces composants.

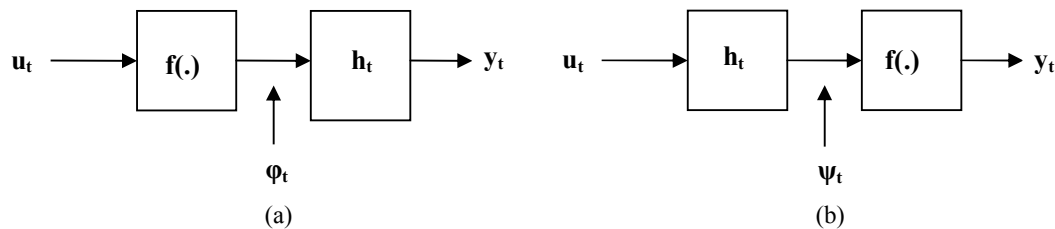


Figure 2.1. Modèle de Hammerstein (a) et Wiener (b).

Dans cette figure, la non linéarité statique est représentée par $f(\cdot)$ et le système linéaire dynamique par h_t . Pour le cas du modèle de Hammerstein, l'entrée intermédiaire et la sortie sont données respectivement par les équations suivantes :

$$\phi_t = f(u_t) = \sum_{k=0}^d (u_t)^k \quad y_t = \sum_{\tau=0}^{t_h} h_{\tau} \phi_{t-\tau} \quad (1.18)$$

tandis que pour celui de Wiener on a :

$$\psi_t = \sum_{\tau=0}^{t_h} h_{\tau} u_{t-\tau} \quad y_t = \sum_{k=0}^d (\psi_t)^k \quad (1.19)$$

où d est le degré de la non linéarité du système et t_h la mémoire de la partie linéaire dynamique.

Ces modèles ont été utilisés dans le passé dans divers domaines d'application et ont encouragé la recherche sur le développement de nouvelles techniques d'identification. Par exemple la technique de sous espaces MOESP a été étendue à l'identification des modèles de Wiener par le développement de deux versions : celle nommée PI (Past Inputs) et celle nommée RS (Reconstructed State variables), où les entrées et les variables d'état estimées sont respectivement utilisées comme des variables instrumentales pour l'identification [West96]. Il a été montré que l'identification d'un système en considérant les modèles de Hammerstein et Wiener est possible en utilisant en premier lieu la méthode de sous espaces pour ensuite réaliser une minimisation au moyen de la décomposition SVD (Singular Value Decomposition) [Gom02]. Plus tard une autre méthode a été développée pour l'identification des modèles Hammerstein en utilisant l'algorithme de sous espaces N4SID et la notion de moindres carrés par des « machines à support vectoriel » insérée dans les étapes de la méthode de sous espaces [Goeth05].

Récemment, une structure en bloc de type Wiener-Hammerstein a été utilisée pour l'identification des noyaux de Volterra associés à cette structure avec l'implémentation de la méthode des moindres carrés « alternés » [Kiban06]. Ainsi, le modèle employé est donné par :

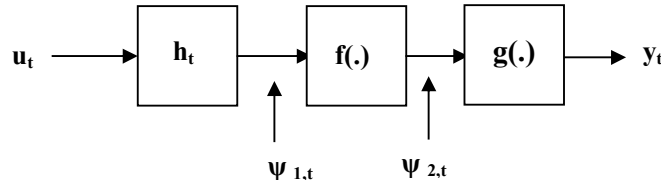


Figure 2.2. Modèle de Wiener- Hammerstein.

où

$$\psi_{1,t} = \sum_{\tau=0}^{t_h} h_{\tau} u_{t-\tau} \quad \psi_{2,t} = \sum_{k=0}^d (\psi_t)^k \quad (1.20)$$

$$y_t = \sum_{\tau=0}^{t_g} g_{\tau} \psi_{2,t-\tau} \quad (1.21)$$

avec t_h et t_g respectivement les mémoires des blocs linéaires d'entrée et de sortie.

Ainsi, les noyaux de Volterra associés à cette structure peuvent être complètement déterminés à partir seulement des coefficients dans la diagonal. Donc, dans la série de Volterra de degré d (voir chapitre 2 pour plus de détails) :

$$y_t = \sum_{k=1}^d \sum_{\tau_1=0}^t \sum_{\tau_2=0}^t \cdots \sum_{\tau_k=0}^t W_k(\tau_1, \dots, \tau_k) u_{t-\tau_1} u_{t-\tau_2} \cdots u_{t-\tau_k}$$

les coefficients hors la diagonal des noyaux sont donnés par :

$$W_k(\tau_1, \dots, \tau_k) = \frac{1}{t_h + 1} \sum_{\tau_g=0}^{t_g} \sum_{\tau_h=0}^{t_h} g(\tau_g) \frac{h_d(\tau_h)}{(h_1(\tau_h))^d} \prod_{k=1}^d h_1(\tau_{h,k} - \tau_g) \quad (1.22)$$

où $h_d(\tau_h), \tau_g = 0, \dots, t_h$ sont les coefficients dans la diagonal du noyau de Volterra de degré non linéaire d associés à la partie de Wiener du model (système linéaire h_t + non linéarité statique $f(\cdot)$) :

$$W_d(\tau) = W_d(\tau, \dots, \tau) = \sum_{\tau_g=0}^{t_g} g(\tau_g) h_d(\tau - \tau_g), \quad \tau = 0, \dots, t \quad (1.23)$$

où $t = t_h + t_g$.

1.3 Représentation Entrée – Sortie

1.3.1 Cas linéaire

Pour réaliser l'identification d'un système, on dispose de mesures des trajectoires d'entrée-sortie $u_\tau, y_\tau \in \mathfrak{R}$, $\tau = 1, \dots, t$ (cas ARX), ou seulement de la mesure de la trajectoire de sortie $y_\tau \in \mathfrak{R}$, $\tau = 1, \dots, t$ (cas ARMAX). Comme il c'est toujours le cas, les mesures de sortie sont bruitées :

$$y_t = \sum_{\tau=0}^{\infty} h_\tau u_{t-\tau} + v_t \quad (1.24)$$

Dans un modèle paramétrique, l'objectif est de caractériser au moyen de fonctions de transfert et de ses paramètres, le transfert de l'entrée et du bruit :

$$y_t = H(z) u_t + N(z) e_t \quad (1.25)$$

Pour cela, la décomposition des fonctions de transfert en des quotients de polynômes pour expliquer la dynamique du système est réalisée. L'expression des modèles couramment utilisés pour l'identification des systèmes linéaires est la suivante [Lju99] :

$$A(z)y_t = \frac{B(z)}{F(z)}u_t + \frac{C(z)}{D(z)}e_t \quad (1.26)$$

où u_t, y_t sont respectivement l'entrée et la sortie du système, et e_t est un bruit blanc quelconque. $A(z), B(z), F(z), C(z)$, et $D(z)$ sont des polynômes en z^{-1} , où z désigne l'opérateur décalage. Ils sont donnés par :

$$\begin{aligned} A(z) &= 1 + a_1 z^{-1} + \dots + a_{n_a} z^{-n_a}, \\ B(z) &= b_1 z^{-1} + b_2 z^{-2} + \dots + b_{n_b} z^{-n_b}, \\ F(z) &= 1 + f_1 z^{-1} + \dots + f_{n_f} z^{-n_f}, \\ C(z) &= 1 + c_1 z^{-1} + \dots + c_{n_c} z^{-n_c}, \\ D(z) &= 1 + d_1 z^{-1} + \dots + d_{n_d} z^{-n_d}. \end{aligned}$$

où $b_0=0$ dans le polynôme $B(z)$ si nous considérons des fonctions de transfert *strictement singulières*.

Le prédicteur du modèle (1.24) trouve son principe dans la considération du bruit v_t comme un processus, lequel peut être modélisé par :

$$v_t = N(z)e_t = \sum_{\tau=0}^{\infty} n_\tau e_{t-\tau} \quad (1.27)$$

où e_t est une séquence de variables aléatoires indépendantes.

Le prédicteur d'un pas en avance dépend de la connaissance de v_τ pour $\tau \leq t-1$, ce qu'implique la connaissance de e_τ pour $\tau \leq t-1$. Ainsi, dans l'expression suivante :

$$v_t = N(z)e_t = e_t + \sum_{\tau=1}^{\infty} n_\tau e_{t-\tau} \quad (1.28)$$

on peut nommer :

$$\hat{v}_{t|t-1} = \sum_{\tau=1}^{\infty} n_\tau e_{t-\tau}$$

comme la *valeur conditionnelle espérée*, basée sur la meilleure prédiction du v_t au temps $t-1$, avec e_t un processus de moyenne nulle. D'où on peut écrire d'après (1.27-1.28):

$$\hat{v}_{t|t-1} = [N(z)-1]e_t = [1-N^{-1}(z)]v_t \quad (1.29)$$

Si on prend ensuite l'expression (1.24) pour $\tau \leq t-1$, on a :

$$v_\tau = y_\tau - H(z)u_\tau \quad (1.30)$$

et par là même, la valeur conditionnelle espérée de y_t peut être donné par :

$$\hat{y}_{t|t-1} = H(z)u_t + \hat{v}_{t|t-1}$$

Ainsi, avec les expressions (1.29, 1.30) on obtient :

$$\begin{aligned} \hat{y}_{t|t-1} &= H(z)u_t + [1-N^{-1}(z)]v_t \\ &= H(z)u_t + [1-N^{-1}(z)][y_t - H(z)u_t] \end{aligned}$$

pour arriver finalement à l'expression :

$$\hat{y}_{t|t-1} = N^{-1}(z)H(z)u_t + [1-N^{-1}(z)]y_t \quad (1.31)$$

Donc, en dépendant du type de modèle considéré pour analyser le système de départ, le prédicteur résultant sera donné, en remplaçant dans l'expression (1.25) :

$$H(z) = \frac{B(z)}{A(z)F(z)} \quad \& \quad N(z) = \frac{C(z)}{A(z)D(z)}$$

où l'on retrouve (1.26) et de cette forme le prédicteur général à partir du (1.31) comme :

$$\hat{y}_t = y_t - \frac{D(z)}{C(z)} \left[A(z)y_t - \frac{B(z)}{F(z)}u_t \right] \quad (1.32)$$

Ensuite nous pouvons définir les composants agissant dans la structure du prédicteur comme :

$$\tilde{y}_t = y_t - \hat{y}_t = \frac{D(z)}{C(z)} \left[A(z)y_t - \frac{B(z)}{F(z)}u_t \right]$$

puis

$$\hat{y}_{u,t} = \frac{B(z)}{F(z)}u_t$$

et

$$\tilde{y}_{u,t} = A(z)y_t - \hat{y}_{u,t}$$

d'où finalement

$$\tilde{y}_t = \frac{D(z)}{C(z)}\tilde{y}_{u,t}$$

Le prédicteur du modèle (1.32) peut être récrit de forme compacte comme suit :

$$\hat{y}_t = \theta^T \varphi_{t,\theta} \quad (1.33)$$

où les vecteurs de paramètres θ et de régression $\varphi_{t,\theta}$, sont exprimés par :

$$\theta^T = (a_1 \dots a_{n_a} \quad b_1 \dots b_{n_b} \quad f_1 \dots f_{n_f} \quad c_1 \dots c_{n_c} \quad d_1 \dots d_{n_d}) \quad (1.34)$$

$$\varphi_{t,\theta}^T = (y_{t-1} \dots y_{t-n_a} \quad u_{t-1} \dots u_{t-n_b} \quad \hat{y}_{u,t-1} \dots \hat{y}_{u,t-n_f} \quad \tilde{y}_{t-1} \dots \tilde{y}_{t-n_c} \quad \tilde{y}_{u,t-1} \dots \tilde{y}_{u,t-n_d}) \quad (1.35)$$

et où les indexes n_a, n_b, n_f, n_c, n_d désignent le nombre de retards temporels des données entrée, sortie, et ceux obtenues pendant le calcul de l'estimateur et de l'erreur. Chaque type de donnée est lié à son coefficient respectif.

En désignant par n le retard apparaissant dans (1.35), la relation que ces derniers entretiennent avec chaque polynôme est la suivante :

- y_{t-n} , sortie mesurée (associé au polynôme A)
- u_{t-n} , entrée mesurée (associé au polynôme B)
- $\hat{y}_{u,t-n}$, prédiction basée sur les entrées passées (associé au polynôme F)
- $\tilde{y}_{t-n} = y_{t-n} - \hat{y}_{t-n}$, erreur de prédiction (associé au polynôme C)
- $\tilde{y}_{u,t-n} = y_{t-n} - \hat{y}_{u,t-n}$, erreur de simulation (associé au polynôme D)

où nous devons signaler que dans le cas dans lequel la matrice $A \neq 1$, la sortie simulée $\hat{y}_{u,t-n}$ utilisé dans le vecteur de régression devient $A(z)y_t$.

Il est clair que nous pouvons choisir le type de modèle employé, selon l'approche à implémenter. Ce choix dépend directement des éléments inclus dans le vecteur de régression (1.35) ; et pour ce faire, nous cherchons à déterminer les polynômes associés dans l'expression (1.26) lesquels étant différents de la valeur unité, donnent lieu aux quelques modèles particuliers suivants :

1. *FIR* (Finite Impulse Response), polynômes $A = F = C = D = 1$

2. *BJ* (Box Jenkins), polynôme $A = 1$
3. *ARMAX* (Autoregressive Moving Average eXogenous input), polynômes $F = D = 1$
4. *OE* (Output Error), polynômes $A = C = D = 1$
5. *ARX* (Autoregressive eXogenous input), avec polynômes $F = C = D = 1$

Il est important de signaler que le prédictor défini dans l'expression (1.33) est en général une *régression pseudo linéaire*. En effet, $\varphi_{t,\theta}$ dépendant lui-même du vecteur de paramètres θ , la relation entre prédictor et paramètres est donc clairement non linéaire. Une approche classique consiste à estimer d'abord les paramètres θ par un moyen quelconque pour ensuite réactualiser les nouvelles données $\varphi_{t,\theta}$.

Le critère d'évaluation pour décider de la capacité d'un modèle à décrire des données observées est sa capacité de prédiction. Pour ce faire, on utilise généralement l'erreur de prédiction :

$$\tilde{y}_t = y_t - \hat{y}_t \quad (1.36)$$

Il est clair que l'objectif principal d'un algorithme d'identification est minimiser cette erreur en un certain sens, pour ainsi récupérer les paramètres du système en question. Ainsi, le critère généralement utilisé à cet effet est celui des moindres carrés qui, sur un horizon fixé t est défini par :

$$V_t(\theta, \varphi_t) = \frac{1}{2} E[\tilde{y}_t^2] = \frac{1}{2t} \sum_{\tau=1}^t (y_\tau - \hat{y}_\tau)^2 \quad (1.37)$$

Pour minimiser ce critère, il est nécessaire d'assurer la nullité de la dérivée par rapport aux paramètres, soit :

$$\frac{\partial V_t(\theta, \varphi_t)}{\partial \theta} = -\frac{1}{t} \sum_{\tau=1}^t (y_\tau - \hat{y}_\tau) \left(\frac{\partial \hat{y}_\tau}{\partial \theta} \right) = 0$$

puis en y substituant le prédictor (1.33), et en faisant la répartition des produits à chaque côté de l'égalité, nous trouvons :

$$\hat{\theta}_t = \left[\frac{1}{t} \sum_{\tau=1}^t \varphi_\tau \varphi_\tau^T \right]^{-1} \frac{1}{t} \sum_{\tau=1}^t \varphi_\tau y_\tau \quad (1.38)$$

si on suppose que $\varphi_{t,\theta}$ ne contient pas des données dépendantes de θ comme c'est cependant le cas pour $\hat{y}_{u,t-n}, \tilde{y}_{t-n}, \tilde{y}_{u,t-n}$. L'équation (1.38) montre seulement de façon simple la solution d'un problème de moindres carrés pour une *régression*.

1.3.2 Cas non linéaire

De l'ensemble de représentations d'entrée-sortie des systèmes non linéaires on peut retenir principalement trois cas :

- La représentation de type NARMAX (Nonlinear Autoregressive Moving Average with eXogenous input), lequel est un modèle polynomial aux différences récursives.

- La représentation en séries de Volterra.
- La représentation par un modèle bilinéaire général.

Il existe d'autres représentations entrée-sortie pour les systèmes non linéaires, mais on a choisi celles-ci parce qu'elles ont été étudiées plus amplement par le passé dans divers publications. Leur fiabilité pour approcher des systèmes non linéaires a d'ores et déjà été établie. D'ailleurs, le modèle bilinéaire général est un type de représentation polynomial aux différences récursives, avec lequel on peut aussi approcher les noyaux d'une série de Volterra [Haber99].

1.3.2.1 Représentation de type NARMAX

Le modèle NARMAX a été développé dans les années quatre-vingt pour de systèmes SISO (single input, single output) et sa forme est définie par l'équation :

$$y_t = f(y_{t-1}, \dots, y_{t-n_y}, u_t, \dots, u_{t-n_u}, e_t, \dots, e_{t-n_e}) + e_t \quad (1.39)$$

où f est une fonction non linéaire, u_t , y_t , e_t sont respectivement l'entrée, sortie, et le bruit du système. n_u , n_y , n_e sont ainsi l'ordre du modèle de l'entrée, de sortie, et de bruit ; e_t est considéré une séquence de bruit blanc. Si la non linéarité est décrite par un polynôme de degré d , on peut écrire :

$$y_t = \sum_{i=1}^{n_\theta} \theta_i x_{i,t} + \tilde{y}_t \quad (1.40)$$

où

$$n_\theta = \sum_{i=0}^d n_i \quad (1.41)$$

est le nombre total des composantes dans le polynôme, dont n_i :

$$n_0 = 1 \quad n_i = n_{i-1}(n_y + n_u + n_e + i - 1)/i, \quad i = 1, \dots, d$$

représente le nombre des termes de degré $i = 1, \dots, d$ par rapport aux paramètres :

$$\theta = (\theta_1, \dots, \theta_{n_\theta})^T \quad (1.42)$$

Ainsi, chaque composante d'état se construit de la façon suivante :

$$\begin{aligned} x_{1,t} &= 1 \\ x_{i,t} &= \prod_{j=1}^p y_{t-n_{yj}} \cdot \prod_{k=1}^q u_{t-n_{uk}} \cdot \prod_{m=1}^r \tilde{y}_{t-n_{em}} \\ i &= 2, \dots, n_\theta, \quad p, q, r \geq 0, \quad 1 \leq p + q + r \leq d \quad 1 \leq n_{yj} \leq n_y \\ 1 &\leq n_{uk} \leq n_u, \quad 1 \leq n_{em} \leq n_e \end{aligned} \quad (1.43)$$

où par convention, $p=0$, $q=0$ et/ou $r=0$, indique que $x_{i,t}$ ne contient pas des termes $y(\cdot)$, $u(\cdot)$, ou $\tilde{y}(\cdot)$ (l'erreur de prédiction), respectivement.

Le prédicteur du système (1.40) au temps t s'écrit alors :

$$\hat{y}_t = \theta^T \varphi_{t,\theta} \quad (1.44)$$

où

$$\varphi_{t,\theta}^T = (x_{1,t}, \dots, x_{n_\theta,t}) \quad (1.45)$$

L'erreur de prédiction est donné par :

$$\tilde{y}_t = y_t - \hat{y}_t \quad (1.46)$$

et le gradient du prédicteur par :

$$\psi_{t,\theta} = \frac{\partial \hat{y}_t}{\partial \theta_i} = [x_{1,t}, \dots, x_{n_\theta,t}]^T, \quad i = 1, \dots, n_\theta \quad (1.47)$$

lequel est un vecteur colonne de dimension n_θ .

Au début de ce chapitre, nous avons mentionné que ce modèle a été dérivé comme une extension du domaine linéaire au non linéaire par des modèles ARMA. De plus, il a été comparé en détail aux autres représentations de modèles non linéaires dans quelques travaux [Chen89a]. Un algorithme de prédiction d'erreur récursif a été développé pour ce modèle et appliqué à des données de simulation et réels [Chen89b].

Une extension du cas linéaire au non linéaire a été effectuée avec un modèle ARMA (Autoregressive Moving Average) dans plusieurs publications. Dans une première approche les données mesurées de sortie et d'entrée ont été développées dans un modèle contenant des expansions polynomiales de la sortie et de l'entrée, plus des termes croisés :

$$\begin{aligned} \hat{y}_t = & \sum_{i_1=1}^{\infty} a_{i_1} y_{t-i_1} + \dots + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} a_{i_1, i_2} y_{t-i_1} y_{t-i_2} + \sum_{i_1=1}^{\infty} \dots \sum_{i_p=1}^{\infty} a_{i_1, \dots, i_p} y_{t-i_1} \dots y_{t-i_p} \\ & + \sum_{j_1=0}^{\infty} b_{j_1} u_{t-j_1} + \dots + \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} b_{j_1, j_2} u_{t-j_1} u_{t-j_2} + \sum_{j_1=0}^{\infty} \dots \sum_{j_q=0}^{\infty} b_{j_1, \dots, j_q} u_{t-j_1} \dots u_{t-j_q} \\ & + \sum_{i_1=1}^{\infty} \sum_{j_1=0}^{\infty} c_{i_1, j_1} y_{t-i_1} u_{t-j_1} + \dots + \sum_{i_1=1}^{\infty} \dots \sum_{i_p=1}^{\infty} \sum_{j_1=0}^{\infty} \dots \sum_{j_q=0}^{\infty} c_{i_1, \dots, i_p, j_1, \dots, j_q} y_{t-i_1} \dots y_{t-i_p} u_{t-j_1} \dots u_{t-j_q} \end{aligned} \quad (1.48)$$

où u_t, y_t, i_p, j_q sont respectivement l'entrée, la sortie, et les *ordres dynamiques* du système.

Les produits croisés entrée-sortie ont été ajoutés dans l'intention de réduire le nombre important des paramètres requis normalement pour calculer les noyaux dans une fonctionnelle polynomiale. Le critère à minimiser pour l'identification a été celui de la moyenne quadratique de l'erreur d'équation en dérivant un modèle multicanaux en treillis par algorithme de Levinson [Park81].

1.3.2.2 Représentation en Séries de Volterra

La représentation générale d'une série de Volterra de degré infini est donnée par l'équation :

$$y_t = \sum_{k=1}^{\infty} \sum_{\tau_1=1}^{\infty} \sum_{\tau_2=0}^{\infty} \cdots \sum_{\tau_k=0}^{\infty} W_k(t, \tau_1, \dots, \tau_k) u_{t-\tau_1} u_{t-\tau_2} \cdots u_{t-\tau_k} \quad (1.49)$$

où u_t est l'entrée et y_t la sortie, k est le *degré* de la non linéarité de la série.

Une approche pour identifier les noyaux d'une série de Volterra de degré deux a été développée en utilisant l'algorithme de Levinson [Kho85]. Ailleurs une technique rapide d'orthogonalisation a été développée pour calculer les noyaux de Volterra avec pour entrée un bruit blanc Gaussien [Koren91]. On a trouvé ainsi le développement des filtres de Volterra adaptatifs en treillis pour l'identification de systèmes non linéaires où l'algorithme utilisé a été construit à partir d'une entrée Gaussienne et d'une décomposition orthogonale par l'algorithme de Gram-Schmidt [Math96]. Toutes ces approches partent de la forme (1.48).

Dans une autre approche le développement en séries de Volterra de l'entrée Gaussienne du système a été utilisé. Les équations normales obtenues dans l'identification pour le calcul des noyaux ont été mises en termes des *cumulants* plutôt que des *moments*. D'après la formulation multicanaux présentée dans ce cas, l'algorithme des moindres carrés transversal avec une structure en treillis a été implémentée [Glen99].

Une décomposition orthogonale d'une entrée Gaussienne par des polynômes d'Hermite a été développée avec l'intention d'identifier les noyaux de Volterra par un modèle discret non linéaire de Wiener en partenariat avec l'implémentation du critère d'optimisation de moindres carrés adaptif. Cette décomposition orthogonale suppose améliorer la vitesse de convergence de l'algorithme d'optimisation et réduire le nombre de coefficients pour représenter les noyaux de Volterra [Ogun01].

Finalement nous avons trouvé dans les références la méthode d'identification « aveugle » des modèles de Volterra et Hammerstein. Avec cette méthode, même si on dispose seulement des données de sortie, deux cas ont été prévus : un dans lequel on dispose de la statistique de l'entrée et l'autre dans lequel on ne la connaît pas. Dans le premier cas, la statistique d'entrée est celle d'un bruit blanc, cependant sa densité de probabilité n'est pas connue avec précision. Dans le deuxième, il a été assumé qu'il s'agissait d'un bruit blanc Gaussien. Le calcul des noyaux de Volterra a été effectué par l'utilisation d'une forme linéaire multivariable en calculant les cumulants de la sortie par un produit de Kronecker [Kaloup05].

1.3.2.3 Représentation par un Modèle Bilinéaire Général

Le modèle bilinéaire général, étant une forme particulière du modèle NARMAX et ayant la possibilité d'approcher les noyaux de séries de Volterra, a été étudié et testé pour l'identification de systèmes non linéaires. Il est donné par l'équation :

$$y_t = \sum_{i=1}^{n_y} a_i y_{t-i} + \sum_{j=0}^{n_u} b_j u_{t-j} + \sum_{i=1}^{n_y} \sum_{j=0}^{n_u} c_{ij} y_{t-i} u_{t-j} \quad (1.50)$$

où u_t est l'entrée et y_t la sortie, n_y , n_u sont, respectivement, l'ordre du modèle pour l'entrée et la sortie.

Les travaux menés sur ce modèle ont été motivés par sa forte capacité à décrire des systèmes physiques réels ; il apparaît comme une simple extension d'un système linéaire au domaine non linéaire [Mohl80]. Une approche pour son identification a été d'estimer les noyaux de Volterra générés par un système bilinéaire pour obtenir les paramètres concernés par l'algorithme de moindres carrés [Inag84].

Par ailleurs, les méthodes récursives, utilisées dans le passé pour l'identification des systèmes linéaires, ont été étendues à ce modèle bilinéaire entrée-sortie. Ces extensions portent, par exemple, sur les méthodes des moindres carrés récursifs, des moindres carrés étendus, des variables instrumentales récursives, et de prédiction d'erreur récursive [Fna87]. Ces résultats ont été améliorés par l'introduction d'une méthode récursive robuste qui utilise les moindres carrés avec des poids modifiés dans le critère pour améliorer sa convergence [Dai89].

On a trouvé l'estimation des paramètres d'un modèle bilinéaire de séries temporelles, ayant aussi pour objet, l'estimation de la variance d'un processus d'excitation de moyenne nulle et Gaussien. Ceci a été effectué par le calcul des fonctions d'autocovariance et des cumulants du troisième ordre de la série temporelle [Math91]. Dans une autre approche de l'identification de systèmes bilinéaires, l'inférence Bayésienne a été utilisée pour calculer la fonction de densité de probabilité à posteriori des paramètres par moyen d'un échantillonner de Gibbs. Ce dernier, en générant un vecteur aléatoire de l'entrée-sortie du système, utilise la densité de probabilité conjointe pour ainsi obtenir la densité de probabilité de l'entrée [Medd98].

Un algorithme dit des « moindres carrés améliorés » a été développé pour l'estimation des coefficients du modèle bilinéaire par un algorithme adaptif des moindres carrés basé sur la méthode du gradient stochastique. Ici, la fonction de minimisation a été modifiée pour réduire le biais [Zhu99]. Un autre algorithme a été implémenté au moyen des moindres carrés. Avec celui-ci, l'estimation de la variance du bruit a été employée pour compenser le biais trouvé dans l'algorithme classique de moindres carrés récursifs. De cette façon l'inversion des matrices n'est plus requise [Medd00].

Comme il a été mentionné dans la section 1.3.4, les cumulants ont été utilisés pour l'identification des noyaux de Volterra. Ceci a été aussi le cas pour les systèmes bilinéaires prenant comme entrée un processus indépendant de distribution uniforme, où les cumulants croisés entre la sortie et l'entrée du système ont été calculés pour obtenir les coefficients du modèle. Une comparaison a été effectuée avec l'approche d'erreur de sortie pour montrer que le calcul des cumulants permet d'éviter les désavantages dans l'utilisation de la forme d'erreur de sortie, comme les minima locaux et la dépendance dans les conditions initiales avec un bruit de mesure Gaussien [Tsou01].

Le cas de l'identification « aveugle » a été aussi étudié pour les systèmes bilinéaires avec le calcul des cumulants à partir de la sortie du système pour une entrée assumée statistiquement comme un bruit blanc non Gaussien d'ordre élevé. L'identification a été développée en cinq étapes et il y a été souligné le besoin de bonnes conditions expérimentales nécessaires pour le design de l'entrée assumée d'ordre élevé [Kaloup03].

2. Structures Polynomiales

2.1 Introduction

Une structure polynomiale permet d'exprimer le signal de sortie d'un système quelconque comme la fonctionnelle des signaux d'entrée et sortie élevées à une puissance donnée, ainsi que des produits croisés entre ces signaux. Le degré résultant dépend de la structure de la représentation employée, ainsi que de l'ordre dynamique lié aux échantillons enregistrés en mémoire au fil du temps [Haber99]. Comme il a été exprimé au début de ce manuscrit, les séries de Volterra offrent une représentation importante pour l'analyse des systèmes non linéaires, surtout quand il est possible de réduire le nombre des paramètres concernés dans les noyaux du polynôme en dimension finie.

Dans ce chapitre, nous analyserons les moyens de réaliser des séries de Volterra dont nous verrons qu'elles ne sont réalisables en dimension finie que si les noyaux sont séparables. Dans cette hypothèse, elles peuvent être réalisées par des systèmes bilinéaires. Nous nous attarderons alors sur les systèmes bilinéaires homogènes (nilpotents) qui permettent de réaliser des séries de Volterra de degré fini (à noyaux séparable) et étudierons les formes canoniques possibles pour ces systèmes. Parmi celles-ci, une forme privilégiée sera retenue pour la réalisation et l'identification des systèmes non linéaires pour sa linéarité en sortie vis-à-vis des paramètres. Enfin, nous développerons une forme NARMA multi-degré qui possède cette propriété dans le cas quadratique.

2.2 Définition d'une série de Volterra

Commençons cette section en donnant l'expression générale entrée-sortie d'un système *dynamique* causal linéaire stationnaire à entrée monodimensionnelle :

$$y_t = \sum_{\tau=0}^{\infty} W(\tau) u_{t-\tau} \quad (2.1)$$

où y_t , u_t sont respectivement l'entrée et sortie du système, et $W(\tau)$ représente le noyau ou fonction de poids du système linéaire. A noter que cette fonction s'apparente à la notion classique de réponse impulsionnelle du système.

A l'opposé, dans le cas d'un système *statique* à non linéarité « faible », la relation entrée-sortie peut être exprimée par une série de Taylor sous la forme suivante :

$$y_t = w_0 + w_1 u_t + w_2 (u_t)^2 + \dots = \sum_{i=0}^{\infty} w_i (u_t)^i \quad (2.2)$$

Un système *dynamique* stationnaire à non linéarité faible peut donc *a priori* être approximé, au voisinage d'un point de fonctionnement, par une série de Volterra. Celle-ci se représente comme la somme de convolutions multidimensionnelles ayant la forme suivante :

$$\begin{aligned} y_t &= \sum_{\tau_1=0}^{\infty} W^1(\tau_1) u_{t-\tau_1} + \sum_{\tau_1=0}^{\infty} \sum_{\tau_2=0}^{\infty} W^2(\tau_1, \tau_2) u_{t-\tau_1} u_{t-\tau_2} + \dots \\ &= \sum_{k=1}^{\infty} \sum_{\tau_1=0}^{\infty} \sum_{\tau_2=0}^{\infty} \dots \sum_{\tau_k=0}^{\infty} W^k(\tau_1, \tau_2, \dots, \tau_k) u_{t-\tau_1} u_{t-\tau_2} \dots u_{t-\tau_k} \end{aligned} \quad (2.3)$$

La série peut-être décomposée comme suit :

$$y_t = V_t^1(u_{\bullet}) + V_t^2(u_{\bullet}) + \dots = \sum_{i=1}^{\infty} V_t^i(u_{\bullet}) \quad (2.4)$$

Chaque terme de cette série est dit *homogène*, ce qui signifie que si l'entrée est multipliée par un coefficient α scalaire, αu_t , alors la sortie sera multipliée par α^d , $\alpha^d y_t$. L'expression d'un terme homogène de degré k est donc la suivante :

$$V_t^k(u_{\bullet}) = \sum_{\tau_1=0}^{\infty} \sum_{\tau_2=0}^{\infty} \dots \sum_{\tau_k=0}^{\infty} W^k(\tau_1, \tau_2, \dots, \tau_k) u_{t-\tau_1} u_{t-\tau_2} \dots u_{t-\tau_k} \quad (2.5)$$

Cette série peut être réécrite, après changement de variable, comme suit :

$$\mathbf{V}_t^k(u_\bullet) = \sum_{\tau_1=-\infty}^t \sum_{\tau_2=-\infty}^{\tau_1} \cdots \sum_{\tau_k=-\infty}^{\tau_{k-1}} W^k(t-\tau_1, t-\tau_2, \dots, t-\tau_k) u_{\tau_1} u_{\tau_2} \cdots u_{\tau_k} \quad (2.6)$$

En modifiant la définition du noyau, et en tenant compte de la commutativité du produit $u_{\tau_1} u_{\tau_2} \cdots u_{\tau_k}$ (les noyaux de la série ont le même coefficient pour toute permutation des instants $\tau_1, \dots, \tau_k$), on peut réécrire cette série comme une somme sur un « hyper triangle » où on impose arbitrairement $\tau_1 \geq \tau_2 \geq \dots \geq \tau_k$ comme suit :

$$\mathbf{V}_t^k(u_\bullet) = \sum_{\tau_1=-\infty}^t \sum_{\tau_2=-\infty}^{\tau_1} \cdots \sum_{\tau_k=-\infty}^{\tau_{k-1}} \tilde{W}^k(t, \tau_1, \tau_2, \dots, \tau_k) u_{\tau_1} u_{\tau_2} \cdots u_{\tau_k} \quad (2.7)$$

Si on considère de telles séries très générales, le nombre de paramètres qui décrit celle-ci est évidemment prohibitif pour pouvoir être traité d'un point de vue pratique. Aussi, il est important de signaler un cas particulier de séries homogènes, où le noyau est dit séparable s'il peut être développé comme une somme finie de produits de fonctions d'une seule variable. Cela signifie que la i -ème composante du noyau W^k est alors définie par (en notation vectorielle) :

$$W^k(t, \tau_1, \dots, \tau_k)_i = \sum_{i_k=1}^{n_k} \cdots \sum_{i_1=1}^{n_1} [K_t^{k+1}]_i^{i_k} [K_{\tau_k}^k]_{i_k}^{i_{k-1}} \cdots [K_{\tau_2}^2]_{i_2}^{i_1} [K_{\tau_1}^1]_{i_1} \quad (2.8)$$

où $n_1, \dots, n_k$ sont respectivement les ordres de séparabilité en $(\tau_1, \tau_2), \dots, (\tau_k, t)$. Cette hypothèse de séparabilité est essentielle pour assurer la réalisabilité en dimension finie de cette série. Par exemple, pour $k=2$, on a :

$$W^2(t, \tau_1, \tau_2)_i = \begin{bmatrix} \kappa_{1,1}^3 & \cdots & \kappa_{1,i_2}^3 \\ \vdots & & \vdots \\ \kappa_{i,1}^3 & \cdots & \kappa_{i,i_2}^3 \end{bmatrix}_t \begin{bmatrix} \kappa_{1,1}^2 & \cdots & \kappa_{1,i_1}^2 \\ \vdots & & \vdots \\ \kappa_{i_2,1}^2 & \cdots & \kappa_{i_2,i_1}^2 \end{bmatrix}_{\tau_1} \begin{bmatrix} \kappa_{1,1}^1 & \cdots & \kappa_{1,n_1}^3 \\ \vdots & & \vdots \\ \kappa_{i_1,1}^1 & \cdots & \kappa_{i_1,n_1}^3 \end{bmatrix}_{\tau_2}$$

où $\kappa_{i,i_2}^3, \kappa_{i_2,i_1}^2, \kappa_{i_1,n_1}^1$ sont les variables de la dynamique de sortie, quadratique, et d'entrée respectivement, voir [Rugh81, p. 150] et [Lamna94, p. 124]. Ainsi, on peut observer clairement à la sortie la i -ème composante du noyau.

2.3 Réalisation d'une Série de Volterra à Noyaux Séparables par un Système Bilinéaire

2.3.1 Série de degré infini

Considérons le système bilinéaire mono-entrée/mono-sortie suivant

$$\begin{aligned} x_t &= F_t x_{t-1} + G_t x_{t-1} u_t + g_t u_t \\ y_t &= h_t x_t \end{aligned} \quad (2.9)$$

A noter que la présence du terme linéaire $g_t u_t$ permet d'éviter au système d'être dégénéré. En effet si celui-ci était absent, pour une condition initiale nulle, le système resterait à zéro. De même, si le système passe par une valeur nulle, rien ne peut lui faire quitter cette valeur.

Par ailleurs, il est important de noter que la forme utilisée ici s'exprime par

$$\begin{aligned} x_t &= f(x_{t-1}, u_t) \\ y_t &= h(x_t) \end{aligned} \quad (2.10)$$

contrairement au cas classique représenté par

$$\begin{aligned} x_t &= f(x_{t-1}, u_{t-1}) \\ y_t &= h(x_t, u_t) \end{aligned} \quad (2.11)$$

Si on réalise le changement de variable défini par :

$$z_t = \phi_F(0, t) x_t$$

où $\phi_F(0, t)$ désigne la matrice de transition associée à la matrice F_t :

$$\phi_F(t, \tau) = F_t \phi_F(t-1, \tau) = \phi_F(t, \tau+1) F_{\tau}$$

on obtient pour le nouveau système l'expression suivante :

$$\begin{aligned} z_t &= \phi_F(0, t) \left(\overbrace{F_t \phi_F(t-1, 0)}^{\phi_F(t, 0)} z_{t-1} + G_t \phi_F(t-1, 0) z_{t-1} u_t + g_t u_t \right) \\ &= z_{t-1} + \phi_F(0, t) g_t u_t + \phi_F(0, t) G_t \phi_F(t-1, 0) z_{t-1} u_t \end{aligned}$$

en utilisant la propriété des matrices de transition suivante

$$\phi_F(t, \tau)^{-1} = \phi_F(\tau, t)$$

Si on somme sur l'horizon $]-\infty, t]$ ce nouveau système afin de s'abstraire de l'effet d'une condition initiale non nulle, on obtient :

$$z_t = \sum_{\tau=-\infty}^t \phi_F(0, \tau) g_{\tau} u_{\tau} + \sum_{\tau=-\infty}^t \phi_F(0, \tau) G_{\tau} \phi_F(\tau-1, 0) z_{\tau-1} u_{\tau}$$

Si on applique à nouveau cette égalité dans la somme sur $z_{\tau-1}$, on peut récrire :

$$z_t = \sum_{\tau_1=-\infty}^t \phi_F(0, \tau_1) g_{\tau_1} u_{\tau_1} + \sum_{\tau_1=1}^t \phi_F(0, \tau_1) G_{\tau_1} \phi_F(\tau_1-1, 0) \left(\sum_{\tau_2=-\infty}^{\tau_1-1} \phi_F(0, \tau_2) g_{\tau_2} u_{\tau_2} + \sum_{\tau_2=-\infty}^{\tau_1-1} \phi_F(0, \tau_2) G_{\tau_2} \phi_F(\tau_2-1, 0) z_{\tau_2-1} u_{\tau_2} \right)$$

En itérant ce processus indéfiniment (intégration de Picard), et en utilisant l'expression de la sortie en fonction de z_t ($y_t = H_t \phi_F(t, 0) z_t$), on montre alors que la sortie du système peut s'écrire comme la série de Volterra de degré infini suivante :

$$y_t = \sum_{k=1}^{\infty} \sum_{\tau_1=-\infty}^t \sum_{\tau_2=-\infty}^{\tau_1} \cdots \sum_{\tau_k=-\infty}^{\tau_{k-1}} W^k(t, \tau_1, \tau_2, \dots, \tau_k) u_{\tau_1} u_{\tau_2} \cdots u_{\tau_k} \quad (2.12)$$

où les noyaux sont par définition séparables d'ordre de la dimension du vecteur d'état et sont définis par l'expression suivante :

$$W^k(t, \tau_1, \tau_2, \dots, \tau_k) = h_t \phi_F(t, \tau_1) G_{\tau_1} \phi_F(\tau_1-1, 0) \dots \phi_F(0, \tau_k) G_{\tau_k} \phi_F(\tau_k-1, 0) g_{\tau_k} \quad (2.13)$$

A noter que dans cette représentation, les ordres de séparabilité des noyaux sont tous égaux à la dimension de l'état x_t .

2.3.2 Série homogène (de degré fini)

Chaque terme de degré homogène d de la série de Volterra à noyaux séparables peut être obtenu par le système bilinéaire (nilpotent) suivant [Brock76a,b,c] :

$$\begin{aligned} x_t^1 &= A_t^1 x_{t-1}^1 + b_t u_t \\ x_t^2 &= A_t^2 x_{t-1}^2 + D_t^1 x_t^1 u_t \\ &\dots \\ x_t^d &= A_t^d x_{t-1}^d + D_t^{d-1} x_t^{d-1} u_t \\ y_t^d &= c_t x_t^d \end{aligned} \quad (2.14)$$

où chaque x_t^k est un vecteur réel de dimension n_k . A noter que cette représentation est en quelque sorte implicite dans la mesure où dans le second membre apparaissent à la fois x_{t-1}^d et x_t^{d-1} . Cependant, cette écriture permet d'être davantage synthétique qu'une écriture plus lourde équivalente sous la forme suivante :

$$\begin{aligned}
x_t^1 &= A_t^1 x_{t-1}^1 + b_t u_t \\
x_t^2 &= A_t^2 x_{t-1}^2 + D_t^1 A_t^1 x_{t-1}^1 u_t + D_t^1 b_t u_t \\
&\dots \\
x_t^d &= A_t^d x_{t-1}^d + D_t^{d-1} A_t^{d-1} x_{t-1}^{d-1} u_t + D_t^{d-1} D_t^{d-2} A_t^{d-2} x_{t-1}^{d-2} (u_t)^2 \dots + D_t^{d-1} D_t^{d-2} \dots D_t^1 b_t (u_t)^d \\
y_t^d &= c_t x_t^d
\end{aligned} \tag{2.15}$$

Grâce à cette définition, chaque terme x_t^i est un terme de degré i en l'entrée u_t . Plus précisément, en faisant la somme du premier terme (linéaire), on obtient l'expression classique :

$$x_t^1 = \sum_{\tau_1=-\infty}^t \phi^1(t, \tau_1) b_{\tau_1} u_{\tau_1}$$

Le terme de degré deux x_t^2 , peut alors être vu comme un système linéaire dont l'entrée est égale à $x_{t-1}^1 u_t$. En faisant la somme du terme de degré deux, on obtient :

$$\begin{aligned}
x_t^2 &= \sum_{\tau_1=-\infty}^t \phi^2(t, \tau_1) D_{\tau_1}^1 x_{\tau_1}^1 u_{\tau_1} \\
&= \sum_{\tau_1=-\infty}^t \phi^2(t, \tau_1) D_{\tau_1}^1 \sum_{\tau_2=-\infty}^{\tau_1} \phi^1(\tau_1, \tau_2) b_{\tau_2} u_{\tau_1} u_{\tau_2}
\end{aligned}$$

et de manière générale, en itérant ce processus :

$$\begin{aligned}
x_t^k &= \sum_{\tau_1=-\infty}^t \phi^k(t, \tau_1) D_{\tau_1}^k x_{\tau_1}^{k-1} u_{\tau_1} \\
&= \sum_{\tau_1=-\infty}^t \sum_{\tau_2=-\infty}^{\tau_1} \dots \sum_{\tau_k=-\infty}^{\tau_{k-1}} \phi^k(t, \tau_1) D_{\tau_1}^{k-1} \dots D_{\tau_{k-1}}^1 \phi^1(\tau_{k-1}, \tau_k) b_{\tau_k} u_{\tau_1} u_{\tau_2} \dots u_{\tau_k}
\end{aligned} \tag{2.16}$$

Ceci conduit finalement à l'écriture de la sortie sous la forme suivante :

$$y_t^d = \sum_{\tau_1=-\infty}^t \sum_{\tau_2=-\infty}^{\tau_1} \dots \sum_{\tau_d=-\infty}^{\tau_{d-1}} W^d(t, \tau_1, \tau_2, \dots, \tau_d) u_{\tau_1} u_{\tau_2} \dots u_{\tau_d} \tag{2.17}$$

où le noyau est défini par :

$$W^d(t, \tau_1, \tau_2, \dots, \tau_k) = c_t \phi^d(t, \tau_1) D_{\tau_1}^{k-1} \phi^{k-1}(\tau_1, \tau_2) \dots D_{\tau_{k-1}}^1 \phi(\tau_{k-1}, \tau_k) b_{\tau_k} \tag{2.18}$$

Il est donc clair qu'une série de Volterra de degré fini d peut s'obtenir comme la somme de termes de ce type de degré inférieur ou égal à d .

2.4 Transformées en z et fonctions de transfert multivariables

Dans le contexte des systèmes bilinéaires, la transformée en z doit être définie comme étant multi-dimensionnelle. En effet, la transformée en z doit avoir la propriété d'un opérateur linéaire. Or, dans les équations d'évolution des systèmes bilinéaires certaines variables apparaissent sous la forme de produit, ce qui donnerait lieu à une convolution entre leurs transformées en z ce qui lui ferait perdre son caractère d'opérateur linéaire.

Pour définir une telle transformée qui permette de donner une représentation entrée/sortie des systèmes bilinéaires sous la forme d'une fonction de transfert, il est nécessaire de définir un nouvel état multivariable de la manière suivante. A noter que nous utiliserons ici une approche « séries formelles » sans aborder les éventuels problèmes de convergence. Définissons l'ensemble des états suivants :

$$\begin{aligned}\tilde{x}_{\tau_1}^1 &= A^1 \tilde{x}_{\tau_1-1}^1 + bu_{\tau_1} \\ \tilde{x}_{\tau_1, \dots, \tau_k}^k &= A^k \tilde{x}_{\tau_1-1, \dots, \tau_k-1}^k + D^{k-1} \tilde{x}_{\tau_k, \dots, \tau_{k-1}}^{k-1} u_{\tau_k}, \forall k = 2 \dots d \\ \tilde{y}_{\tau_1, \dots, \tau_d}^d &= c \tilde{x}_{\tau_1, \dots, \tau_d}^d\end{aligned}\quad (2.19)$$

où les $\tau_1, \dots, \tau_d$ sont des temps indépendants introduits artificiellement. Par définition, les composantes de degré k x_t^k de l'état et la sortie y_t^d s'obtiennent par la formule suivante :

$$\begin{aligned}x_t^k &= \tilde{x}_{\tau_1, \dots, \tau_k}^k \mid_{\tau_1 = \dots = \tau_k = t}, \forall k = 1 \dots d \\ y_t^d &= c \tilde{x}_{\tau_1, \dots, \tau_d}^d \mid_{\tau_1 = \dots = \tau_d = t}\end{aligned}\quad (2.20)$$

Notons $Z(\bullet)$ l'opérateur transformée en z . La transformée en z de l'entrée u_t est alors définie par l'expression suivante:

$$\begin{aligned}Z(u) &= \sum_{\tau=-\infty}^t u_{\tau} z^{\tau-t} \\ &= u_t + u_{t-1} z^{-1} + \dots + u_{t-i} z^{-i} + \dots\end{aligned}\quad (2.21)$$

et la transformée en z multivariable de $\tilde{x}_{\tau_1, \dots, \tau_k}^k$ par :

$$Z(\tilde{x}_{\tau_1, \dots, \tau_k}^k) = \sum_{\tau_1=-\infty}^t \dots \sum_{\tau_k=-\infty}^t \tilde{x}_{\tau_1, \dots, \tau_k}^k z_1^{\tau_1-t} \dots z_k^{\tau_k-t} \quad (2.22)$$

où $z_1, \dots, z_k$ sont des variables formelles. L'opérateur transformée en z multivariable est clairement un opérateur linéaire et l'opération « multiplier par $z_i^{-\tau}$ » s'interprète alors comme l'opérateur retard de τ sur la variable τ_i . En effet, avec une telle définition, il est clair que :

$$\begin{aligned}
Z(\tilde{x}_{\tau_1-\tau, \dots, \tau_k-\tau}^k) &= \sum_{\tau_1=-\infty}^t \dots \sum_{\tau_k=-\infty}^t \tilde{x}_{\tau_1-\tau, \dots, \tau_k-\tau}^k z_1^{\tau_1-\tau} \dots z_k^{\tau_k-\tau} \\
&= \sum_{\tau_1=-\infty}^t \dots \sum_{\tau_k=-\infty}^t \tilde{x}_{\tau_1, \dots, \tau_k}^k z_1^{\tau_1-\tau} \dots z_k^{\tau_k-\tau}
\end{aligned} \tag{2.23}$$

Par conséquent, on a :

$$\begin{aligned}
Z(\tilde{x}_{\tau_1-\tau, \dots, \tau_k-\tau}^k) &= z_1^{-\tau} \dots z_k^{-\tau} \sum_{\tau_1=-\infty}^t \dots \sum_{\tau_k=-\infty}^t \tilde{x}_{\tau_1, \dots, \tau_k}^k z_1^{\tau_1} \dots z_k^{\tau_k} \\
&= z_1^{-\tau} \dots z_k^{-\tau} Z(\tilde{x}_{\tau_1, \dots, \tau_k}^k)
\end{aligned} \tag{2.24}$$

A noter que la multiplication par un polynôme quelconque en $z_1, \dots, z_k$ ne peut s'interpréter par un retard que si ce polynôme est homogène en le monôme $z_1 \dots z_k$.

Soient

$$\begin{aligned}
U(z) &= Z(u) \\
X^1(z_1, \dots, z_k) &= Z(\tilde{x}_{\tau_1, \dots, \tau_k}^k)
\end{aligned}$$

les transformées en z ainsi définies de l'entrée u_t et de l'état $\tilde{x}_{\tau_1, \dots, \tau_k}^k$. Si on applique cette propriété du retard au système multivariable défini plus haut, on obtient l'écriture formelle suivante :

$$\begin{aligned}
X^1(z_1) &= z_1^{-1} A^1 X^1(z_1) + b U(z_1) \\
X^k(z_1, \dots, z_k) &= z_1^{-1} \dots z_k^{-1} (A^k X^k(z_1, \dots, z_k)) + D^{k-1} X^{k-1}(z_1, \dots, z_{k-1}) U(z_k), \forall k = 1 \dots d
\end{aligned} \tag{2.25}$$

Ainsi, en multipliant ces équations respectivement par $z_1, z_1 z_2, \dots, z_1 \dots z_k$, on obtient les fonctions de transfert de $X^1(z_1), \dots, X^d(z_1, \dots, z_d)$ sous la forme suivante :

$$\begin{aligned}
X^1(z_1) &= (z_1 I^1 - A^1)^{-1} b z_1 U(z_1) \\
X^k(z_1, \dots, z_k) &= (z_1 \dots z_k I^k - A^k)^{-1} D^{k-1} X^{k-1}(z_1, \dots, z_{k-1}) z_1 \dots z_k U(z_k), \forall k = 1 \dots d \\
Y^d(z_1, \dots, z_d) &= c X^d(z_1, \dots, z_d)
\end{aligned} \tag{2.26}$$

où I^k désigne la matrice identité de dimension n_k . Pour simplifier les écritures, notons le produit

$$\tilde{z}_k = z_1 \dots z_k \tag{2.27}$$

Avec cette notation, les fonctions de transfert s'écrivent alors :

$$\begin{aligned}
X^1(z_1) &= (z_1 I^1 - A^1)^{-1} z_1 b U(z_1) \\
X^k(z_1, \dots, z_k) &= (\tilde{z}_k I^k - A^k)^{-1} \tilde{z}_k D^{k-1} X^{k-1}(z_1, \dots, z_{k-1}) U(z_k), \forall k = 1 \dots d
\end{aligned} \tag{2.28}$$

En substituant chaque $X^k(z_1, \dots, z_k)$ à son expression dans celle de la sortie $Y^d(z_1, \dots, z_d)$, on obtient la fonction de transfert entrée/sortie du système bilinéaire homogène sous la forme suivante :

$$Y^d(z_1, \dots, z_d) = c(\tilde{z}_d I^d - A^d)^{-1} \tilde{z}_d D^{d-1} (\tilde{z}_{d-1} I^{d-1} - A^{d-1})^{-1} \tilde{z}_{d-1} D^{d-2} \dots (\tilde{z}_2 I^2 - A^2)^{-1} \tilde{z}_2 D^1 (\tilde{z}_1 I^1 - A^1)^{-1} z_1 b U(z_1) \dots U(z_d) \quad (2.29)$$

2.5 Formes canoniques des systèmes bilinéaires homogènes

Comme il apparaîtra par la suite, la résolution des problèmes de réalisation et d'identification des systèmes bilinéaires dont la série de Volterra est de degré fini impose deux contraintes. La première est que la représentation paramétrée de ces systèmes soit sous une forme canonique dans le sens où le nombre de paramètres de la représentation doit être minimal afin d'éviter une indétermination des paramètres calculés. Bien que cette condition ne soit pas strictement indispensable (possibilité d'utiliser des pseudo-inversion à la place des strictes inversions), elle assure un bon conditionnement des algorithmes. En second lieu, la paramétrisation doit être linéaire en la sortie observée du système afin de donner lieu à un problème convexe dans la démarche de minimisation de l'erreur de sortie.

2.5.1 Forme observateur

La forme canonique observateur peut s'obtenir par la proposition suivante :

Proposition 1 : Le système bilinéaire homogène défini par l'équation (2.14) possède la forme canonique observateur suivante :

$$\begin{aligned}x_t^1 &= A^1 x_{t-1}^1 + b u_t \\x_t^2 &= A^2 x_{t-1}^2 + D^1 x_t^1 u_t \\&\dots \\x_t^d &= A^d x_{t-1}^d + D^{d-1} x_t^{d-1} u_t \\y_t^d &= c x_t^d\end{aligned}$$

où

$$b = [b_1 \quad \dots \quad b_{n_1}]^T \quad (2.30)$$

$$A^k = \begin{bmatrix} a_1^k & 1 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & & \ddots & 1 \\ a_{n_k}^k & 0 & \dots & \dots & 0 \end{bmatrix}, \forall k = 1 \dots d \quad (2.31)$$

$$D^k = \begin{bmatrix} 1 & 0 & \dots & 0 \\ d_{2,1}^k & \dots & \dots & d_{2,n_{k-1}}^k \\ \vdots & & & \vdots \\ d_{n_k,1}^k & \dots & \dots & d_{n_k,n_{k-1}}^k \end{bmatrix}, \forall k = 1 \dots d-1 \quad (2.32)$$

$$c = [1 \quad 0 \quad \dots \quad 0] \quad (2.33)$$

Preuve : Il a été établi plus haut que la fonction de transfert multivariable de ces systèmes pouvait être écrite de la manière suivante :

$$H(z_1, \dots, z_d) = c \left(\tilde{z}_d I^d - A^d \right)^{-1} \tilde{z}_d D^{d-1} \left(\tilde{z}_{d-1} I^{d-1} - A^{d-1} \right)^{-1} \tilde{z}_{d-1} D^{d-2} \dots \left(\tilde{z}_2 I^2 - A^2 \right)^{-1} \tilde{z}_2 D^1 \left(\tilde{z}_1 I^1 - A^1 \right)^{-1} z_1 b$$

Par ailleurs, le théorème de Cayley-Hamilton permet d'écrire l'équation caractéristique de chaque matrice sous la forme suivante :

$$(A^k)^{n_k} = \sum_{i=0}^{n_k-1} a_{i+1}^k (A^k)^i, \forall k = 1 \dots d$$

Considérons à présent la matrice $P^d \in \mathbb{R}^{n_d \times n_d}$ de changement de variable utilisée dans le contexte des systèmes linéaires pour obtenir leur forme canonique observateur. Celle-ci est définie par :

$$P^d = \begin{bmatrix} c \\ c(A^d - a_n^d I^d) \\ \vdots \\ c((A^d)^{n-1} + a_n^d (A^d)^{n-2} + \dots a_3^d A^d + a_2^d I^d) \end{bmatrix}$$

Si la paire (c, A^d) est observable, c'est-à-dire si le rang de la matrice

$$O^d = \begin{bmatrix} c \\ cA^d \\ \vdots \\ c(A^d)^{n_d-1} \end{bmatrix}$$

est égal à la dimension n_d du vecteur x_t^d , alors la matrice P^d est inversible. Cette matrice, par construction, possède les propriétés suivantes :

$$P^d A^d = \tilde{A}^d P^d \\ \tilde{C} P^d = c$$

où

$$\tilde{A}^d = \begin{bmatrix} a_1^d & 1 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & & \ddots & 1 \\ a_{n_d}^d & 0 & \dots & \dots & 0 \end{bmatrix}$$

$$\tilde{C} = [1 \quad 0 \quad \dots \quad 0]$$

Si on insère $(P^d)^{-1} P^d = I^d$ dans l'expression de la fonction de transfert $H(z_1, \dots, z_d)$, on obtient :

$$\begin{aligned}
c(\tilde{z}_d I^d - A^d)^{-1} D^{d-1} &= c(P^d)^{-1} P^d (\tilde{z}_d I^d - A^d)^{-1} (P^d)^{-1} P^d D^{d-1} \\
&= \tilde{c}(\tilde{z}_d I^d - \tilde{A}^d)^{-1} P^d D^{d-1}
\end{aligned}$$

et la fonction de transfert peut-être réécrite

$$H(z_1, \dots, z_d) = \tilde{c}(\tilde{z}_d I^d - \tilde{A}^d)^{-1} P^d D^{d-1} (\tilde{z}_{d-1} I^{d-1} - A^{d-1})^{-1} D^{d-2} \dots D^1 (\tilde{z}_1 I^1 - A^1)^{-1} b$$

Considérons la nouvelle paramétrisation $D'^{d-1} = P^d D^{d-1}$ qui peut se décomposer en lignes selon l'expression suivante :

$$D'^{d-1} = \begin{bmatrix} D_1'^{d-1} \\ D_2'^{d-1} \\ \vdots \\ D_{n_d}'^{d-1} \end{bmatrix}$$

Soit à présent $P^{d-1} \in \mathbb{R}^{n_{d-1} \times n_{d-1}}$ associé à la paire $\{D_1'^{d-1}, A^{d-1}\}$ définie par :

$$P^{d-1} = \begin{bmatrix} D_1'^{d-1} \\ D_1'^{d-1} (A^{d-1} - a_{n_{d-1}}^{d-1} I^{d-1}) \\ \vdots \\ D_1'^{d-1} \left((A^{d-1})^{n_{d-1}-1} + a_{n_{d-1}}^{d-1} (A^{d-1})^{n_{d-1}-2} + \dots + a_3^{d-1} A^{d-1} + a_2^{d-1} I^{d-1} \right) \end{bmatrix}$$

A nouveau, cette matrice possède les propriétés suivantes :

$$\begin{aligned}
P^{d-1} A^{d-1} &= \tilde{A}^{d-1} P^{d-1} \\
\tilde{c} P^{d-1} &= D_1'^{d-1}
\end{aligned}$$

Par conséquent, on a :

$$\begin{bmatrix} \tilde{c} \\ D_2'^{d-1} (P^{d-1})^{-1} \\ \vdots \\ D_{n_d}'^{d-1} (P^{d-1})^{-1} \end{bmatrix} P^{d-1} = \begin{bmatrix} D_1'^{d-1} \\ D_2'^{d-1} \\ \vdots \\ D_{n_d}'^{d-1} \end{bmatrix} = D'^{d-1}$$

Si on définit

$$\tilde{D}^{d-1} = \begin{bmatrix} \tilde{c} \\ D_2'^{d-1} (P^{d-1})^{-1} \\ \vdots \\ D_{n_d}'^{d-1} (P^{d-1})^{-1} \end{bmatrix}$$

alors la fonction de transfert peut s'exprimer selon :

$$H(z_1, \dots, z_d) = \tilde{c}(\tilde{z}_d I - \tilde{A}^d)^{-1} \tilde{D}^{d-1} (\tilde{z}_{d-1} I - \tilde{A}^{d-1})^{-1} D'^{d-2} \dots D^1 (\tilde{z}_1 I - A^1)^{-1} \tilde{b}$$

où D^{d-2} désigne la matrice $P^{d-1}D^{d-2}$. En itérant cette procédure jusqu'à la partie linéaire, on obtient :

$$H(z_1, \dots, z_d) = \tilde{c}(\tilde{z}_d I - \tilde{A}^d)^{-1} \tilde{D}^{d-1}(\tilde{z}_{d-1} I - \tilde{A}^{d-1})^{-1} \tilde{D}^{d-2} \dots \tilde{D}^1(\tilde{z}_1 I - \tilde{A}^1)^{-1} \tilde{b}$$

avec $\tilde{b} = P_1 b$, ce qui est bien la forme annoncée. ■

2.5.2 Forme contrôleur

Proposition 2 : Le système bilinéaire homogène défini par l'équation (2.14) possède la forme canonique contrôleur suivante :

$$\begin{aligned} x_t^1 &= A^1 x_{t-1}^1 + b u_t \\ x_t^2 &= A^2 x_{t-1}^2 + D^1 x_t^1 u_t \\ &\dots \\ x_t^d &= A^d x_{t-1}^d + D^{d-1} x_t^{d-1} u_t \\ y_t^d &= c x_t^d \end{aligned}$$

où

$$b = [1 \quad 0 \quad \dots \quad 0]^T \quad (2.34)$$

$$A^k = \begin{bmatrix} a_1^k & \dots & \dots & \dots & a_{n_k-1}^k \\ 1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{bmatrix}, \forall k = 1 \dots d \quad (2.35)$$

$$D^k = \begin{bmatrix} 1 & d_{1,1}^k & \dots & d_{1,n_k-1}^k \\ 0 & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & d_{n_k,1}^k & \dots & d_{n_k,n_k-1}^k \end{bmatrix}, \forall k = 1 \dots d-1 \quad (2.36)$$

$$c = [c_1 \quad \dots \quad c_n] \quad (2.37)$$

Preuve : Considérons à présent la matrice $P^1 \in \mathbb{R}^{n_1 \times n_1}$ utilisée classiquement pour dégager les formes canoniques de contrôlabilité dans le contexte des systèmes linéaires. Les équations caractéristiques des matrices A^k s'écrit :

$$(A^k)^{n_k} = \sum_{i=0}^{n_k-1} a_{i+1}^k (A^k)^i, \forall k = 1 \dots d$$

Celle-ci est définie par :

$$P^1 = \begin{bmatrix} b & (A^1 - a_1^1 I)b & \dots & ((A^1)^{n-1} - a_1^1 (A^1)^{n-2} - \dots - a_{n-1}^1 I)b \end{bmatrix}$$

et possède les propriétés suivantes :

$$P^1 A^1 = \tilde{A}^1 P^1$$

$$P^1 b = \tilde{b}$$

où

$$\tilde{A}^1 = \begin{bmatrix} a_1^1 & \dots & \dots & \dots & a_{n_1}^1 \\ 1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{bmatrix}$$

$$\tilde{b} = [1 \quad 0 \quad \dots \quad 0]^T$$

En insérant $(P^1)^{-1} P^1 = I^1$ dans l'expression de la fonction de transfert, on obtient :

$$\begin{aligned} D^1 (z_1 I - A^1)^{-1} b &= D^1 P^1 (P^1)^{-1} (z_1 I^1 - A^1)^{-1} P^1 (P^1)^{-1} b \\ &= D^1 P^1 \left((P^1)^{-1} (z_1 I^1 - A^1) P^1 \right)^{-1} (P^1)^{-1} b \\ &= D^1 P^1 (z_1 I^1 - \tilde{A}^1)^{-1} \tilde{b} \end{aligned}$$

et la fonction de transfert peut être réécrite sous la forme suivante :

$$H(z_1, \dots, z_d) = c (\tilde{z}_d I^1 - A^d)^{-1} \dots D^2 (\tilde{z}_2 I^2 - A^2)^{-1} D^1 (z_1 I^1 - \tilde{A}^1)^{-1} \tilde{b}$$

où $D'^1 = D^1 P^1$. Considérons cette nouvelle paramétrisation et décomposons cette matrice en vecteurs comme suit :

$$D'^1 = \begin{bmatrix} D_1'^1 & \dots & D_{n_1}'^1 \end{bmatrix}$$

Définissons à présent la matrice de changement de variable $P^2 \in \mathbb{R}^{n_2 \times n_2}$ associée à la paire $\{D_1'^1, A^2\}$. Cette matrice possède les propriétés suivantes :

$$P^2 A^2 = \tilde{A}^2 P^2$$

$$P^2 D'^1 = \begin{bmatrix} \tilde{b} & \tilde{D}_2^1 & \dots & \tilde{D}_{n_2}^1 \end{bmatrix}$$

où par définition $\tilde{D}_i^1 = P^2 D_i'^1, \forall i = 2 \dots n_2$. En itérant cette procédure, on obtient la forme canonique annoncée, soit :

$$H(z_1, \dots, z_d) = \tilde{c} (\tilde{z}_d I^d - \tilde{A}^d)^{-1} \tilde{D}^{d-1} (\tilde{z}_{d-1} I^{d-1} - \tilde{A}^{d-1})^{-1} \tilde{D}^{d-2} \dots \tilde{D}^1 (z_1 I^1 - \tilde{A}^1)^{-1} \tilde{b}$$

avec $\tilde{c} = cP^d$. ■

2.5.3 Forme hybride

Pour chaque matrice dynamique A^k , il est possible d'opter soit pour la forme canonique d'observabilité soit pour la forme canonique contrôleur. Cela signifie que dans la représentation sous la forme :

$$\begin{aligned}x_t^1 &= A_t^1 x_{t-1}^1 + b_t u_t \\x_t^2 &= A_t^2 x_{t-1}^2 + D_t^1 x_t^1 u_t \\&\dots \\x_t^d &= A_t^d x_{t-1}^d + D_t^{d-1} x_t^{d-1} u_t \\y_t^d &= c_t x_t^d\end{aligned}$$

les matrices A^k peuvent prendre l'une des formes suivantes, (2.31) ou (2.35):

$$A^k = \begin{bmatrix} a_1^k & 1 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & & \ddots & 1 \\ a_{n_k}^k & 0 & \cdots & \cdots & 0 \end{bmatrix} \quad \text{ou} \quad A^k = \begin{bmatrix} a_1^k & \cdots & \cdots & \cdots & a_{n_{k-1}}^k \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}$$

Si A^1 a la forme observateur, alors on a (2.30) :

$$b = [b_1 \quad \cdots \quad b_{n_1}]^T$$

sinon on a (2.34) :

$$b = [1 \quad 0 \quad \cdots \quad 0]^T$$

De même, si A^k a la forme observateur pour $k \geq 2$, et si A^{k-1} a aussi la forme observateur, alors on a (2.32) :

$$D^{k-1} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ d_{2,1}^{k-1} & \cdots & \cdots & d_{2,n_{k-1}}^{k-1} \\ \vdots & & & \vdots \\ d_{n_k,1}^{k-1} & \cdots & \cdots & d_{n_k,n_{k-1}}^{k-1} \end{bmatrix}$$

sinon

$$D^{k-1} = \begin{bmatrix} 1 & d_{1,1}^{k-1} & \cdots & d_{1,n_{k-1}}^{k-1} \\ 0 & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & d_{n_k,1}^{k-1} & \cdots & d_{n_k,n_{k-1}}^{k-1} \end{bmatrix} \quad (2.38)$$

Enfin, si A^d a la forme observateur, alors on a (2.33) :

$$c = [1 \quad 0 \quad \cdots \quad 0]$$

sinon on a (2.37) :

$$c = [c_1 \quad \cdots \quad c_n]$$

2.5.4 Forme linéaire en les paramètres

La mise en œuvre de l'algorithme héréditaire nécessite que les paramètres apparaissent linéairement dans l'expression de la sortie du système. Il est possible d'atteindre cet objectif simplement à condition d'imposer que les ordres de séparabilité de chaque sous-système soient égaux entre eux.

Proposition 3 : Le système bilinéaire homogène défini par l'équation (2.14) dont les ordres de séparabilité sont égaux à n possède la forme canonique suivante :

$$\begin{aligned} x_t^1 &= A^1 x_{t-1}^1 + b u_t \\ x_t^2 &= A^2 x_{t-1}^2 + x_t^1 u_t \\ &\dots \\ x_t^d &= A^d x_{t-1}^d + x_t^{d-1} u_t \\ y_t^d &= c x_t^d \end{aligned}$$

avec (2.30) :

$$b = [b_1 \quad \cdots \quad b_{n_1}]^T$$

$$A^k = \begin{bmatrix} a_{1,1}^k & \cdots & a_{1,n}^k \\ \vdots & & \vdots \\ a_{n,1}^k & \cdots & a_{n,n}^k \end{bmatrix}, \forall k = 1, \dots, d-1, \quad (2.39)$$

$$\tilde{A}^d = \begin{bmatrix} a_1^d & 1 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & & \ddots & 1 \\ a_{n_d}^d & 0 & \cdots & \cdots & 0 \end{bmatrix} \quad (2.40)$$

et (2.33) :

$$c = [1 \quad 0 \quad \cdots \quad 0]$$

Preuve : Considérons la forme canonique d'observabilité définie plus haut par :

$$\begin{aligned}x_t^1 &= A^1 x_{t-1}^1 + b u_t \\x_t^2 &= A^2 x_{t-1}^2 + D^1 x_t^1 u_t \\&\dots \\x_t^d &= A^d x_{t-1}^d + D^{d-1} x_t^{d-1} u_t \\y_t^d &= c x_t^d\end{aligned}$$

et supposons que les ordres de séparabilité n_i soient tous égaux à n . Supposons de plus que les matrices $D^k, \forall k = 1, \dots, d-1$ soient inversibles. Cette hypothèse n'est pas restrictive dans la mesure où la présence d'une matrice D^k singulière revient à considérer que $D^k x_t^k$ est équivalent à un vecteur de dimension réduite, égale au rang de D^k , et donc de modifier l'ordre de séparabilité.

Considérons la nouvelle variable d'état $\tilde{x}_t^{d-1} = D^{d-1} x_t^{d-1}$. Il est clair qu'alors :

$$x_t^d = A^d x_{t-1}^d + \tilde{x}_t^{d-1} u_t$$

L'équation d'évolution de $\tilde{x}_t^{d-1}$ s'écrit alors :

$$\tilde{x}_t^{d-1} = D^{d-1} A^{d-1} (D^{d-1})^{-1} \tilde{x}_{t-1}^{d-1} + D^{d-1} D^{d-2} x_t^{d-2} u_t$$

Définissons alors $\tilde{A}^{d-1} = D^{d-1} A^{d-1} (D^{d-1})^{-1}$ et le changement de variable $\tilde{x}_t^{d-2} = D^{d-1} D^{d-2} x_t^{d-2}$. On a alors :

$$\tilde{x}_t^{d-1} = \tilde{A}^{d-1} \tilde{x}_{t-1}^{d-1} + \tilde{x}_t^{d-2} u_t$$

En poursuivant ce processus, c'est-à-dire en définissant $\tilde{A}^k = D^k A^k (D^k)^{-1}$ et les changements de variable successifs $\tilde{x}_t^k = D^{d-1} \dots D^k x_t^k$, on obtient la forme désirée avec $\tilde{b} = D^{d-1} \dots D^1 b$. ■

Cette forme canonique atteint bien l'objectif souhaité.

Proposition 4 : Le système bilinéaire homogène défini par la proposition 3 fait apparaître linéairement les paramètres dans l'expression de la sortie.

Preuve : La forme canonique développée dans la proposition 3 permet de décomposer x_t^{d-1} sous la forme suivante :

$$\begin{aligned}
x_t^{d-1} &= A^{d-1}x_{t-1}^{d-1} + A^{d-2}x_{t-1}^{d-2}u_t + \dots + A^1x_{t-1}^1(u_t)^{d-2} + b(u_t)^{d-1} \\
&= \sum_{k=1}^{d-1} A^k x_{t-1}^k (u_t)^{d-k-1} + b(u_t)^{d-1}
\end{aligned}$$

La forme compagne de A^d permet d'écrire la sortie y_t comme :

$$y_t^d = \sum_{i=1}^n a_i^d y_{t-i}^d + \sum_{i=1}^n [x_{t-i+1}^{d-1}]_i u_{t-i+1}$$

où $[x_t^{d-1}]_i$ est par définition la i -ème composante du vecteur x_t^{d-1} . Si on décompose la matrice A^{d-1} en lignes sous la forme :

$$A^{d-1} = \begin{bmatrix} A_1^{d-1} \\ \vdots \\ A_n^{d-1} \end{bmatrix}$$

on peut récrire la sortie comme suit :

$$y_t^d = \sum_{i=1}^n a_i^d y_{t-i}^d + \sum_{k=1}^{d-1} \sum_{i=1}^n A_i^k x_{t-i}^k (u_{t-i+1})^{d-k-1} + \sum_{i=1}^n b_i (u_{t-i+1})^d$$

qui est bien linéaire vis-à-vis des paramètres du système. ■

2.6 Formes entrée/sorties multi-degré

2.6.1 Cas général

Considérons la forme canonique d'observabilité définie plus haut. L'objectif de cette partie est de donner une représentation entrée/sortie du système qui fasse apparaître les sorties canoniques de chacun des sous-systèmes de degré homogène de degré 1 à d . Pour ce faire, il suffit de remarquer que si on définit la sortie du k -ième système (scalaire) comme étant la première composante du vecteur x_t^k , soit :

$$y_t^k = c x_t^k$$

avec (2.33) :

$$\tilde{c} = [1 \quad 0 \quad \dots \quad 0]$$

Avec cette écriture, l'évolution de l'état x_t^k de chaque sous-système peut être définie à partir de la définition de cette sortie de la manière suivante :

$$x_t^k = E^k x_{t-1}^k + a^k y_{t-1}^k + D^{k-1} x_t^{k-1} u_t \quad (2.41)$$

où

$$E^k = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & 0 & \dots & \dots & 0 \end{bmatrix} \quad (2.42)$$

et

$$a^k = \begin{bmatrix} a_1^k \\ \vdots \\ \vdots \\ \vdots \\ a_{n_k}^k \end{bmatrix} \quad (2.43)$$

En effet, il est clair que la matrice dynamique définie par l'équation (2.40) de chaque sous-système peut être décomposée en :

$$\tilde{A}^k = E^k + a^k c \quad (2.44)$$

Ainsi, on a bien à partir de l'équation (2.14) :

$$\begin{aligned} x_t^k &= (E^k + a^k c) x_{t-1}^k + D^{k-1} x_t^{k-1} u_t \\ &= E^k x_{t-1}^k + a^k c x_{t-1}^k + D^{k-1} x_t^{k-1} u_t \end{aligned} \quad (2.45)$$

qui est conforme à l'expression (2.41) et où $cx_{t-1}^k = y_{t-1}^k$.

La transformée en z multivariable peut donc être réécrite sous la forme suivante :

$$\begin{aligned} X^1(z_1) &= z_1^{-1} (E^1 X^1(z_1) + a^1 Y^1(z_1)) + bU(z_1) \\ X^k(z_1, \dots, z_k) &= \tilde{z}_k^{-1} (E^k X^k(z_1, \dots, z_k) + a^k Y^k(z_1, \dots, z_k)) + D^{k-1} X^{k-1}(z_1, \dots, z_{k-1}) U(z_k), \forall k = 2 \dots d \end{aligned} \quad (2.46)$$

où par exemple, la transformée en z de la partie linéaire procède de la forma suivante :

$$\begin{aligned} x_t^1 &= E^1 x_{t-1}^1 + a^1 y_{t-1}^1 + bu_t \\ z_1 X^1(z_1) &= E^1 X^1(z_1) + a^1 Y^1(z_1) + bz_1 U(z_1) \\ X^1(z_1) &= z_1^{-1} (E^1 X^1(z_1) + a^1 Y^1(z_1)) + bU(z_1) \end{aligned}$$

Le regroupement des termes en $X^k(z_1, \dots, z_k)$ conduit alors à l'expression :

$$\begin{aligned} X^1(z_1) &= (z_1 I^1 - E^1)^{-1} (a^1 Y^1(z_1) + bz_1 U(z_1)) \\ X^k(z_1, \dots, z_k) &= (\tilde{z}_k I^k - E^k)^{-1} (a^k Y^k(z_1, \dots, z_k) + D^{k-1} X^{k-1}(z_1, \dots, z_{k-1}) \tilde{z}_k U(z_k)), \forall k = 2 \dots d \end{aligned} \quad (2.47)$$

Par ailleurs, il est aisé de montrer que :

$$(\tilde{z}I - E)^{-1} = \begin{bmatrix} \tilde{z}^{-1} & \tilde{z}^{-2} & \dots & \tilde{z}^{-n} \\ 0 & \tilde{z}^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \tilde{z}^{-2} \\ 0 & \dots & 0 & \tilde{z}^{-1} \end{bmatrix} \quad (2.48)$$

et que par conséquent :

$$c(\tilde{z}I - E)^{-1} = [\tilde{z}^{-1} \quad \tilde{z}^{-2} \quad \dots \quad \dots \quad \tilde{z}^{-n_d}] \quad (2.49)$$

Les sorties homogènes de degré k peuvent donc être réécrites sous la forme suivante :

$$Y^k(z_1, \dots, z_k) = c(\tilde{z}_k I - E)^{-1} (a^k Y^k(z_1, \dots, z_k) + D^{k-1} X^{k-1}(z_1, \dots, z_{k-1}) \tilde{z}_k U(z_k)), \forall k = 2 \dots d \quad (2.50)$$

Si on note

$$\begin{aligned} G^k(\tilde{z}_k) &= 1 - c(\tilde{z}_k I - E)^{-1} a^k \\ &= 1 - \sum_{i=1}^{n_k} a_i^k \tilde{z}_k^{-i} \end{aligned} \quad (2.51)$$

chaque sortie peut alors s'exprimer selon :

$$Y^k(z_1, \dots, z_k) = \frac{1}{G^k(\tilde{z}_k)} c(\tilde{z}_k I - E)^{-1} D^{k-1} X^{k-1}(z_1, \dots, z_{k-1}) \tilde{z}_k U(z_k), \forall k = 1 \dots d \quad (2.52)$$

Si on itère l'expression (2.47), on obtient alors :

$$\begin{aligned} X^k(z_1, \dots, z_k) &= (\tilde{z}_k I^k - E^k)^{-1} a^k Y^k(z_1, \dots, z_k) \\ &\quad + (\tilde{z}_k I^k - E^k)^{-1} D^{k-1} (\tilde{z}_{k-1} I^{k-1} - E^{k-1})^{-1} a^{k-1} Y^{k-1}(z_1, \dots, z_{k-1}) \tilde{z}_k U(z_k) \\ &\quad + (\tilde{z}_k I^k - E^k)^{-1} D^{k-1} (\tilde{z}_{k-1} I^{k-1} - E^{k-1})^{-1} D^{k-2} X^{k-2}(z_1, \dots, z_{k-2}) \tilde{z}_{k-1} \tilde{z}_k U(z_{k-1}) U(z_k) \end{aligned}$$

En épuisant ce processus de substitution, on aboutit à l'expression de chaque sous-système sous la forme suivante :

$$Y^k(z_1, \dots, z_k) = \frac{\sum_{i=1}^k (\tilde{z}_k I - E)^{-1} D^{k-1} (\tilde{z}_{k-1} I^{k-1} - E^{k-1})^{-1} a^{k-1} Y^{k-1}(z_1, \dots, z_{k-1}) D^{k-2} \dots D^{k-i} (\tilde{z}_{k-i} I^{k-i} - E^{k-i})^{-1} a^{k-i} Y^{k-i}(z_1, \dots, z_{k-i}) \tilde{z}_{k-i+1} \dots \tilde{z}_k U(z_{k-i+1}) \dots U(z_k)}{G^k(\tilde{z}_k)} \quad (2.53)$$

Le coefficient de chaque terme $Y^{k-i}(z_1, \dots, z_{k-i}) \tilde{z}_{k-i+1} U(z_{k-i+1}) \dots U(z_k)$ est donc un polynôme dont les paramètres sont construits comme des combinaisons algébriques des éléments des matrices $D^{k-1}, \dots, D^{k-i}$ et du vecteur a^{k-i} . Du fait du produit avec le monôme $\tilde{z}_{k-i+1} \dots \tilde{z}_k$, les termes de ce polynôme ont la structure suivante :

- pour tout $j = k-i+1, \dots, k$, ce polynôme a pour base les monômes $1, \tilde{z}_j^{-1}, \dots, \tilde{z}_j^{-n_j+1}$
- pour tout $j = 1, \dots, k-i$, ce polynôme a pour base les monômes $\tilde{z}_j^{-1}, \tilde{z}_j^{-2}, \dots, \tilde{z}_j^{-n_j}$

Pour simplifier les écritures, notons l'ensemble des puissances de $\tilde{z}$ sous la forme suivante :

$$\tilde{z}^{i, \dots, j} = \{\tilde{z}^i, \dots, \tilde{z}^j\} \quad (2.54)$$

Détaillons alors, avec cette notation, les termes apparaissant dans l'expression (2.53). En vertu de la définition de $\tilde{z}$ par l'expression, ce polynôme est construit sur les termes suivants:

- $Y^1(z_1) U(z_2) \dots U(z_k) :$
 $\{\tilde{z}_1^{-1, \dots, -n_1}, \tilde{z}_2^{0, \dots, -n_2+1}, \dots, \tilde{z}_k^{0, \dots, -n_k+1}\} = \{\tilde{z}_1^{-1, \dots, -(n_1+n_2+\dots+n_k-k+1)}, \tilde{z}_2^{0, \dots, -(n_1+n_2+\dots+n_k-k+1)}, \dots, \tilde{z}_k^{0, \dots, -n_k+1}\}$
- $Y^2(z_1, z_2) U(z_3) \dots U(z_k) :$
 $\{\tilde{z}_2^{-1, \dots, -n_2}, \tilde{z}_3^{0, \dots, -n_3+1}, \dots, \tilde{z}_k^{0, \dots, -n_k+1}\} = \{\tilde{z}_2^{-1, \dots, -(n_2+n_3+\dots+n_k-k+2)}, \tilde{z}_3^{0, \dots, -(n_3+\dots+n_k-k+2)}, \dots, \tilde{z}_k^{0, \dots, -n_k+1}\}$
- ...
- $Y^{k-1}(z_1, \dots, z_{k-1}) U(z_k) :$
 $\{\tilde{z}_{k-1}^{-1, \dots, -n_{k-1}}, \tilde{z}_k^{0, \dots, -n_k+1}\} = \{\tilde{z}_{k-1}^{-1, \dots, -(n_{k-1}+n_k-1)}, \tilde{z}_k^{0, \dots, -n_k+1}\}$

d'après [Rugh81, p. 154-155] et l'expression (2.27).

Rappelons que, par définition de l'opérateur retard, la transformée en z inverse a la propriété suivante :

$$\begin{aligned} Z^{-1}(\tilde{z}_k^{-j} Y^k(z_1, \dots, z_k)) &= y_{t-j}^k \\ Z^{-1}(\tilde{z}_k^{-j} U(z_k)) &= u_{t-j} \end{aligned}$$

Par conséquent, cette écriture permet de déterminer l'ensemble des retards intervenant dans l'écriture de y_t^k , c'est-à-dire des termes de la forme

$$y_{t-i}^{k-1} u_{t-j}, y_{t-i}^{k-2} u_{t-i} u_{t-j}, \dots, y_{t-i}^1 u_{t-i_1} \dots u_{t-i_{k-1}}$$

De plus on peut représenter le système bilinéaire homogène de degré k comme une cascade de sous-systèmes dont l'entrée de chacun des sous-systèmes comporte le terme $y_t^{k-1} u_t$, soit un transfert direct de la sortie de degré $k-1$ sur la sortie de degré k .

En effet, rappelons que, d'après l'expression (2.52), la sortie du k -ième système s'écrit sous la forme suivante :

$$Y^k(z_1, \dots, z_k) = \frac{1}{G^k(\tilde{z}_k)} c(\tilde{z}_k I^k - E^k)^{-1} D^{k-1} X^{k-1}(z_1, \dots, z_{k-1}) \tilde{z}_k U(z_k), \forall k = 1 \dots d \quad (2.55)$$

La forme canonique de D^{k-1} fait que cette matrice peut se décomposer comme suit :

$$D^{k-1} = \begin{bmatrix} c \\ \tilde{D}^{k-1} \end{bmatrix} \quad (2.56)$$

où $\tilde{D}^{k-1}$ contient l'ensemble des paramètres du sous système (preuve de la proposition 1, section 2.5.1). Par conséquent, le produit $D^{k-1} X^{k-1}(z_1, \dots, z_{k-1})$ se décompose également en :

$$D^{k-1} X^{k-1}(z_1, \dots, z_{k-1}) = \begin{bmatrix} c X^{k-1}(z_1, \dots, z_{k-1}) \\ \tilde{D}^{k-1} X^{k-1}(z_1, \dots, z_{k-1}) \end{bmatrix} = \begin{bmatrix} Y^{k-1}(z_1, \dots, z_{k-1}) \\ \tilde{D}^{k-1} X^{k-1}(z_1, \dots, z_{k-1}) \end{bmatrix} \quad (2.57)$$

De même, écrivons la décomposition suivante :

$$c(\tilde{z}_k I - E)^{-1} \tilde{z}_k = [1, \tilde{Z}_k^{-1}] \quad (2.58)$$

avec

$$\tilde{Z}_k^{-1} = [\tilde{z}_k^{-1} \quad \dots \quad \tilde{z}_k^{-n_k+1}] \quad (2.59)$$

Avec ces notations, on peut décomposer la sortie comme suit :

$$Y^k(z_1, \dots, z_k) = \frac{1}{G^k(\tilde{z}_k)} (Y^{k-1}(z_1, \dots, z_{k-1}) U(z_k) + \tilde{Z}_k^{-1} \tilde{D}^{k-1} X^{k-1}(z_1, \dots, z_{k-1}) U(z_k)) \quad (2.60)$$

Or, d'après (2.47), on a :

$$X^{k-1}(z_1, \dots, z_{k-1}) = (\tilde{z}_{k-1} I^{k-1} - E^{k-1})^{-1} (a^{k-1} Y^{k-1}(z_1, \dots, z_{k-1}) + D^{k-2} X^{k-2}(z_1, \dots, z_{k-2}) \tilde{z}_{k-1} U(z_{k-1})) \quad (2.61)$$

Par conséquent, on obtient :

$$\begin{aligned} Y^k(z_1, \dots, z_k) &= \frac{1}{G^k(\tilde{z}_k)} Y^{k-1}(z_1, \dots, z_{k-1}) U(z_k) \\ &+ \frac{1}{G^k(\tilde{z}_k)} \tilde{Z}_k^{-1} \tilde{D}^{k-1} (\tilde{z}_{k-1} I^{k-1} - E^{k-1})^{-1} a^{k-1} Y^{k-1}(z_1, \dots, z_{k-1}) U(z_k) \\ &+ \frac{1}{G^k(\tilde{z}_k)} \tilde{Z}_k^{-1} \tilde{D}^{k-1} (\tilde{z}_{k-1} I^{k-1} - E^{k-1})^{-1} D^{k-2} X^{k-2}(z_1, \dots, z_{k-2}) \tilde{z}_{k-1} U(z_{k-1}) U(z_k) \end{aligned} \quad (2.62)$$

En itérant ce processus comme précédemment, on obtient la forme suivante :

$$\begin{aligned} Y^k(z_1, \dots, z_k) &= \frac{1}{G^k(\tilde{z}_k)} Y^{k-1}(z_1, \dots, z_{k-1}) U(z_k) \\ &+ \frac{1}{G^k(\tilde{z}_k)} \sum_{i=1}^{k-1} \tilde{Z}_k^{-1} \tilde{D}^{k-1} \dots D^{k-i} (\tilde{z}_{k-i} I^{k-i} - E^{k-i})^{-1} a^{k-i} Y^{k-i}(z_1, \dots, z_{k-i}) \tilde{z}_{k-i+1} \dots \tilde{z}_k U(z_{k-i+1}) \dots U(z_k) \\ &+ \frac{1}{G^k(\tilde{z}_k)} \tilde{D}^{k-1} \dots D^{k-i} (\tilde{z}_{k-i} I^{k-i} - E^{k-i})^{-1} b \tilde{z}_1 \dots \tilde{z}_k U(z_1) \dots U(z_k) \end{aligned} \quad (2.63)$$

On peut alors trouver l'expression de la sortie de degré k du système en fonction de produits de sorties de degré $\leq k$ et des entrées retardées conformément à l'expression ci-dessus. Malheureusement, cette représentation n'est pas linéaire vis-à-vis des paramètres. Cependant, dans le cas quadratique, il existe une représentation « hybride » qui permet d'assurer cette linéarité souhaitée.

2.6.2 Forme linéaire en les paramètres

Il est clair que la paramétrisation proposée au paragraphe précédent ne donne pas lieu à une relation linéaire en les paramètres à la sortie du système. Cependant celle-ci permet de nous éclairer sur une structure possible qui n'aurait pas cet inconvénient. Pour ce faire, il suffit de considérer la forme suivante :

$$\begin{aligned}
y_t^1 &= \sum_{i=1}^{n_1} a_i^1 y_{t-i}^1 + \sum_{i=1}^{n_1} b_i u_{t-i+1} \\
y_t^2 &= \sum_{i=1}^{n_2} a_i^2 y_{t-i}^2 + y_t^1 u_t + \sum_{i=1}^{n_2} \sum_{j=1}^{n_1} d_{i,j}^1 y_{t-i-j+1}^1 u_{t-i+1} \\
&\dots \\
y_t^d &= \sum_{i=1}^{n_d} a_i^d y_{t-i}^d + y_t^{d-1} u_t + \sum_{i=1}^{n_d} \sum_{j=1}^{n_{d-1}} d_{i,j}^{d-1} y_{t-i-j+1}^{d-1} u_{t-i+1}
\end{aligned} \tag{2.64}$$

En effet, il apparaît qu'avec cette représentation, le développement de la sortie en les paramètres prend la forme multi-degré suivante :

$$\begin{aligned}
y_t^d &= \sum_{i=1}^{n_d} a_i^d y_{t-i}^d + y_t^{d-1} u_t + \sum_{i=1}^{n_d} \sum_{j=1}^{n_{d-1}} d_{i,j}^{d-1} y_{t-i-j+1}^{d-1} u_{t-i+1} \\
&= \sum_{i=1}^{n_d} a_i^d y_{t-i}^d + \sum_{i=1}^{n_d} \sum_{j=1}^{n_{d-1}} d_{i,j}^{d-1} y_{t-i-j+1}^{d-1} u_{t-i+1} + \sum_{i=1}^{n_{d-1}} a_i^{d-1} y_{t-i}^{d-1} u_t + \sum_{i=1}^{n_{d-1}} \sum_{j=1}^{n_{d-2}} d_{i,j}^{d-2} y_{t-i-j+1}^{d-2} u_{t-i+1} u_t + y_t^{d-2} (u_t)^2 \\
&\dots \\
&= \sum_{k=1}^d \sum_{i=1}^{n_k} a_i^k y_{t-i}^k (u_t)^{d-k} + \sum_{k=1}^{d-1} \sum_{i=1}^{n_{k+1}} \sum_{j=1}^{n_k} d_{i,j}^k y_{t-i-j+1}^k u_{t-i+1} (u_t)^{d-k} + \sum_{i=1}^{n_1} b_i u_{t-i+1} (u_t)^{d-1}
\end{aligned} \tag{2.65}$$

qui est bien linéaire vis-à-vis des paramètres du système.

3. Réalisation et Identification par l'approche héréditaire

3.1 Introduction

Ce chapitre présente le corps de nos travaux théoriques sur la réalisation et l'identification des systèmes bilinéaires homogènes décrits au chapitre 2.

La partie intitulée réalisation des systèmes linéaire et bilinéaire s'attache à donner un algorithme destiné à construire un système linéaire et/ou bilinéaire homogène dont on connaît la (les) fonctions d'autocorrélation entrée/sortie quand celui-ci est attaqué par un bruit blanc, c'est-à-dire les fonctions du type

$$h^k(\tau_1, \dots, \tau_k) = E[y_t u_{t-\tau_1} \dots u_{t-\tau_k}] \quad (3.1)$$

Dans le cas $k=1$, cette fonction s'identifie à la réponse impulsionnelle d'un système linéaire.

On montre que l'approche héréditaire étendue aux systèmes bilinéaires homogènes de degré k exhibe un nouveau type de calcul « multi-héréditaire » sous la forme de boucles héréditaires imbriquées et qui à chaque pas de l'algorithme balaye l'ensemble des intercorrélations données sur le système.

Une deuxième partie est consacrée à l'identification des systèmes linéaire et/ou bilinéaires homogènes dont on connaît une trajectoire entrée/sortie bruitée mesurée sur le système physique. Nous établissons dans cette partie que l'algorithme s'étend facilement dans ce cas à des systèmes de degré quelconque.

A noter que pour les deux problèmes étudiés, il convient, sur des cas physiques, de ne pas considérer qu'un seul noyau homogène de degré k mais la somme de systèmes homogènes de degré $\leq k$.

3.2 Réalisation des systèmes linéaires

3.2.1 Principe

Le principe de réalisation proposé ici repose sur une vision stochastique des systèmes dynamiques causaux (markoviens). La notion de réalisation héréditaire (ou à mémoire augmentante avec le temps) est une technique exacte de réalisation des systèmes linéaire à partir de la donnée de l'autocorrélation d'un signal de type ARMA [Mon94] (système linéaire à entrée bruit blanc), de l'intercorrélation entrée-sortie (réponse impulsionnelle d'un système linéaire) ou des deux (système ARMAX). En d'autres termes, le problème consiste à trouver un système linéaire de dimension fixée dont l'autocorrélation de la sortie (systèmes ARMA), l'intercorrélation entrée/sortie (systèmes ARX) ou les deux (systèmes ARMAX) soit aussi proche possible (au sens de L^2) que celle(s) d'un système donné. A noter que cette technique permet de réaliser une réduction optimale de la dimension d'un modèle linéaire, même si celui-ci est de dimension infinie.

Sa genèse repose sur la théorie du filtrage optimal de Kalman. En effet, considérons le cas ARMAX. La sortie du système linéaire peut être définie comme suit :

$$y_t = \sum_{i=1}^n a^i y_{t-i} + \sum_{i=1}^n c^i v_{t-i} + \sum_{i=1}^n b^i u_{t-i+1} + v_t \quad (3.2)$$

où u_t désigne l'entrée du système, v_t est un bruit blanc (non nécessairement gaussien) et y_t est la sortie du système. A noter que $c_i = 0, \forall i = 1, \dots, n$ conduit à la forme ARX et $b_i = 0, \forall i = 1, \dots, n$ conduit à la forme ARMA. On peut alors montrer que le prédicteur optimal d'un tel système peut être réalisé par le système linéaire suivant :

$$\hat{y}_{t|t-1} = \sum_{i=1}^n a^i \hat{y}_{t-i|t-i-1} + \sum_{i=1}^n k^i (y_{t-i} - \hat{y}_{t-i|t-i-1}) + \sum_{i=1}^n b^i u_{t-i+1} \quad (3.3)$$

Dans cette représentation, $\hat{y}_{t|t-1}$ désigne le prédicteur optimal, soit :

$$\hat{y}_{t|t-1} = E[y_t | y_{0:t-1}] \quad (3.4)$$

le terme $\tilde{y}_{t|t-1} = y_t - \hat{y}_{t|t-1}$ est un bruit blanc nommé « innovation », $\{k^i, i = 1, \dots, n\}$ sont les composantes du gain du prédicteur optimal. Si le bruit blanc v_t est gaussien, alors l'innovation est également gaussienne. Par conséquent, la sortie y_t peut être réalisée à partir de ce prédicteur, en utilisant le fait que $y_t = \hat{y}_{t|t-1} + \tilde{y}_{t|t-1}$ et en utilisant une nouvelle paramétrisation sous la forme suivante :

$$y_t = \sum_{i=1}^n a_t^i y_{t-i} + \sum_{i=1}^n c_t^i \tilde{y}_{t-i|t-i-1} + \sum_{i=1}^n b_t^i u_{t-i+1} + \tilde{y}_{t|t-1} \quad (3.5)$$

Cette représentation est appelée « forme filtre ».

Or, le prédictor optimal est aussi le meilleur filtre linéaire, au sens du minimum de variance. En d'autres termes, si on recherche dans la classe paramétrique des systèmes linéaires de dimension égale à celle du système de départ le système linéaire de la forme (3.3) qui minimise :

$$E\left[\left(y_t - \hat{y}_{t|t-1}\right)^2\right] \quad (3.6)$$

on obtient le meilleur prédictor du système, donc le meilleur modèle. Il apparaît alors clairement, comme en commande optimale, que le calcul des paramètres optimaux à l'instant t s'appuie sur les estimations et innovations précédentes, sans remettre en cause le passé de celles-ci. En commande optimale, ceci se traduit par le principe d'optimalité qui dit que toute sous trajectoire d'une trajectoire optimale est optimale.

Pour obtenir le modèle optimal, on considère un prédictor de la sortie **variable avec le temps**, comme c'est le cas en filtrage linéaire de Kalman. Les valeurs asymptotiques des paramètres du filtre sont alors transformée en forme filtre.

Cette opération s'interprète comme une **projection** de y_t sur ses prédictions $\hat{y}_{t-i|t-i-1}$, ses innovations $\tilde{y}_{t-i|t-i-1} = y_{t-i} - \hat{y}_{t-i|t-i-1}$ et ses entrées u_{t-i} passées. Si on définit le vecteur de régression :

$$\varphi_{t-1} = [\hat{y}_{t-1|t-2}, \dots, \hat{y}_{t-n|t-n-1}, \tilde{y}_{t-1|t-2}, \dots, \tilde{y}_{t-n|t-n-1}, u_t, \dots, u_{t-n+1}]^T \quad (3.7)$$

l'équation du prédictor est alors linéaire en les paramètres sous la forme suivante :

$$\hat{y}_{t|t-1} = \theta_t^T \varphi_{t-1} \quad (3.8)$$

avec

$$\theta_t = \begin{bmatrix} a_t \\ c_t \\ b_t \end{bmatrix} \quad (3.9)$$

Les paramètres optimaux sont alors obtenus par simple inversion d'un système linéaire comme suit :

$$\theta_t = E[\varphi_{t-1} \varphi_{t-1}^T]^{-1} E[\varphi_{t-1} y_t] \quad (3.10)$$

Si on considère le produit extérieur entre deux variables aléatoires x_t et y_t vectorielles défini par :

$$\langle x_t | y_t \rangle = E[x_t (y_t)^T] \quad (3.11)$$

le prédictor optimal peut être vu comme la projection, au sens de ce produit, de la sortie sur les sorties, innovations et entrées précédentes :

$$\hat{y}_{t|t-1} = P(y_t | \hat{y}_{t-1|t-2}, \tilde{y}_{t-1|t-2}, u_t, \dots, \hat{y}_{t-n|t-n-1}, \tilde{y}_{t-1|t-n-1}, u_{t-n+1}) \quad (3.12)$$

Il apparaît donc nécessaire de calculer à chaque instant les autocorrélations du vecteur de régression ainsi que ses intercorrélations avec l'entrée et la sortie. Or, il est clair que ces espérances n'ont pas une écriture récursive en dimension finie (en fait en dimension non bornée). En effet, pour calculer les valeurs à l'instant t des autocorrélations et intercorrélations, il est nécessaire d'avoir la connaissance des statistiques de $\hat{y}_{t-1|t-2} = \theta_{t-1}^T \varphi_{t-2}$, donc de θ_{t-1} qui lui-même dépend de celles de φ_{t-3} , et ainsi de suite. En définitive, le calcul de θ_t fait appel à toutes les valeurs passées des paramètres $\theta_\tau, \tau=1, \dots, t-1$, 1 étant l'instant d'initialisation.

Le critère à minimiser est quadratique et, par conséquent, annuler la dérivée de celui-ci est également une condition suffisante d'optimalité. De fait, la dépendance non linéaire en les paramètres est résolue au prix du caractère héréditaire de l'algorithme (en pratique, à mémoire augmentante).

3.2.2 Algorithme

Nous décrivons ici l'algorithme dans le cas du modèle ARX. Le problème consiste donc à rechercher le système linéaire causal sous la forme suivante :

$$\hat{y}_t = \sum_{i=1}^n a_t^i \hat{y}_{t-i} + \sum_{i=1}^n b_t^i u_{t-i+1} \quad (3.13)$$

dont on connaît l'intercorrélation entrée/sortie, à partir de laquelle on obtient la réponse impulsionnelle h_t , soit la fonction :

$$E[y_t u_\tau] = h(t - \tau), \forall \tau \leq t \quad (3.14)$$

le système étant supposé stationnaire. Définissons les vecteurs de paramètres suivants :

$$\begin{aligned} a_t &= [a_t^1, \dots, a_t^n]^T \\ b_t &= [b_t^1, \dots, b_t^n]^T \end{aligned} \quad (3.15)$$

et les vecteurs de régression suivants :

$$\begin{aligned} \hat{Y}_{t-1} &= [\hat{y}_{t-1}, \dots, \hat{y}_{t-n}]^T \\ U_t &= [u_t, \dots, u_{t-n+1}]^T \end{aligned} \quad (3.16)$$

Avec ces notations, le modèle de prédiction s'écrit de la manière synthétique suivante :

$$\hat{y}_t = (a_t)^T \hat{Y}_{t-1} + (b_t)^T U_t \quad (3.17)$$

La minimisation de la variance de l'erreur $E[(y_t - \hat{y}_{t|t-1})^2]$ conduit, après dérivation par rapport aux paramètres, aux équations d'orthogonalité suivantes :

$$\begin{aligned} E[(y_t - \hat{y}_t)\hat{Y}_{t-1}] &= 0 \\ E[(y_t - \hat{y}_t)U_t] &= 0 \end{aligned} \quad (3.18)$$

En développant l'expression de $\hat{y}_t$ en fonction des paramètres, on obtient le système linéaire suivant :

$$\begin{aligned} E[\hat{Y}_{t-1}(\hat{Y}_{t-1})^T]a_t + E[\hat{Y}_{t-1}(U_t)^T]b_t &= E[\hat{Y}_{t-1}y_t] \\ E[U_t(\hat{Y}_{t-1})^T]a_t + E[U_t(U_t)^T]b_t &= E[U_t y_t] \end{aligned} \quad (3.19)$$

qui par inversion donne la valeur des paramètres à l'instant t . Par conséquent, le calcul de ces paramètres nécessite la connaissance de l'ensemble des espérances mathématiques suivantes :

$$E[\hat{Y}_{t-1}(\hat{Y}_{t-1})^T], E[\hat{Y}_{t-1}(U_t)^T], E[U_t(U_t)^T], E[\hat{Y}_{t-1}y_t], E[U_t y_t]$$

Détaillons à présent comment ces espérances sont mises à jour :

- $E[\hat{Y}_t(\hat{Y}_t)^T]$. Cette espérance ne nécessite que la mise à jour de sa première colonne, c'est-à-dire de $E[\hat{y}_{t-i}\hat{y}_t] \forall i=0..n-1$ qui se réduit elle-même à l'ensemble $E[\hat{y}_{t-i}y_t] \forall i=0..n-1$. En effet, les équations d'orthogonalité permettent d'écrire

$$E[(y_t - \hat{y}_t)\hat{Y}_t] = 0 \Rightarrow E[\hat{y}_t \hat{Y}_t] = E[y_t \hat{Y}_t]$$

qui conduit à $E[\hat{y}_{t-i}\hat{y}_t] = E[\hat{y}_{t-i}y_t] \forall i=0..n-1$

- $E[\hat{Y}_{t-1}(U_t)^T]$. Chaque composante de cette matrice s'exprime en fonction de l'intercorrélation entrée/sortie selon :

$$E[\hat{y}_{t-i}u_{t-j+1}] = \begin{cases} E[\hat{y}_{t-i}u_{t-j+1}] = 0 & \text{if } i \geq j \\ E[y_{t-i}u_{t-j+1}] = h(j-i-1) & \text{if } i < j \end{cases}$$

- $E[U_t(U_t)^T]$. L'entrée étant considérée comme un bruit blanc, on a simplement :

$$E[u_{t-i+1}u_{t-j+1}] = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

- $E[U_t y_t]$ est donné par l'intercorrélation par :

$$E[u_{t-i+1} y_t] = h(i-1), \forall i = 1 \dots n$$

En définitive, on constate que la connaissance de seules quantités $E[\hat{y}_{t-i} y_t], \forall i = 0 \dots n$ permet de déterminer tout le système linéaire nécessaire à la détermination des paramètres. Elles constituent de fait la partie héréditaire de l'algorithme. Si on initialise le calcul de cette espérance avec $E[\hat{y}_0 y_t] = 0$, on est amené à calculer $\forall \tau = 1 \dots t-1$ l'expression suivante :

$$\begin{aligned} E[\hat{y}_\tau y_t] &= \sum_{i=1}^n a_\tau^i E[\hat{y}_{\tau-i} y_t] + \sum_{i=1}^n b_\tau^i h(t - \tau + i - 1) \\ E[\hat{Y}_\tau y_t] &= \left[E[\hat{y}_\tau y_t], E\left[\left(\hat{Y}_{\tau-1}\right)^T y_t\right] \right]^T \end{aligned} \quad (3.20)$$

Apparaît ici clairement la nécessité de conserver en mémoire l'ensemble des paramètres préalablement calculés.

L'algorithme se résume donc aux étapes suivantes :

1. Initialisation :

$$t = 1$$

$$\begin{aligned} E[\hat{Y}_0 (\hat{Y}_0)^T] &= 0 \\ E[\hat{Y}_0 (U_1)^T] &= \begin{cases} 0 & \text{if } i \geq j \\ h(j - i - 1) & \text{if } i < j \end{cases} \\ E[\hat{Y}_0 y_t] &= 0 \\ E[U_t (U_t)^T] &= I \\ E[U_t y_t] &= [h(0), \dots, h(n-1)]^T \end{aligned}$$

2. Inversion du système linéaire suivant :

$$\begin{aligned} E[\hat{Y}_{t-1} (\hat{Y}_{t-1})^T] a_t + E[\hat{Y}_{t-1} (U_t)^T] b_t &= E[\hat{Y}_{t-1} y_t] \\ E[U_t (\hat{Y}_{t-1})^T] a_t + E[U_t (U_t)^T] b_t &= E[U_t y_t] \end{aligned}$$

et calcul des paramètres a_t et b_t . Ces valeurs sont mises en mémoire

3. Calcul héréditaire de $E[\hat{y}_{t-i} y_t], \forall i = 0 \dots n$

- $E[\hat{Y}_0 y_t] = 0$
- Pour $\tau = 1, \dots, t-1$, $E[\hat{y}_\tau y_t] = \sum_{i=1}^n a_\tau^i E[\hat{y}_{\tau-i} y_t] + \sum_{i=1}^n b_\tau^i h(t - \tau + i - 1)$

4. Mise à jour des espérances $E\left[\hat{Y}_{t-1}(\hat{Y}_{t-1})^T\right]$, $E\left[\hat{Y}_{t-1}(U_t)^T\right]$, $E\left[\hat{Y}_{t-1}y_t\right]$.

Pour la matrice :

$$E\left[\hat{Y}_{t-1}(\hat{Y}_{t-1})^T\right]$$

seule sa première colonne est obtenue à partir du pas 3, c'est-à-dire de $E\left[\hat{y}_{t-i-1}y_{t-1}\right]$, $\forall i = 0..n-1$.

La matrice :

$$E\left[\hat{Y}_{t-1}(U_t)^T\right] = E\left[\left[\hat{y}_{t-1}, \dots, \hat{y}_{t-n}\right]^T \left[u_t, \dots, u_{t-n+1}\right]\right]$$

est mise à jour à partir de $h(t-\tau)$, voir (3.14) :

$$E\left[\hat{y}_{t-i}u_{t-j+1}\right] = \begin{cases} E\left[\hat{y}_{t-i}u_{t-j+1}\right] = 0 & \text{if } i \geq j \\ E\left[\hat{y}_{t-i}u_{t-j+1}\right] = h(j-i-1) & \text{if } i < j \end{cases}$$

Finalement, la matrice

$$E\left[\hat{Y}_{\tau}y_t\right] = \left[E\left[\hat{y}_{\tau}y_t\right], E\left[\left(\hat{Y}_{\tau-1}\right)^T y_t\right] \right]^T$$

est mise à jour à partir de l'information recueilli au pas 3 et aux valeurs passées.

5. $t+1 \rightarrow t$ et retour en 2

3.3 Identification des systèmes linéaires

3.3.1 Principe

Le principe de l'identification héréditaire reprend la même idée en substituant à l'espérance mathématique de l'erreur de prédiction son espérance expérimentale [Mon96a, Mon04]. Dans ce cas, le critère à minimiser s'interprète comme une distance entre trajectoires d'estimation. On peut alors montrer que l'algorithme converge sur un critère d'ergodicité.

Plus précisément, notons :

$$E^t[x_\bullet] = \frac{1}{t} \sum_{\tau=1}^t x_\tau \quad (3.21)$$

cette espérance expérimentale de la trajectoire de x_τ sur l'intervalle $[1, t]$. En reprenant le principe explicité plus haut, soit $\hat{y}_\tau^t$ la trajectoire du prédicteur optimal obtenu à l'instant t à partir des données sur l'intervalle $[1, t]$. Cette trajectoire est obtenue en s'appuyant sur les valeurs passées de l'estimateur, soit :

$$\hat{y}_{\tau|t-1}^t = \sum_{i=1}^n a_i^t \hat{y}_{\tau-i|t-1}^{t-i} + \sum_{i=1}^n k_i^t (y_{\tau-i} - \hat{y}_{\tau-i|t-1}^{t-i}) + \sum_{i=1}^n b_i^t u_{\tau-i+1} \quad (3.22)$$

Si on définit la trajectoire du vecteur de régression par

$$\phi_\tau^{t-1} = [\hat{y}_{\tau-1|t-2}^{t-1}, \dots, \hat{y}_{\tau-n|t-n-1}^{t-n}, \tilde{y}_{\tau-1|t-2}^{t-1}, \dots, \tilde{y}_{\tau-n|t-n-1}^{t-n}, u_{\tau-1}, \dots, u_{\tau-n+1}]^T \quad (3.23)$$

la minimisation de l'erreur expérimentale de prédiction conduit à :

$$\theta_t = E^t \left[\phi_\bullet^{t-1} (\phi_\bullet^{t-1})^T \right]^{-1} E^t [\phi_\bullet^{t-1} y_\bullet] \quad (3.24)$$

Comme pour le cas du problème de réalisation, le calcul des espérances expérimentales nécessaires à l'obtention des paramètres à l'instant t n'est pas récursif mais héréditaire. En effet, il apparaît clairement que pour calculer θ_t , il est nécessaire de connaître les autocorrélations expérimentales de ϕ_τ^{t-1} et intercorrélations expérimentales entre y_τ et ϕ_τ^{t-1} . Or ces quantités dépendent directement des trajectoires passées des estimateurs obtenus aux étapes précédentes. Ainsi, pour calculer les paramètres du système identifié à l'instant t , il est nécessaire de garder en mémoire toutes les trajectoires des estimateurs $\hat{y}_{\tau|t-1}^{t-1}, \dots, \hat{y}_{\tau|t-1}^{t-n}$.

Si on considère le produit extérieur entre deux trajectoires x_τ et y_τ vectorielles défini par

$$\langle x_\bullet | y_\bullet \rangle^t = \frac{1}{t} \sum_{\tau=1}^t x_\tau^T y_\tau \quad (3.25)$$

le prédictor optimal peut être vu comme la projection, au sens de ce produit, de la trajectoire de sortie sur les trajectoires suivantes :

$$\hat{y}_\tau^t = P_\tau^t \left(y_\bullet \middle| \hat{y}_{\bullet-1|\bullet-2}^{t-1}, \tilde{y}_{\bullet-1|\bullet-2}^{t-1}, u_{\bullet-1}, \dots, \hat{y}_{\bullet-n|\bullet-n-1}^{t-n}, \tilde{y}_{\bullet-n|\bullet-n-1}^{t-n}, u_{\bullet-n+1} \right) \quad (3.26)$$

Comme dans le cas de la réalisation des systèmes linéaires, annuler la dérivée du critère est une condition suffisante d'optimalité. A nouveau, la dépendance non linéaire en les paramètres est résolue au prix du caractère héréditaire de l'algorithme.

3.3.2 Algorithme

L'algorithme d'identification est structuré ainsi :

1. Initialisation. Les données antérieures à $t=1$ étant inconnues, la projection sur celles-ci donne naturellement une valeur nulle (division par une covariance *a priori* infinie). Il est donc équivalent de fixer ces valeurs à zéro : $\hat{y}_{\tau|\tau-1}^{1-i} = 0, \forall \tau < 1, \forall i = 1, \dots, n$. A l'instant $t-1$, nous disposons des données des trajectoires préalablement optimisées :

$$\hat{y}_{\tau|\tau-1}^{t-i}, \forall \tau = 1, \dots, t-1, \forall i = 1, \dots, n$$

2. A l'instant t , les nouvelles données y_t, u_t , permettent de mettre à jour la matrice d'autocorrélation expérimentale et le vecteur d'intercorrrelation respectivement :

$$E^t [\varphi_\bullet^{t-1} \varphi_\bullet^{t-1T}] = \frac{1}{t} \sum_{\tau=1}^t \varphi_\tau^{t-1} \varphi_\tau^{t-1T}$$

$$E^t [y_\bullet \varphi_\bullet^{t-1}] = \frac{1}{t} \sum_{\tau=1}^t y_\tau \varphi_\tau^{t-1}$$

3. On obtient les nouvelles valeurs des paramètres du système à l'instant t par inversion du système linéaire suivant :

$$\theta_t \triangleq E^t [\varphi_\bullet^{t-1} \varphi_\bullet^{t-1T}]^{-1} \times E^t [y_\bullet \varphi_\bullet^{t-1}]$$

4. (Partie héréditaire). Une fois les paramètres obtenus, nous faisons le calcul des nouvelles trajectoires de l'estimateur $\hat{y}_{\tau|\tau-1}^t$ qui s'appuie sur la trajectoire $\hat{y}_{\tau|\tau-1}^{t-i}$ calculée à l'étape précédente en utilisant le système :

$$\forall \tau = 1 \dots t : \hat{y}_{\tau|\tau-1}^t = \sum_{i=1}^n a_\tau^i \hat{y}_{\tau-i|\tau-i-1}^{t-i} + \sum_{i=1}^n k_\tau^i (y_{\tau-i} - \hat{y}_{\tau-i|\tau-i-1}^{t-i}) + \sum_{i=0}^n b_\tau^i u_{\tau-i+1}$$

5. Retour en 2.

3.4 Réalisation des systèmes bilinéaires homogènes

3.4.1 Principe général

L'objectif poursuivi est de trouver un système bilinéaire qui explique au mieux une fonction d'intercorrélation entrée/sortie donnée. Pour les systèmes linéaires, nous avons vu au paragraphe 3.2 que la réalisation optimale d'un système linéaire causal pouvait être obtenue par une modélisation stochastique des systèmes (entrée et bruits d'observations blancs) à partir de la donnée de sa fonction d'autocorrélation entrée/sortie :

$$h(t - \tau) = E[y_t u_\tau] \forall \tau \leq t \quad (3.27)$$

Nous avons vu au chapitre 2 que les systèmes bilinéaires constituent une classe important des systèmes non linéaires de par leur capacité à approximer tout système non linéaire analytique. Cependant, il est clair que pour un système bilinéaire général donnant lieu à une série de Volterra de degré infini, la définition de l'intercorrélation entrée/sortie n'est pas utilisable en pratique, le nombre de fonctions à manipuler étant en nombre infini. En revanche, si on s'intéresse à un seul ou à un nombre fini de noyaux homogènes, leur intercorrélation entrée/sortie peut être décrite par un nombre fini de fonctions, égale au nombre de termes homogènes considérés.

Nous nous intéresserons donc ici à la réalisation d'un de ces termes homogène de degré fixé d et donc à la détermination d'un système bilinéaire homogène ayant une intercorrélation entrée/sortie aussi proche que possible d'une fonction donnée sous la forme :

$$h^d(t - \tau_1, \dots, t - \tau_d) = E[y_t u_{\tau_1} \dots u_{\tau_d}] \quad (3.28)$$

où $h^d(t - \tau_1, \dots, t - \tau_d)$ sont des données supposés connues.

Rappelons que l'approche héréditaire de réalisation des systèmes suppose une représentation telle que les paramètres variables avec le temps apparaissent linéairement dans l'expression de la sortie. En effet, ceci suffit à assurer la convexité du problème d'optimisation, l'optimisation concernant uniquement les paramètres à l'instant t , leurs valeurs précédentes étant inchangées (principe d'optimalité).

Rappelons que la forme canonique linéaire en les paramètres développée au chapitre 2 est réalisée par le système suivant :

$$\begin{aligned} x_t^1 &= A^1 x_{t-1}^1 + b u_t \\ x_t^2 &= A^2 x_{t-1}^2 + x_t^1 u_t \\ &\dots \\ x_t^d &= A^d x_{t-1}^d + x_t^{d-1} u_t \\ y_t^d &= c x_t^d \end{aligned}$$

et la sortie a pour expression :

$$y_t^d = \sum_{i=1}^n a_i^d y_{t-i}^d + \sum_{k=1}^{d-1} \sum_{i=1}^n A_i^k x_{t-i}^k (u_{t-i+1})^{d-k-1} + \sum_{i=1}^n b_i (u_{t-i+1})^d$$

où

$$\{a_i^d \in R, i=1, \dots, n\} \cup \{A_i^k \in R^n, i=1, \dots, n, k=1, \dots, d-1\} \cup \{b_i \in R, i=1, \dots, n\}$$

constitue l'ensemble des paramètres du système homogène de degré d . A remarquer que la forme (2.65) est une structure particulière de celle en dessus.

L'objectif poursuivi consiste donc à trouver un système bilinéaire homogène du type ci-dessus dont la fonction d'intercorrélation soit aussi proche possible que celle donnée. Comme dans le cas linéaire, il s'agit de calculer le prédicteur variable avec le temps ayant la structure suivante :

$$\begin{aligned} \hat{y}_t^d &= \sum_{i=1}^n a_{i,t}^d \hat{y}_{t-i}^d + \sum_{k=1}^{d-1} \sum_{i=1}^n A_{i,t}^k \hat{x}_{t-i}^k (u_{t-i+1})^{d-k-1} + \sum_{i=1}^n b_{i,t} (u_{t-i+1})^d \\ \hat{x}_t^{d-1} &= A_t^{d-1} \hat{x}_{t-1}^{d-1} + \sum_{k=1}^{d-2} A_t^k \hat{x}_{t-i}^k (u_t)^{d-k-1} + \sum_{i=1}^n b_{i,t} (u_{t-i+1})^{d-1} \\ &\dots \\ \hat{x}_t^2 &= A_t^2 \hat{x}_{t-1}^2 + A_t^1 \hat{x}_{t-1}^1 u_t + b_t (u_t)^2 \\ \hat{x}_t^1 &= A_t^1 \hat{x}_{t-1}^1 + b_t u_t \end{aligned} \quad (3.29)$$

et qui minimise la variance de l'erreur $E[(y_t - \hat{y}_t^d)^2]$.

L'optimisation des paramètres à l'instant t conduit aux équations d'orthogonalité suivantes :

$$\begin{aligned} E[(y_t - \hat{y}_t^d) \hat{y}_{t-j}^d] &= 0, \quad \forall j=1, \dots, n \\ E[(y_t - \hat{y}_t^d) [\hat{x}_{t-j}^l]_m (u_{t-j+1})^{d-l-1}] &= 0, \quad \forall l=1, \dots, d-1, \quad \forall j=1, \dots, n, \quad \forall m=1, \dots, n \\ E[(y_t - \hat{y}_t^d) (u_{t-j+1})^d] &= 0, \quad \forall j=1, \dots, n \end{aligned} \quad (3.30)$$

Remarquons à nouveau que l'optimisation prend cette forme grâce au fait que les paramètres sont variables avec le temps et que par conséquent $\hat{x}_{t-i}^k, \forall i \geq 1$ dépendent des paramètres passés et non présents.

Si on développe l'expression de $\hat{y}_t^d$ dans ces équations d'orthogonalité, on aboutit clairement à un système linéaire en les paramètres dont les coefficients sont des espérances calculables héréditairement. Ce système d'équations linéaires est le suivant :

- $\forall j=1, \dots, n$

$$\sum_{i=1}^n a_{i,t}^d E [\hat{y}_{t-i}^d \hat{y}_{t-j}^d] + \sum_{k=1}^{d-1} \sum_{i=1}^n A_{i,t}^k E [\hat{y}_{t-j}^d \hat{x}_{t-i}^k (u_{t-i+1})^{d-k-1}] + \sum_{i=1}^n b_{i,t} E [\hat{y}_{t-j}^d (u_{t-i+1})^d] = E [\hat{y}_{t-j}^d y_t]$$

- $\forall l=1, \dots, d-1, \forall j=1, \dots, n, \forall m=1, \dots, n$

$$\sum_{i=1}^n a_{i,t}^d E [\hat{y}_{t-i}^d [\hat{x}_{t-j}^l]_m (u_{t-j+1})^{d-l-1}] + \sum_{k=1}^{d-1} \sum_{i=1}^n A_{i,t}^k E [\hat{x}_{t-i}^k [\hat{x}_{t-j}^l]_m (u_{t-j+1})^{d-l-1} (u_{t-i+1})^{d-k-1}]$$

$$+ \sum_{i=1}^n b_{i,t} E [[\hat{x}_{t-j}^l]_m (u_{t-j+1})^{d-l-1} (u_{t-i+1})^d] = E [y_t [\hat{x}_{t-j}^l]_m (u_{t-j+1})^{d-l-1}]$$
- $\forall j=1, \dots, n$

$$\sum_{i=1}^n a_{i,t}^d E [\hat{y}_{t-i}^d (u_{t-j+1})^d] + \sum_{k=1}^{d-1} \sum_{i=1}^n A_{i,t}^k E [\hat{x}_{t-i}^k (u_{t-i+1})^{d-k-1} (u_{t-j+1})^d]$$

$$+ \sum_{i=1}^n b_{i,t} E [(u_{t-i+1})^d (u_{t-j+1})^d] = E [y_t (u_{t-j+1})^d]$$

où $[\hat{x}_t^l]_m$ désigne la m -ième composante du vecteur $\hat{x}_t^l$.

L'écriture complète de l'algorithme de mise à jour des espérances est quelque peu compliquée, aussi nous nous limiterons ici à sa description complète dans le cas quadratique.

3.4.2 Cas Quadratique

3.4.2.1 Equations de projection

Examinons plus en détails le cas quadratique, c'est à dire $d=2$. Dans ce cas particulier, l'état linéaire et la sortie quadratique sont définis par :

$$\begin{aligned} \hat{y}_t^2 &= \sum_{i=1}^n a_{i,t}^2 \hat{y}_{t-i}^2 + \sum_{i=1}^n A_{i,t}^1 \hat{x}_{t-i}^1 u_{t-i+1} + \sum_{i=1}^n b_{i,t} (u_{t-i+1})^2 \\ \hat{x}_t^1 &= A_t^1 \hat{x}_{t-1}^1 + b_t u_t \end{aligned} \quad (3.31)$$

Définissons

$$\begin{aligned} \hat{Y}_{t-1}^2 &= [\hat{y}_{t-1}^2, \dots, \hat{y}_{t-n}^2]^T \\ U_t^2 &= [(u_t)^2, \dots, (u_{t-n+1})^2]^T \\ X_{t-1}^1 \otimes U_t^1 &= [(\hat{x}_{t-1}^1)^T u_t, \dots, (\hat{x}_{t-n}^1)^T u_{t-n+1}]^T \\ \tilde{a}_t^2 &= [a_{1,t}^2, \dots, a_{n,t}^2]^T \\ b_t &= [b_{1,t}, \dots, b_{n,t}]^T \end{aligned} \quad (3.32)$$

et le vecteur de dimension n^2 suivant :

$$\tilde{A}_t^1 = [A_{1,t}^1, \dots, A_{n,t}^1] \quad (3.33)$$

L'expression de la sortie peut alors être réécrite sous forme matricielle comme suit :

$$\hat{y}_t^2 = (\tilde{a}_t^2)^T \hat{Y}_{t-1}^2 + (\tilde{A}_t^1)^T X_{t-1}^1 \otimes U_t^1 + (\tilde{b}_t^2)^T U_t^2 \quad (3.34)$$

Les équations d'optimisation s'obtiennent alors par les orthogonalités suivantes :

$$\begin{aligned} E[(y_t - \hat{y}_t^2) \hat{Y}_{t-1}^2] &= 0 \\ E[(y_t - \hat{y}_t^2) X_{t-1}^1 \otimes U_t^1] &= 0 \\ E[(y_t - \hat{y}_t^2) U_t^2] &= 0 \end{aligned} \quad (3.35)$$

En développant l'expression de $\hat{y}_t^2$, on obtient le système linéaire en les paramètres suivant :

$$\begin{aligned} E[\hat{Y}_{t-1}^2 (\hat{Y}_{t-1}^2)^T] \tilde{a}_t + E[\hat{Y}_{t-1}^2 (X_{t-1}^1 \otimes U_t^1)^T] \tilde{A}_t^1 + E[\hat{Y}_{t-1}^2 (U_t^2)^T] \tilde{b}_t &= E[\hat{Y}_{t-1}^2 y_t] \\ E[X_{t-1}^1 \otimes U_t^1 (\hat{Y}_{t-1}^2)^T] \tilde{a}_t + E[(X_{t-1}^1 \otimes U_t^1) (X_{t-1}^1 \otimes U_t^1)^T] \tilde{A}_t^1 + E[X_{t-1}^1 \otimes U_t^1 (U_t^2)^T] \tilde{b}_t &= E[X_{t-1}^1 \otimes U_t^1 y_t] \\ E[U_t^2 (\hat{Y}_{t-1}^2)^T] \tilde{a}_t + E[U_t^2 (X_{t-1}^1 \otimes U_t^1)^T] \tilde{A}_t^1 + E[U_t^2 (U_t^2)^T] \tilde{b}_t &= E[U_t^2 y_t] \end{aligned} \quad (3.36)$$

3.4.2.2 Mise à jour héréditaire

Observons à présent comment calculer récursivement l'ensemble des espérances apparaissant dans ce système linéaire.

- $E[\hat{Y}_t^2 (\hat{Y}_t^2)^T]$ s'obtient à partir de $E[\hat{Y}_{t-1}^2 (\hat{Y}_{t-1}^2)^T]$ en ajoutant à celle-ci la première colonne de celle-la, c'est à dire $E[\hat{Y}_{t-1}^2 \hat{y}_t^2]$. Cet ensemble se ramène à l'ensemble $\{E[\hat{y}_{t-i}^2 y_t] \mid \forall i = 0, \dots, n-1\}$ grâce au principe d'orthogonalité :

$$E[(y_t - \hat{y}_t^2) \hat{Y}_t^2] = 0 \Rightarrow E[\hat{Y}_t^2 \hat{y}_t^2] = E[\hat{Y}_t^2 y_t] \quad (3.37)$$

- $E[\hat{Y}_t^2 (X_t^1 \otimes U_{t+1}^1)^T]$: Cette matrice est composée des sous-vecteurs $E[\hat{y}_{t-i}^2 \hat{x}_{t-j}^1 u_{t-j+1}]$ $\forall i = 0, \dots, n-1, j = 0, \dots, n-1$. Or, si par hypothèse u_t est un bruit blanc on a :

$$E[\hat{y}_{t-i}^2 \hat{x}_{t-j}^1 u_{t-j+1}] = \begin{cases} 0 & \text{si } i \geq j \\ E[y_{t-i} \hat{x}_{t-j}^1 u_{t-j+1}] & \text{si } i < j \end{cases} \quad (3.38)$$

- $E[\hat{Y}_t^2 (U_{t+1}^2)^T]$ est composé des termes $E[\hat{y}_{t-i}^2 (u_{t-j+1})^2]$, $\forall i=0, \dots, n-1, j=0, \dots, n-1$.
Si u_t est un bruit blanc unitaire ($E[(u_t)^2]=1$), alors :

$$E[\hat{y}_{t-i}^2 (u_{t-j+1})^2] = \begin{cases} E[y_{t-i}^2 (u_{t-j+1})^2] = 0 & \text{si } i \geq j \\ E[y_{t-i}^2 (u_{t-j+1})^2] = h^2(j-i-1, j-i-1) & \text{si } i < j \end{cases} \quad (3.39)$$

où $h^2(t, \tau)$ fait partie des données du problème.

- $E[X_t^1 \otimes U_{t+1}^1 (X_t^1 \otimes U_{t+1}^1)^T]$ est composé des sous-matrices $E[\hat{x}_{t-i}^1 (\hat{x}_{t-j}^1)^T u_{t-i+1} u_{t-j+1}]$, $\forall i=0, \dots, n-1, j=0, \dots, n-1$. Une fois encore, puisque u_t est un bruit blanc, on a :

$$E[\hat{x}_{t-i}^1 (\hat{x}_{t-j}^1)^T u_{t-i+1} u_{t-j+1}] = \begin{cases} E[\hat{x}_{t-i}^1 (\hat{x}_{t-i}^1)^T] & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad (3.40)$$

- $E[X_t^1 \otimes U_{t+1}^1 (U_{t+1}^2)^T]$ est composé des sous-vecteurs $E[\hat{x}_{t-i}^1 u_{t-i+1} (u_{t-j+1})^2]$, $\forall i=0, \dots, n-1, j=0, \dots, n-1$. Si la densité de probabilité de u_t est symétrique ($E[(u_t)^3]=0$), on aboutit à :

$$E[\hat{x}_{t-i}^1 u_{t-i+1} (u_{t-j+1})^2] = \begin{cases} E[\hat{x}_{t-i}^1 u_{t-i+1} (u_{t-j+1})^2] & \text{si } i < j \\ 0 & \text{si } i \geq j \end{cases} \quad (3.41)$$

- $E[U_t^2 (U_t^2)^T]$ est composé de $E[(u_{t-i+1})^2 (u_{t-j+1})^2]$, $\forall i=0, \dots, n-1, j=0, \dots, n-1$ et de par le caractère bruit blanc de u_t , est égale à :

$$E[(u_{t-i+1})^2 (u_{t-j+1})^2] = \begin{cases} E[(u_t)^4] & \text{si } i = j \\ 1 & \text{si } i \neq j \end{cases} \quad (3.42)$$

Concernant les seconds membres du système linéaire, il nous faut calculer les éléments suivants :

- $E[\hat{Y}_{t-1}^2 y_t]$ est composé des termes $E[\hat{y}_{t-i}^2 y_t]$ $\forall i=1, \dots, n$
- $E[X_{t-1}^1 \otimes U_t^1 y_t]$ est composé des termes $E[y_t \hat{x}_{t-j}^1 u_{t-j+1}]$ $\forall j=1, \dots, n$:

- $E[U_t^2 y_t]$ fait partie des données du problème et $E[y_t(u_{t-j})^2] = h^2(j, j)$

En définitive, il apparaît que le calcul de l'ensemble des espérances apparaissant dans le système linéaire qui fournit les paramètres à l'instant t se réduit à l'ensemble des espérances suivantes :

$$\begin{aligned}
& E[\hat{y}_{t-i}^2 y_t] \forall i = 0 \dots n \\
& E[y_{t-i} \hat{x}_t^1 u_{t+1}] \forall i = 0, \dots, n-1 \\
& E[\hat{y}_t^2] \\
& E[\hat{x}_t^1 (\hat{x}_t^1)^T] \\
& E[\hat{x}_t^1 u_{t+1} (u_{t-j+1})^2] \forall i < j = 0, \dots, n-1
\end{aligned} \tag{3.43}$$

Parmi ces quantités, certaines sont calculables récursivement et d'autres par un algorithme héréditaire.

- Partie récursive de l'algorithme : calcul de $E[\hat{x}_t^1 (\hat{x}_t^1)^T]$, $E[\hat{y}_t^2]$. A partir de l'écriture récursive de $\hat{x}_t^1$ sous la forme $\hat{x}_t^1 = A_t^1 \hat{x}_{t-1}^1 + b_t u_t$, on obtient :

$$\begin{aligned}
E[\hat{x}_t^1 (\hat{x}_t^1)^T] &= A_t^1 E[\hat{x}_{t-1}^1 (\hat{x}_{t-1}^1)^T] (A_t^1)^T + b_t E[u_t (\hat{x}_{t-1}^1)^T] (A_t^1)^T \\
&\quad + A_t^1 E[\hat{x}_{t-1}^1 u_t] b_t + b_t E[u_t^2] (b_t)^T
\end{aligned} \tag{3.44}$$

En utilisant le fait que u_t est un bruit blanc unitaire, ceci se simplifie en :

$$E[\hat{x}_t^1 (\hat{x}_t^1)^T] = A_t^1 E[\hat{x}_{t-1}^1 (\hat{x}_{t-1}^1)^T] (A_t^1)^T + b_t (b_t)^T \tag{3.45}$$

De même, $\hat{y}_t^2 = (\tilde{a}_t)^T \hat{Y}_{t-1}^2 + (\tilde{A}_t^1)^T X_{t-1}^1 \otimes U_t^1 + (\tilde{b}_t)^T U_t^2$ conduit à

$$\begin{aligned}
E[\hat{y}_t^2] &= (\tilde{a}_t)^T E[\hat{Y}_{t-1}^2] + (\tilde{A}_t^1)^T E[X_{t-1}^1 U_t] + (\tilde{b}_t)^T E[U_t^2] \\
&= (\tilde{a}_t)^T E[\hat{Y}_{t-1}^2] + (\tilde{b}_t)^T E[U_t^2]
\end{aligned} \tag{3.46}$$

où $E[U_t^2] = [1 \quad \dots \quad 1]^T$

- Partie héréditaire de l'algorithme : $E[\hat{y}_{t-i}^2 y_t] \forall i = 0 \dots n$, $E[y_{t-i} \hat{x}_t^1 u_{t+1}] \forall i = 0, \dots, n-1$, et $E[\hat{x}_{t-i}^1 u_{t-i+1} (u_{t-j+1})^2] \forall i = 0, \dots, n-1, j = 0, \dots, n-1$
 - Initialisation avec $E[\hat{y}_0^2 y_t] = 0$
 - Pour $\tau = 1, \dots, t$

- $E[\hat{x}_0^1 u_\tau y_t] = 0, E[\hat{x}_0^1 u_\tau (u_t)^2] = 0$
- Pour $\sigma = 1, \dots, \tau - 1$

$$E[\hat{x}_\sigma^1 u_\tau y_t] = A_\sigma^1 E[\hat{x}_{\sigma-1}^1 u_\tau y_t] + b_\sigma E[y_t u_\sigma u_\tau]$$

$$= A_\sigma^1 E[\hat{x}_{\sigma-1}^1 u_\tau y_t] + b_\sigma h^2(t - \tau, t - \sigma)$$

$$E[\hat{x}_\sigma^1 u_\tau (u_t)^2] = A_\sigma^1 E[\hat{x}_{\sigma-1}^1 u_\tau (u_t)^2] + b_\sigma E[(u_t)^2 u_\sigma u_\tau]$$

$$= A_\sigma^1 E[\hat{x}_{\sigma-1}^1 u_\tau (u_t)^2] + b_\sigma E[(u_t)^4]$$

- $E[\hat{y}_\tau^2 y_t] = (\tilde{a}_\tau)^T E[\hat{Y}_{\tau-1}^2 y_t] + (\tilde{A}_\tau^1)^T E[X_{\tau-1}^1 \otimes U_\tau^1 y_t] + (\tilde{b}_\tau)^T E[U_\tau^2 y_t]$

Il apparaît donc que dans ce cas l'algorithme soit en quelque sorte « doublement » héréditaire du fait de la présence d'une boucle de $\sigma = 1, \dots, \tau - 1$ à l'intérieur de la boucle $\tau = 1, \dots, t$. De fait, on constate que l'ensemble des données $h^2(t, \tau)$ sont effectivement balayées par l'algorithme de réalisation.

3.4.2.3 Algorithme

Nous reprenons ici l'ensemble de l'algorithme de réalisation décrit au paragraphe précédent. La notation $[X]_i^j$ désigne la coordonnée (i, j) de la matrice X et $[X]_{i_1:i_2}^{j_1:j_2}$ désigne la sous matrice de la matrice X définie par :

$$[X]_{i_1:i_2}^{j_1:j_2} = \{X_i^j, i = i_1, \dots, i_2, j = j_1, \dots, j_2\}$$

$I^{n \times n}$ désigne la matrice identité de dimension n

1 Initialisation à $t = 1$: toutes les espérances sont fixées à zéros sauf les suivantes :

$$\forall i, j = 1, \dots, n, \left[E[\hat{Y}_0^2 (U_1^2)^T] \right]^j = \begin{cases} 0 & \text{if } i \geq j \\ h^2(j - i - 1, j - i - 1) & \text{if } i < j \end{cases}$$

$$E[U_0^2 (U_0^2)^T] = E[(u_1)^4] \times I^{n \times n}$$

$$\forall i = 1, \dots, n, [E[U_1^2 y_1]]_i = h^2(i - 1, i - 1)$$

2 Inversion du système linéaire

$$\begin{aligned}
& E\left[\hat{Y}_{t-1}^2(\hat{Y}_{t-1}^2)^T\right]\tilde{a}_t + E\left[\hat{Y}_{t-1}^2(X_{t-1}^1 \otimes U_t^1)^T\right]\tilde{A}_t^1 + E\left[\hat{Y}_{t-1}^2(U_t^2)^T\right]\tilde{b}_t = E\left[\hat{Y}_{t-1}^2 y_t\right] \\
& E\left[X_{t-1}^1 \otimes U_t^1(\hat{Y}_{t-1}^2)^T\right]\tilde{a}_t + E\left[X_{t-1}^1 \otimes U_t^1(X_{t-1}^1 \otimes U_t^1)^T\right]\tilde{A}_t^1 + E\left[X_{t-1}^1 \otimes U_t^1(U_t^2)^T\right]\tilde{b}_t = E\left[X_{t-1}^1 \otimes U_t^1 y_t\right] \\
& E\left[U_t^2(\hat{Y}_{t-1}^2)^T\right]\tilde{a}_t + E\left[U_t^2(X_{t-1}^1 \otimes U_t^1)^T\right]\tilde{A}_t^1 + E\left[U_t^2(U_t^2)^T\right]\tilde{b}_t = E\left[U_t^2 y_t\right]
\end{aligned}$$

et calcul des paramètres $\tilde{a}_t$, $\tilde{A}_t^1$ et $\tilde{b}_t$. Ces valeurs sont stockées en mémoire.

3 Calcul récursif de $E\left[\hat{x}_t^1(\hat{x}_t^1)^T\right]$, $E\left[\hat{y}_t^2\right]$ selon :

$$E\left[\hat{x}_t^1(\hat{x}_t^1)^T\right] = \tilde{A}_t^1 E\left[\hat{x}_{t-1}^1(\hat{x}_{t-1}^1)^T\right](\tilde{A}_t^1)^T + b_t(b_t)^T$$

$$E\left[\hat{y}_t^2\right] = (\tilde{a}_t)^T E\left[\hat{Y}_{t-1}^2\right] + (\tilde{b}_t)^T I^{n \times n}$$

$$E\left[\hat{Y}_t^2\right] = \left[E\left[\hat{y}_t^2\right], E\left[\hat{Y}_{t-1}^2\right] \right]_{1:n-1}^T$$

4 Mise à jour de $E\left[X_t^1 \otimes U_{t+1}^1(X_t^1 \otimes U_{t+1}^1)^T\right]$

$$\left[E\left[X_t^1 \otimes U_{t+1}^1(X_t^1 \otimes U_{t+1}^1)^T\right] \right]_{1:n}^{1:n} = E\left[\hat{x}_t^1(\hat{x}_t^1)^T\right]$$

$$\left[E\left[X_t^1 \otimes U_{t+1}^1(X_t^1 \otimes U_{t+1}^1)^T\right] \right]_{n+1:n^2}^{n+1:n^2} = \left[E\left[X_{t-1}^1 \otimes U_t^1(X_{t-1}^1 \otimes U_t^1)^T\right] \right]_{1:(n^2-n)}^{1:(n^2-n)}$$

5 Mise à jour de $E\left[\hat{Y}_{t-1}^2(U_t^2)^T\right]$

$$\left[E\left[\hat{Y}_{t-1}^2(U_t^2)^T\right] \right]_{1:n}^{1:n} = \left[E\left[\hat{Y}_{t-1}^2\right] \right]_{1:n}^{1:n}$$

6. Calcul héréditaire de $E\left[\hat{y}_{t-i+1}^2 y_{t+1}\right] \forall i=0..n$, $E\left[\hat{x}_{t-i}^1 u_{t-i+1} y_{t+1}\right] \forall i=0, \dots, n-1$,

et $E\left[\hat{x}_{t-i}^1 u_{t-i+1} (u_{t-j+1})^2\right] \forall i=0, \dots, n-1, j=0, \dots, n-1$

- Initialisation avec $E\left[\hat{y}_0^2 y_{t+1}\right] = 0$
- Pour $\tau = 1, \dots, t+1$
 - $E\left[\hat{x}_0^1 u_\tau y_{t+1}\right] = 0, E\left[\hat{x}_0^1 u_\tau (u_{t+1})^2\right] = 0$
 - Pour $\sigma = 1, \dots, \tau-1$

$$E\left[\hat{x}_\sigma^1 u_\tau y_{t+1}\right] = A_\sigma^1 E\left[\hat{x}_{\sigma-1}^1 u_\tau y_{t+1}\right] + b_\sigma h^2(t-\tau, t-\sigma)$$

$$E\left[\hat{x}_\sigma^1 u_\tau (u_{t+1})^2\right] = A_\sigma^1 E\left[\hat{x}_{\sigma-1}^1 u_\tau (u_{t+1})^2\right] + b_\sigma E\left[(u_t)^4\right]$$

- $\left[E \left[U_{\tau}^2 y_{t+1} \right] \right] = h^2(i, i)$
- $E \left[\hat{y}_{\tau}^2 y_{t+1} \right] = (\tilde{a}_{\tau})^T E \left[\hat{Y}_{\tau-1}^2 y_{t+1} \right] + (\tilde{A}_{\tau}^1)^T E \left[X_{\tau-1}^1 \otimes U_{\tau}^1 y_{t+1} \right] + (\tilde{b}_{\tau})^T E \left[U_{\tau}^2 y_{t+1} \right]$

1 Mise à jour de $E \left[X_{t-1}^1 \otimes U_t^1 y_t \right]$ et $E \left[\hat{Y}_{t-1}^2 y_t \right]$

$$E \left[X_t^1 \otimes U_{t+1}^1 y_{t+1} \right] = \left[E \left[\hat{x}_t^1 u_{t+1} y_{t+1} \right]^T, \dots, E \left[\hat{x}_{t-n+1}^1 u_{t-n+2} y_{t+1} \right]^T \right]^T$$

$$E \left[\hat{Y}_t^2 y_{t+1} \right] = \left[E \left[\hat{y}_t^2 y_{t+1} \right], \dots, E \left[\hat{y}_{t-n+1}^2 y_{t+1} \right] \right]^T$$

2 $t+1 \rightarrow t$ et retour en 2.

3.5 Identification des systèmes bilinéaires homogènes

3.5.1 Principe général

L'objectif poursuivi consiste donc à trouver un système dynamique qui explique « au mieux » le comportement entrée/sortie dont on connaît un jeu de trajectoires d'entrées et de sorties. En d'autres termes, on se donne préalablement une structure paramétrique de système dynamique sous la forme :

$$\begin{aligned} x_t &= f(x_{t-1}, u_t, \theta) \\ y_t &= h(x_t) + v_t \end{aligned} \quad (3.47)$$

où u_t désigne l'entrée, y_t la sortie, v_t le bruit (blanc) d'observation et θ le paramètre inconnu. On recherche alors dans cette classe le paramètre θ qui minimise un certain critère de qualité. De manière générale, ce critère reflète les capacités de prédiction du modèle.

Le principe d'identification héréditaire est de considérer le prédicteur variable avec le temps, comme c'est le cas en filtrage des systèmes dynamiques (le filtre de Kalman, par exemple). Dans ce cas, le prédicteur aura la forme suivante :

$$\begin{aligned} \hat{x}_{t|t-1} &= f(\hat{x}_{t-1|t-2}, u_t, \theta_t) \\ \hat{y}_{t|t-1} &= h(\hat{x}_{t|t-1}) \end{aligned} \quad (3.48)$$

et le critère à minimiser, en référence au bruit d'observation gaussien, s'écrit généralement

$$E^t \left[\left(y_{\bullet} - \hat{y}_{\bullet|t-1} \right)^2 \right] = \frac{1}{t} \sum_{\tau=1}^t \left(y_{\tau} - \hat{y}_{\tau|t-1} \right)^2 \quad (3.49)$$

En effet, si le bruit d'observation est gaussien, ce critère revient à celui du maximum de vraisemblance.

Pour approcher cet estimateur, au sens de l'estimation ergodique, on considère une suite de trajectoires du processus définies sur l'horizon $[1, t]$ sous la forme suivante :

$$\begin{aligned} \hat{x}_{\tau|t-1}^t &= f(\hat{x}_{\tau-1|t-2}^{t-1}, u_{\tau}, \theta_t), \quad \forall \tau = 1, \dots, t \\ \hat{y}_{\tau|t-1}^t &= h(\hat{x}_{\tau|t-1}^t) \end{aligned} \quad (3.50)$$

qui s'appuient sur la trajectoire réalisée à l'instant précédent sur l'horizon $[1, t-1]$ et on cherche la valeur du paramètre θ_t qui minimise le critère.

Pour que l'optimisation du critère quadratique donne lieu à un problème convexe, il suffit que le paramètre θ apparaisse linéairement dans l'expression de la sortie y_t . Considérons

donc la classe des systèmes bilinéaires homogènes à paramétrisation linéaire définie au paragraphe précédant comme :

$\forall \tau = 1, \dots, t :$

$$\begin{aligned}\hat{y}_{\tau|\tau-1}^{t,d} &= \sum_{i=1}^n a_i^{t,d} \hat{y}_{\tau-i|\tau-i-1}^{t-i,d} + \sum_{i=1}^n \sum_{k=1}^{d-1} A_i^{t,k} \hat{x}_{\tau-i|\tau-i-1}^{t-i,k} (u_{\tau-i+1})^{d-k-1} + \sum_{i=1}^n b_i^t (u_{\tau-i+1})^d \\ x_{\tau|\tau-1}^{t,k} &= \sum_{j=1}^k A^{t,j} \hat{x}_{\tau-1|\tau-2}^{t-1,j} (u_{\tau})^{k-j} + b^t (u_{\tau})^k, \quad \forall k = 1, \dots, d-1\end{aligned}\tag{3.51}$$

Pour simplifier les écritures, nous écrirons par la suite $\hat{y}_{\tau}^{t,d}$ en lieu et place $\hat{y}_{\tau|\tau-1}^{t,d}$ ainsi que $x_{\tau}^{t,k}$ à la place de $x_{\tau|\tau-1}^{t,k}$.

De manière analogue au cas réalisation, définissons les trajectoires suivantes :

$$\begin{aligned}\hat{Y}_{\tau-1}^{t-1,d} &= [\hat{y}_{\tau-1}^{t-1,d}, \dots, \hat{y}_{\tau-n}^{t-1,d}]^T \\ U_{\tau}^d &= [(u_{\tau})^d, \dots, (u_{\tau-n+1})^d]^T \\ \hat{X}_{\tau-1}^{t-1,k} \otimes U_{\tau}^{d-k-1} &= \left[(\hat{x}_{\tau-1}^{t-1,k})^T u_{\tau}, \dots, (\hat{x}_{\tau-n}^{t-1,k})^T u_{\tau-n+1} \right]^T \\ a_t^d &= [a_1^{t,d}, \dots, a_n^{t,d}]^T \\ b_t &= [b_1^t, \dots, b_n^t]^T\end{aligned}\tag{3.52}$$

et le vecteur de dimension n^2 suivant :

$$\tilde{A}^{t,k} = [A_1^{t,k}, \dots, A_n^{t,k}]^T\tag{3.53}$$

L'expression de la sortie peut alors être réécrite sous forme matricielle comme suit :

$$\hat{y}_{\tau}^{t,d} = (a_t^d)^T \hat{Y}_{\tau-1}^{t-1,d} + \sum_{k=1}^{d-1} (\tilde{A}^{t,k})^T \hat{X}_{\tau-1}^{t-1,k} \otimes U_{\tau}^{d-k-1} + (b_t)^T U_{\tau}^d\tag{3.54}$$

L'optimisation des paramètres conduit alors aux équations d'orthogonalité suivantes :

$$\begin{aligned}E^t \left[(y_{\tau} - \hat{y}_{\tau}^{t,d}) \hat{Y}_{\tau-1}^{t-1,d} \right] &= 0 \\ E^t \left[(y_{\tau} - \hat{y}_{\tau}^{t,d}) \hat{X}_{\tau-1}^{t-1,l} \otimes U_{\tau}^{d-l-1} \right] &= 0, \quad \forall l = 1, \dots, d-1 \\ E^t \left[(y_{\tau} - \hat{y}_{\tau}^{t,d}) U_{\tau}^d \right] &= 0\end{aligned}\tag{3.55}$$

En développant l'expression de $\hat{y}_{\tau}^{t,d}$, on obtient le système linéaire en les paramètres suivant :

$$\begin{aligned}
E^t \left[\hat{Y}_{\bullet-1}^{t-1,d} \left(\hat{Y}_{\bullet-1}^{t-1,d} \right)^T \right] a_t^d + \sum_{k=1}^{d-1} E^t \left[\hat{Y}_{\bullet-1}^{t-1,d} \left(\hat{X}_{\bullet-1}^{t-1,k} \otimes U_{\bullet}^{d-k-1} \right)^T \right] \tilde{A}^{t,k} + E^t \left[\hat{Y}_{\bullet-1}^{t-1,d} \left(U_{\bullet}^d \right)^T \right] p_t = E^t \left[\hat{Y}_{\bullet-1}^{t-1,d} y_{\bullet} \right] \\
E^t \left[\hat{X}_{\bullet-1}^{t-1,l} \otimes U_{\bullet}^{d-l-1} \left(\hat{Y}_{\bullet-1}^{t-1,d} \right)^T \right] a_t^d + \sum_{k=1}^{d-1} E^t \left[\hat{X}_{\bullet-1}^{t-1,l} \otimes U_{\bullet}^{d-l-1} \left(\hat{X}_{\bullet-1}^{t-1,k} \otimes U_{\bullet}^{d-k-1} \right)^T \right] \tilde{A}^{t,k} \\
+ E^t \left[\hat{X}_{\bullet-1}^{t-1,l} \otimes U_{\bullet}^{d-l-1} \left(U_{\bullet}^d \right)^T \right] p_t = E^t \left[\hat{X}_{\bullet-1}^{t-1,l} \otimes U_{\bullet}^{d-l-1} y_{\bullet} \right] \\
E^t \left[U_{\bullet}^d \left(\hat{Y}_{\bullet-1}^{t-1,d} \right)^T \right] a_t^d + \sum_{k=1}^{d-1} E^t \left[U_{\bullet}^d \left(\hat{X}_{\bullet-1}^{t-1,k} \otimes U_{\bullet}^{d-k-1} \right)^T \right] \tilde{A}^{t,k} + E^t \left[U_{\bullet}^d \left(U_{\bullet}^d \right)^T \right] p_t = E^t \left[U_{\bullet}^d y_{\bullet} \right]
\end{aligned} \tag{3.56}$$

3.5.2 Algorithme héréditaire

L'algorithme héréditaire est structuré ainsi :

1. Initialisation : $t = 1$. Tous les termes du système linéaire sont fixés à zéro sauf :

$$\left[E^1 \left[U_{\bullet}^d \left(U_{\bullet}^d \right)^T \right] \right] = (u_1)^{2d}$$

$$\left[E^1 \left[U_{\bullet}^d y_{\bullet} \right] \right] = (u_1)^d y_1$$

2. Inversion du système linéaire défini par l'ensemble d'équations suivantes :

$$E^t \left[\hat{Y}_{\bullet-1}^{t-1,d} \left(\hat{Y}_{\bullet-1}^{t-1,d} \right)^T \right] a_t^d + \sum_{k=1}^{d-1} E^t \left[\hat{Y}_{\bullet-1}^{t-1,d} \left(\hat{X}_{\bullet-1}^{t-1,k} \otimes U_{\bullet}^{d-k-1} \right)^T \right] \tilde{A}^{t,k} + E^t \left[\hat{Y}_{\bullet-1}^{t-1,d} \left(U_{\bullet}^d \right)^T \right] p_t = E^t \left[\hat{Y}_{\bullet-1}^{t-1,d} y_{\bullet} \right]$$

$$\forall l = 1, \dots, d-1,$$

$$E^t \left[\hat{X}_{\bullet-1}^{t-1,l} \otimes U_{\bullet}^{d-l-1} \left(\hat{Y}_{\bullet-1}^{t-1,d} \right)^T \right] a_t^d + \sum_{k=1}^{d-1} E^t \left[\hat{X}_{\bullet-1}^{t-1,l} \otimes U_{\bullet}^{d-l-1} \left(\hat{X}_{\bullet-1}^{t-1,k} \otimes U_{\bullet}^{d-k-1} \right)^T \right] \tilde{A}^{t,k}$$

$$+ E^t \left[\hat{X}_{\bullet-1}^{t-1,l} \otimes U_{\bullet}^{d-l-1} \left(U_{\bullet}^d \right)^T \right] p_t = E^t \left[\hat{X}_{\bullet-1}^{t-1,l} \otimes U_{\bullet}^{d-l-1} y_{\bullet} \right]$$

$$E^t \left[U_{\bullet}^d \left(\hat{Y}_{\bullet-1}^{t-1,d} \right)^T \right] a_t^d + \sum_{k=1}^{d-1} E^t \left[U_{\bullet}^d \left(\hat{X}_{\bullet-1}^{t-1,k} \otimes U_{\bullet}^{d-k-1} \right)^T \right] \tilde{A}^{t,k} + E^t \left[U_{\bullet}^d \left(U_{\bullet}^d \right)^T \right] p_t = E^t \left[U_{\bullet}^d y_{\bullet} \right]$$

3. Calcul des nouvelles trajectoires de prédiction $\hat{y}_{\tau}^{t,d}$ et $x_{\tau}^{t,k}$

- a. Pour $\tau = 1, \dots, t$,

$$\hat{y}_{\tau}^{t,d} = \sum_{i=1}^n a_i^{t,d} \hat{y}_{\tau-i}^{t-1,d} + \sum_{i=1}^n \sum_{k=1}^{d-1} A_i^{t,k} \hat{x}_{\tau-i}^{t-1,k} (u_{\tau-i+1})^{d-k-1} + \sum_{i=1}^n b_i^t (u_{\tau-i+1})^d$$

- b. Pour $k = 1, \dots, d-1$ et pour $\tau = 1, \dots, t$

$$x_{\tau}^{t,k} = \sum_{j=1}^k A^{t,j} \hat{x}_{\tau-1}^{t-1,j} (u_{\tau})^{k-j} + b^t (u_{\tau})^k$$

Ces nouvelles trajectoires sont mémorisées et les trajectoires $\hat{y}_{\tau}^{t-n,d}$ et $x_{\tau}^{t-n,k}$ sont effacées de la mémoire.

4. Mise à jour des termes du système linéaire :

$$E^{t+1} \left[\hat{Y}_{\bullet-1}^{t,d} \left(\hat{Y}_{\bullet-1}^{t,d} \right)^T \right] = \frac{1}{t+1} \sum_{\tau=1}^{t+1} \hat{Y}_{\bullet-1}^{t,d} \left(\hat{Y}_{\bullet-1}^{t,d} \right)^T$$

...

5. Retour en 2

Remarque : Durant les premières étapes, le système linéaire est naturellement non inversible. Il convient donc de réaliser une pseudo-inversion, non préjudiciable à l'algorithme dans la mesure où la solution est indéterminée, et toute solution du système est admissible.

Remarque : Contrairement au problème de la réalisation, il n'y a pas ici d'aspect « multi-héréditaire » avec des boucles imbriquées.

4. Exemples de mise en oeuvre en Identification Héréditaire

4.1 Exemple linéaire

Afin d'illustrer la performance de l'algorithme héréditaire, nous avons testé celui-ci sur un jeu de données récolté sur un échangeur thermique issu de la base de données DAISY (96-0111). L'algorithme d'identification travaille ici avec un système linéaire de dimension $n = 4$. Nous avons comparé les performances de l'algorithme héréditaire aux résultats obtenus avec la « toolbox identification » de MATLAB : la procédure `n4sid.m` pour la méthode des sous-espaces et les procédures `armax.m` et `oe.m` pour les techniques de type Gauss-Newton. La figure 4.1 montre l'erreur relative de sortie :

$$\sqrt{\sum_{\tau=1}^t (y_{\tau} - \hat{y}_{\tau}^t)^2} / \sqrt{\sum_{\tau=1}^t y_{\tau}^2} \quad (4.1)$$

calculée en fonction du nombre de données traitées. Les erreurs sont fixées à la valeur zéro lorsque le système identifié est instable (en fait, erreur infinie). Il apparaît très clairement sur cette figure que l'algorithme héréditaire donne les meilleures performances en termes de capacité de prédiction. L'erreur se stabilise au environ de 10% avec seulement 140 données. A noter que dans la phase transitoire, l'algorithme donne toujours des systèmes stables ce qui n'est largement pas le cas pour la technique des sous-espaces.

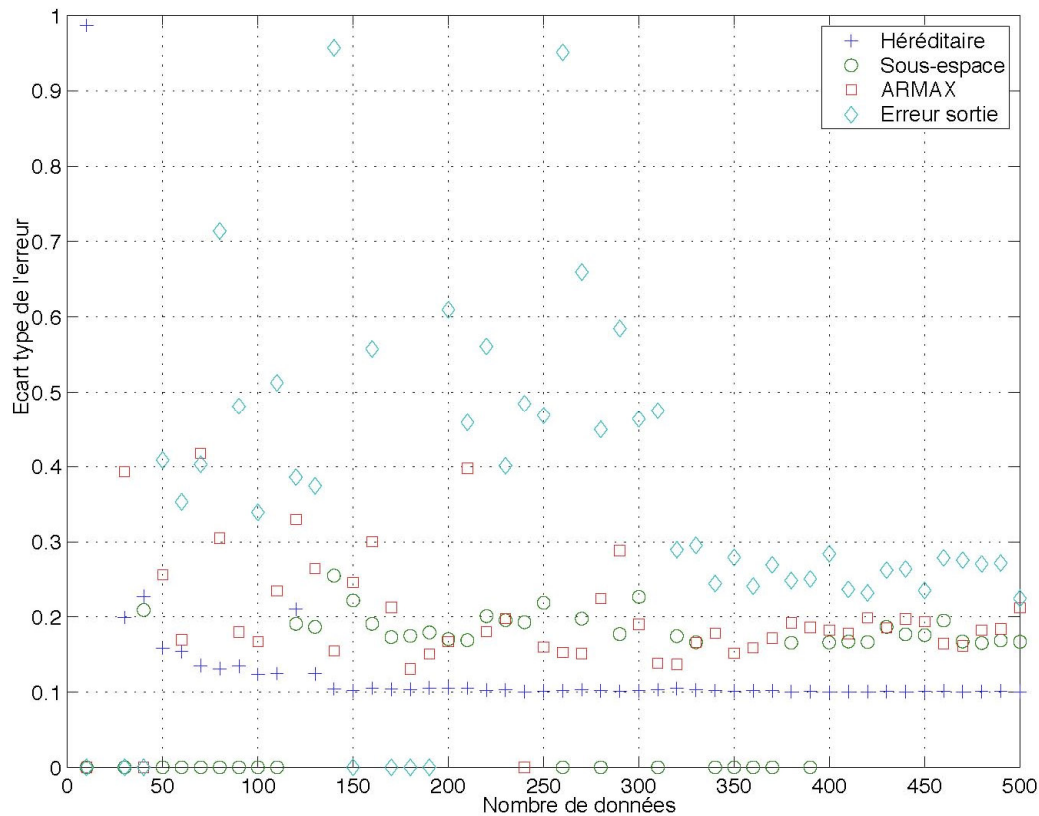


Figure 4.1 Comparaisons des performances des algorithmes

De plus, afin d'illustrer le comportement de l'algorithme en ce qui concerne l'estimation des paramètres, nous avons considéré un système de dimension 6 avec 5 pôles et 6 zéros avec un rapport signal/bruit de 12 dB. L'estimation des deux premiers coefficients du numérateur de ce système par la méthode héréditaire et par une méthode de type Gauss-Newton est illustrée figure 4.2. A nouveau, on constate qu'après environ 150 données, les paramètres estimés par l'approche héréditaire se stabilisent autour des bonnes valeurs alors qu'apparemment, l'algorithme à base de gradient est à plusieurs reprises piégé par un minimum local.

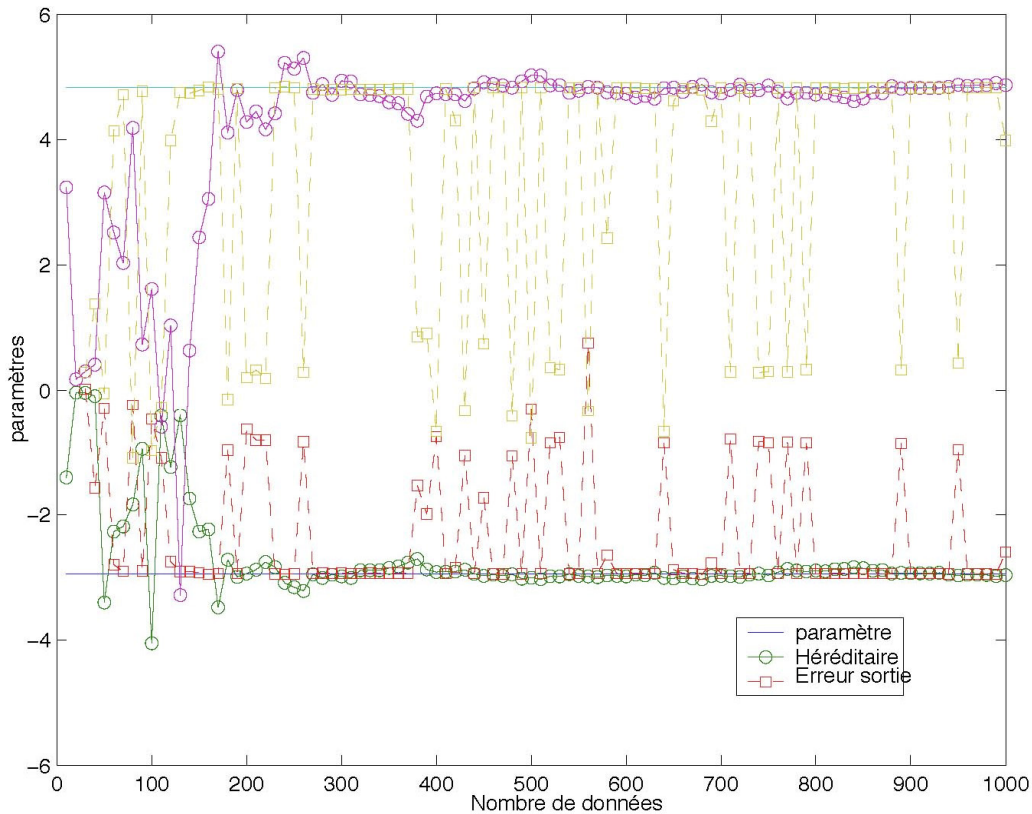


Figure 4.2 Estimations des paramètres

4.2 Exemple quadratique

4.2.1 Système scalaire quadratique

Nous commençons pour comparer l'application de l'approche héréditaire sur un système très simple scalaire quadratique, donc avec dimensions $n_1=n_2=1$ et coefficients $a_1^2=-0,8$, $a_1^1=-0,8$, $b_1=0,8$:

$$\begin{aligned}y_t^1 &= a_1^1 y_{t-1}^1 + b_1 u_t \\ y_t^2 &= a_1^2 y_t^2 + y_t^1 u_t\end{aligned}$$

A la sortie y_t^2 est ajouté un bruit blanc gaussien de sortie pour obtenir un rapport signal sur bruit d'environ 5 décibels. Nous avons utilisé comme entrée un bruit blanc uniforme. Nous avons alors comparé les performances des méthodes de moindres carrés étendus (MCE), d'erreur de prédiction récursif (MEPR), et héréditaire. On a employé les estimateurs correspondant aux formes décrites dans [Fna87].

Les figures suivantes montrent l'estimation des paramètres obtenus par les différents algorithmes. Il apparaît clairement que l'algorithme héréditaire converge rapidement par rapport aux algorithmes classiques, et ce, vers les vraies valeurs (sans biais).

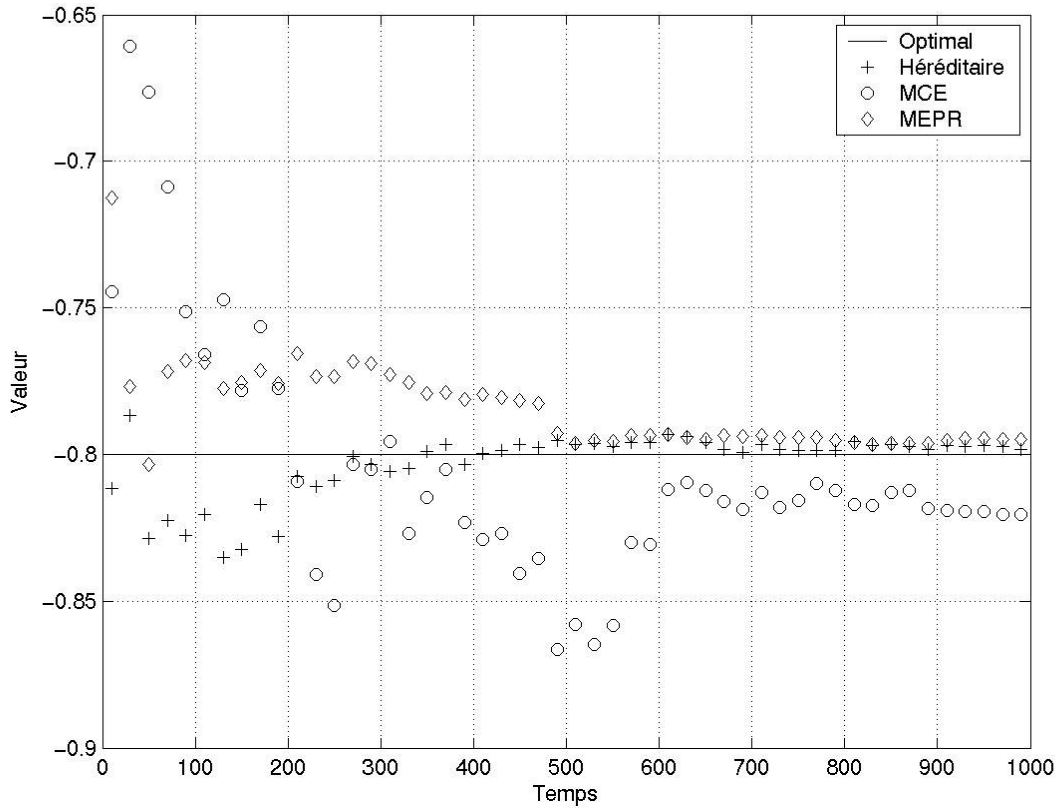


Figure 4.3 Paramètre a_1^2 .

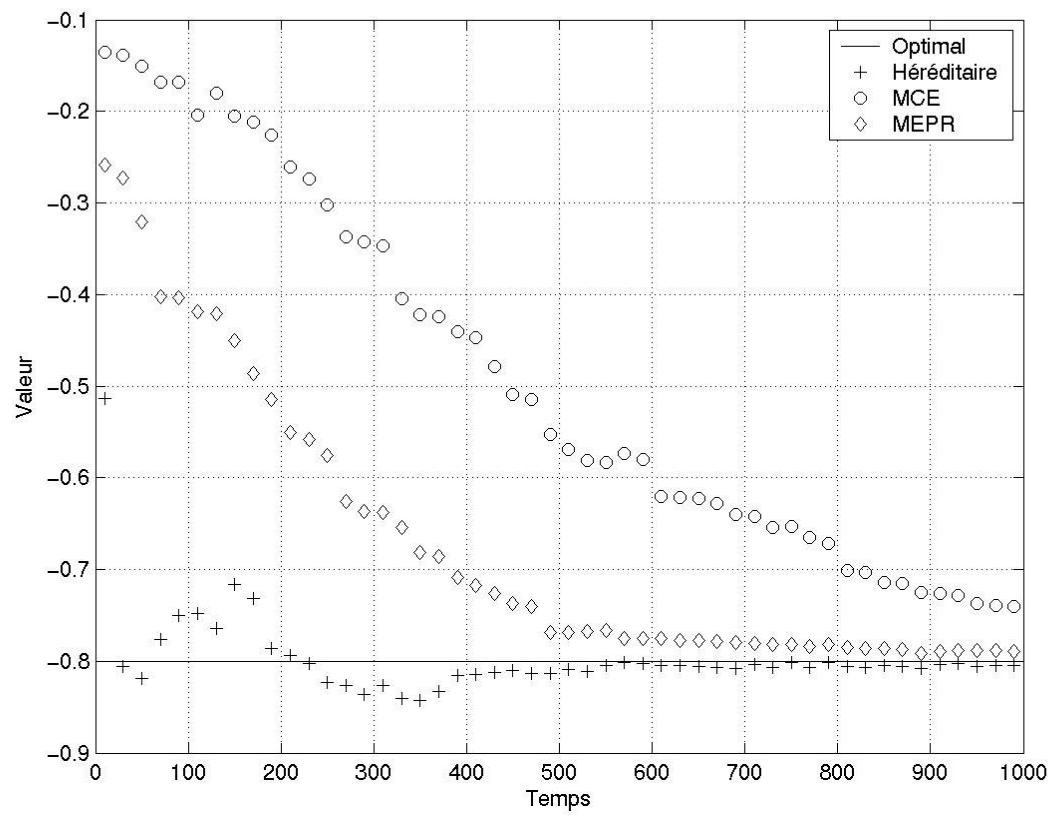


Figure 4.4 Paramètre a_1^1 .

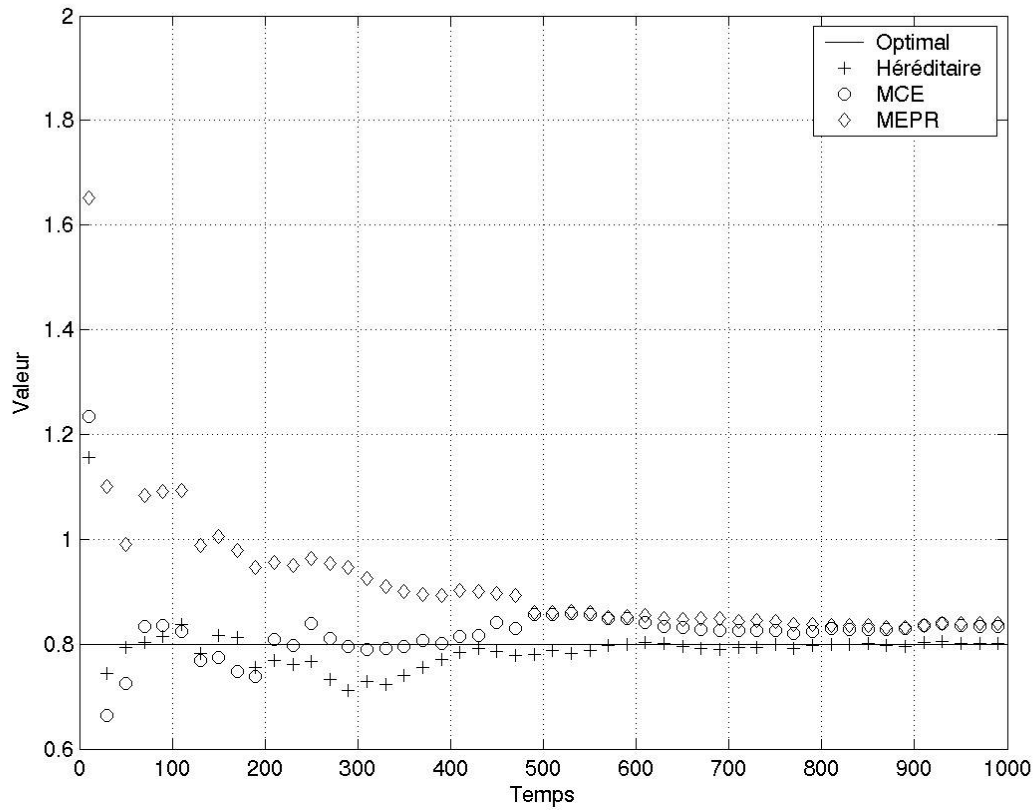


Figure 4.5 Paramètre b_1 .

A nouveau, on observe une convergence beaucoup plus rapide et sans biais de l'estimation des paramètres du système avec l'algorithme héréditaire comparé aux algorithmes classiques.

4.2.2 Approximation d'une Série de Volterra de Degré Infini

Le modèle d'un système non linéaire de dimension finie peut se définir comme un modèle entrée-sortie général du type :

$$y_t = f(y_{t-1}, \dots, y_{t-n_y}, u_t, \dots, u_{t-n_u}) \quad (4.2)$$

où $f(\cdot)$ est une fonction non linéaire quelconque.

Notons que le modèle bilinéaire entrée-sortie que nous manipulons est un cas particulier de ces systèmes non linéaires plus généraux. En plus de sa simplicité, cette classe de modèles est bien adaptée pour de nombreuses applications [Mohl73, Mohl80].

L'équation entrée-sortie générale d'un système bilinéaire est classiquement donnée par [Ki Baik93] :

$$y_t = \sum_{i=1}^{n_y} a_i y_{t-i} + \sum_{j=0}^{n_u} b_j u_{t-j} + \sum_{i=1}^{n_y} \sum_{j=0}^{n_u} c_{ij} y_{t-i} u_{t-j} \quad (4.3)$$

A noter que ce système n'appartient pas à la classe des systèmes homogènes présentés dans le chapitre 2, et donne par conséquent lieu à une série de Volterra de degré infini.

En choisissant les ordres du système $n_y = n_u = 1$, on obtient comme modèle particulier :

$$y_t = a_1 y_{t-1} + b_0 u_t + b_1 u_{t-1} + c_{10} y_{t-1} u_t + c_{11} y_{t-1} u_{t-1}$$

où nous avons fixé les paramètres aux valeurs suivantes $a_1=0.5$, $b_0=b_1=1$, $c_{10}=0.3$, $c_{11}=0.1$.

On a utilisé comme excitation en entrée un bruit blanc uniforme [Ln87]. Un bruit blanc Gaussien a été ajouté à la sortie pour obtenir un rapport signal/bruit $\text{SNR} \cong 5\text{dB}$ $y_{\text{bruitée}}^t = y_t + v_t$, et on a choisi les ordres de séparabilité de l'estimateur égaux $n_1 = n_2 = n$. La classe de système paramétrique a alors la structure suivante [Etchev06] :

$$\begin{aligned} z_t^1 &= \sum_{i=1}^n a_i^{1,1} z_{t-i}^{1,t-1} + \sum_{i=1}^n b_i^1 u_{t-i+1} \\ y_t^1 &= \sum_{i=1}^n a_i^{1,2} \hat{y}_{t-i}^{1,t-1} + \sum_{i=1}^n b_i^2 u_{t-i+1} \\ y_t^2 &= \sum_{i=1}^n a_i^{2,2} y_{t-i}^2 + y_t^1 u + \sum_{i=1}^n \sum_{j=1}^n a_{i,j}^1 y_{t-i-j+1}^1 u_{t-i+1} \\ y_t &= z_t^1 + y_t^2 \end{aligned} \quad (4.4)$$

Naturellement, dans ce cas, les valeurs des paramètres ne sont pas comparables. Aussi, nous utiliserons comme critère de qualité la capacité du système en prédiction sous la forme suivante :

$$R = \left(1 - \frac{\sum_{t=1}^N |y_t - \hat{y}_t|^2}{\sum_{t=1}^N |y_t|^2} \right) \times 100 \quad (4.5)$$

où R est la proportion de la sortie y_t qui est correctement approché par $\hat{y}_t$. Il est nommé *coefficient de corrélation multiple* (au carré) et est souvent donné en pourcentage [Lju99].

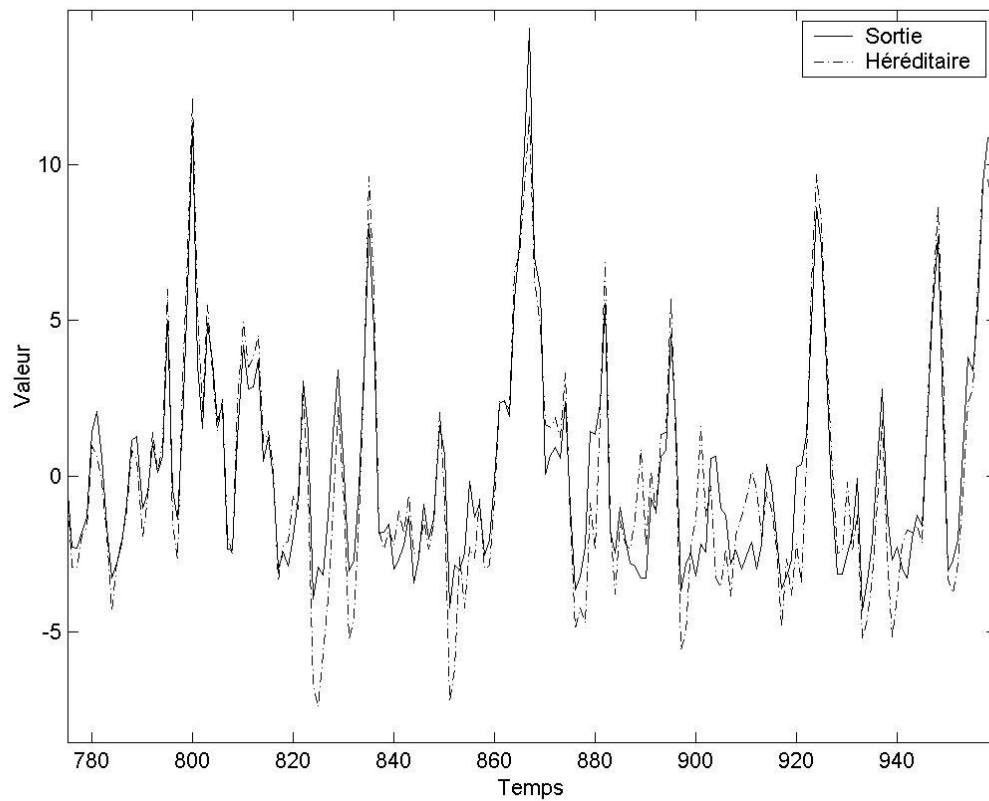


Figure 4.6 Approximation de la sortie sans bruit du modèle non linéaire par (4.4).
Identification/validation (50% des échantillons pour chaque opération).

Le modèle a été approché dans un premier temps sans bruit additif à la sortie afin de mesurer l'efficacité de notre algorithme (voir Figure 4.6). On a pu vérifier qu'avec une dimension $n=3$, l'approximation du modèle non linéaire donne un coefficient $R= 89,25\%$.

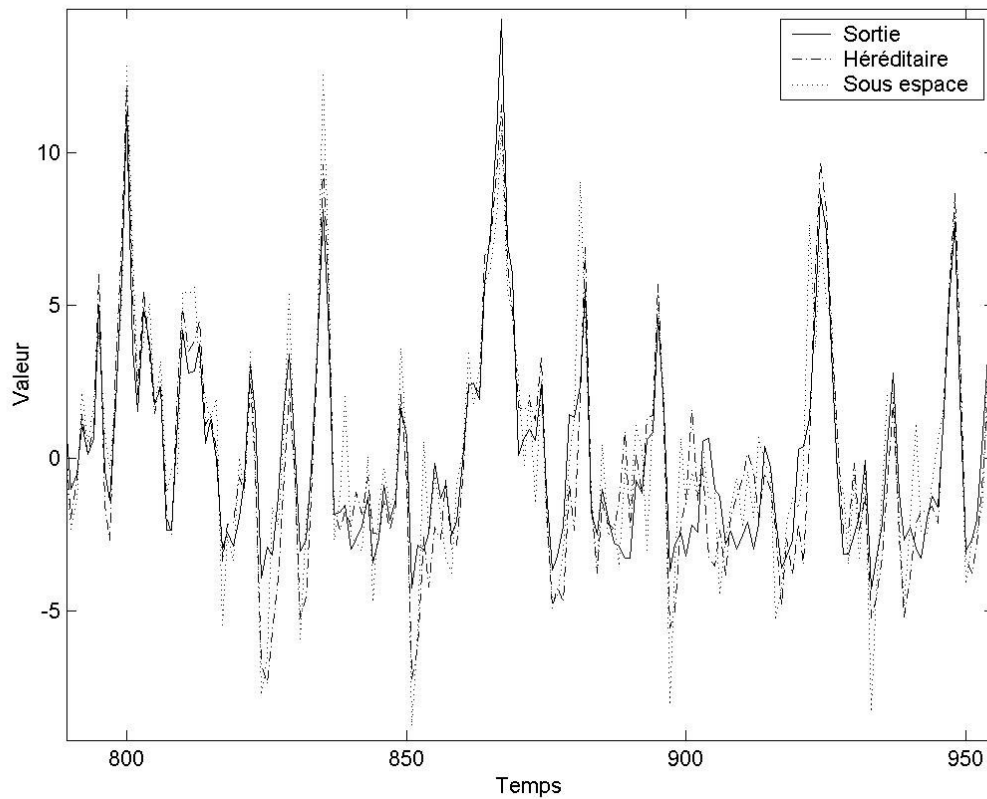


Figure 4.7 Approximation de la sortie bruitée du modèle non linéaire avec les méthodes héréditaire et des sous-espaces. Identification/validation (50% des échantillons pour chaque opération).

La comparaison entre l’algorithme héréditaire et de sous-espaces est illustrée figure 4.7. Le modèle utilisé pour l’algorithme des sous-espaces est aussi bilinéaire homogène avec des matrices sans structure déterminée [Sulei06]. On peut noter sur cette figure que l’approche des sous-espaces (pointillé) en relation à la sortie du modèle non linéaire est moins cadrée sur les sauts et les creux par rapport à l’approche héréditaire (trait pointillé).

Identification du modèle non linéaire			
Rapport SNR 5dB	Points pour Identification et Validation	R (%)	Dimension Totale Lin+Quadr
Héréditaire	50 / 50 (%)	87.15	3+3
Sous- espaces		78.75	4+4

Tableau 4.1

Le Tableau 4.1 montre les dimensions utilisées et le résultat d’approximation R pour les deux techniques d’identification mentionnées. Il en résulte que l’algorithme héréditaire est plus performant de presque 10% avec une dimension total moindre dans la forme polynomiale par rapport à l’approche des sous-espaces.

Les résultats de l'identification du modèle de degré infini par le modèle de degré fini (deux) montrent le besoin d'un ordre de séparabilité égal à 3, comme indiqué dans le tableau 4.1. Ceci a été décidé simplement en augmentant progressivement la dimension du modèle, jusqu'à l'approximation maximale de la sortie du système et en observant les variations du coefficient de corrélation multiple. Si l'approximation ne s'améliore pas significativement à partir d'une certaine dimension, on reste avec la dimension la plus basse pour utiliser ainsi un nombre moindre des coefficients.

Il est clair qu'au fur et à mesure que la dimension employée augmente, la charge de calcul devient plus lourde, même si dans l'utilisation de l'algorithme héréditaire la mémoire employée augmente linéairement à chaque pas [Mon96a]. Pour la série de Volterra quadratique, avec la forme canonique interne linéaire plus quadratique, le nombre de coefficients employés dans l'identification de systèmes est de taille $2n+n^2$.

Conclusion

La contribution principale de ce mémoire réside dans la mise en œuvre d'algorithmes héréditaires destinés à résoudre les problèmes de réalisation et d'identification des systèmes non linéaires. Un des points clés de cette extension s'appuie sur la dérivation d'une forme canonique adaptée à cet approche, laquelle assure un nombre minimal des coefficients à estimer et donne une forme particulière aux matrices du système générant le noyau d'intérêt, avec laquelle il est possible de reconstruire les états des sous systèmes impliqués dans la génération de la fonctionnelle de Volterra.

- La première originalité de ce mémoire réside dans l'utilisation de modèles bilinéaires homogènes (nilpotents) destinés à approcher un système non linéaire à travers la réalisation d'une série de Volterra de degré fini et à horizon infini.
- L'utilisation de l'algorithme héréditaire dans ce contexte rend la tâche d'identification plus robuste, étant donné qu'il n'y a pas de recours à l'optimisation non linéaire pour identifier les paramètres du système, contrairement à d'autres approches [Chen89b]. En particulier, aucun paramètre de réglage ni d'initialisation n'est requis dans l'algorithme. Les seuls choix laissés à l'utilisateur résident dans le choix du degré de la fonctionnelle de Volterra utilisée (linéaire, quadratique, cubique, ..., et la combinaison de ces fonctionnelles) et dans le degré de séparabilité des noyaux à travers la dimension des états intermédiaires.
- L'exploitation de la forme canonique présentée qui autorise des retours de sortie permet de réduire considérablement le nombre de paramètres à estimer, par rapport aux modèles classiquement utilisés à réponse impulsionnelle finie, tout en améliorant le conditionnement numérique du problème. On montre que la définition du système à identifier comme variable avec le temps permet de contourner la non linéarité inhérente à toute procédure d'identification de systèmes bouclés.
- L'identification des paramètres par l'algorithme héréditaire s'avère bénéfique sous plusieurs aspects. En particulier, l'optimisation dans le transitoire permet d'améliorer la précision des paramètres identifiés pour des jeux de données « courts ».

Les perspectives envisagées pour ce travail sont :

- L'extension de la forme canonique dans un premier temps pour un modèle à entrée vectorielle et sortie scalaire, et dans un deuxième temps pour une forme à entrée-sortie vectorielles.
- L'analyse de convergence de la forme canonique pour des modèles de degré supérieur à deux, sachant que le conditionnement de la matrice de corrélation à inverser peut poser un souci par rapport à la qualité de la convergence.
- Envisager une orthogonalisation des espaces de projection afin d'améliorer la robustesse numérique de l'algorithme, comme ceci a déjà été réalisé dans le contexte linéaire grâce aux structures en treillis.

Bibliographie

- [Autom95] "Special Issue on Trends in System Identification," *Automatica*, vol. 31, no. 12, 1995.
- [Bill80] Billings S. A., "Identification of Nonlinear Systems – a Survey," *Proc. IEE*, vol. 127, no.6, pp. 272-284, 1980.
- [Birk65] Birkhoff G. et MacLane S., *A Survey of Modern Algebra*, 3rd Edition, The MacMillan Company, 1965.
- [Brock76a] Brockett R. W., "Volterra Series and Geometric Control Theory," *Automatica*, vol. 12, pp. 167-176, 1976.
- [Brock76b] Brockett R. W., "Nonlinear Systems and Differential Geometry," *Proc. of the IEEE*, vol. 64, no. 1, pp. 61-72, 1976.
- [Brock76c] Brockett R. W., "Finite and Infinite Dimensional Bilinear Realizations," *Journal of The Franklin Institute*, vol. 301, no. 6, pp. 509-520, 1976.
- [Bruni74] Bruni C. et al., "Bilinear Systems: An appealing Class of Nearly Linear Systems in Theory and Applications," *IEEE Trans. on Automatic Control*, vol. AC-19, no. 4, pp. 334-348, 1974.
- [Chen89a] Chen S., Billings S. A., "Representations of Non-linear Systems: the NARMAX model," *Int. J. Control*, vol. 49, no. 3, pp. 1013-1032, 1989.
- [Chen89b] Chen S., Billings S. A., "Recursive Prediction Error Parameter Estimator for Non-linear Models," *Int. J. Control*, vol. 49, no. 2, pp. 569-594, 1989.
- [Chen99] Chen C. T., *Linear System Theory and Design*, 3rd Edition, Oxford University Press, 1999.
- [Chen00] Chen H., Maciejowski J. M., "Subspace Identification Method for Combined Deterministic-Stochastic Bilinear Systems," *Proc. IFAC Symposium on System Identification, SYSID 2000*, Santa Barbara, CA, USA, 2000.
- [Cheng01] Cheng C. H., Powers E. J., "Optimal Volterra Kernel Estimation Algorithms for a Nonlinear Communication System for PSK and QAM Inputs," *IEEE Trans. on Signal Processing*, vol. 49, no. 1, pp. 147-163, 2001.
- [Cher98] Cherry J. A., Snelgrove W. M., "On the Characterization and Reduction of Distortion in Bandpass Filters," *IEEE Trans. Circuits Sys.-I*, vol. 45, no. 5, pp. 523-537, 1998.
- [Chikte79] Chikte S. D., et al., "Optimal Filters for Bilinear Systems with Nilpotent Lie Algebras," *IEEE Trans. on Automatic Control*, vol. AC-24, no. 6, pp. 948-953, 1979.

- [Cho94] Cho Y. S., Powers E. J., "Quadratic System Identification Using High Order Spectra of I.I.D. Signals," *IEEE Trans. Signal Processing*, vol. 42, no. 5, pp. 1268-1271, 1994.
- [Cro81] Crouch P.E., "Dynamical Realizations of Finite Volterra Series," *SIAM J. Control and Optimization*, vol. 19, no. 2, pp. 177-202, 1981.
- [D'Aless74] D'Alessandro P. et al., "Realization and Structure Theory of Bilinear Dynamical Systems," *SIAM Journal of Control*, vol. 12, no. 3, pp.517-535, 1974.
- [Dai89] Dai H., Sinha N. K., "Robust Recursive Least-Squares Method with Modified Weights for Bilinear System Identification," *Proc. IEE*, vol. 136, no. 3, pp. 122-126, 1989.
- [Etchev06] Etcheverry G., et al., "Quadratic System Identification by Hereditary Approach," *2006 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'2006)*, Toulouse (France), 14-19 Mai 2006, pp.III.129-III.132
- [Evans83] Evans M. E., "Bilinear Systems with Homogeneous Input-Output Maps," *IEEE Trans. on Automatic Control*, vol. AC-28, no. 1, pp. 113-115, 1983.
- [Fav99] Favoreel W., De Moor B., Van Overschee P., "Subspace Identification of Bilinear Systmes Subject to White Inputs," *IEEE Trans. Automatic Control*, vol. 44, no. 6, 1999.
- [Fav00] Favoreel W., De Moor B., Van Overschee P., "Subspace Identification for Industrial Processes," *Journal of Process Control*, vol. 10, pp. 149-155, 2000.
- [Fliess79] Fliess M., "A Remark on the Transfer Functions and the Realization of Homogenous Continuous-Time Nonlinear Systems," *IEEE Trans. on Automatic Control*, vol. AC-24, no. 3, pp. 507-508, 1979.
- [Fliess82] Fliess M. et Normand-Cyrot D., "On the Approximation of Nonlinear Systems by Some Simple State-Space Models," *IFAC Symposium on Identification and System Parameter Estimation*, pp. 511-514, 1982.
- [Fliess83] Fliess M., et al., "An Algebraic Approach to Nonlinear Functional Expansions," *IEEE Trans. Circuits Syst.*, vol. CAS-30, no. 8, pp. 554-570, 1983.
- [Fna87] Fnaiech F., Ljung L., "Recursive Identification of Bilinear Systems," *Int. J. Control*, vol. 45, no. 2, pp. 453-470, 1987.
- [Fraz80] Frazho A. E., "A Shift Operator to Bilinear System Theory," *SIAM Journal of Control*, vol. 18, no. 6, pp. 640-658, 1980.

- [Gao89] Gao X. Y., et al., "Nonlinear IIR Adaptive Filtering Using a Bilinear Structure," *Proc. IEEE ISCAS*, pp. 1740-1743, 1989.
- [Gian01] Giannakis G. B., Serpedin E., "A Bibliography on Nonlinear System Identification," *Signal Processing*, vol. 81, pp.533-580, 2001.
- [Gleen79] Gleen E., et al., "On Transfer Functions Representations for Homogeneous Nonlinear Systems," *IEEE Trans. on Automatic Control*, vol. 24, no. 2, pp. 242-249, 1979.
- [Glen99] Glentis G. O. A., Koukoulas P., Kalouptsidis N., "Efficient Algorithms for Volterra System Identification," *IEEE Trans. Signal Processing*, vol. 47, no. 11, pp. 3042-3057, 1999.
- [Goeth05] Goethals I., et al., "Subspace Identification of Hammerstein Systems Using Least Squares Support Vector Machines," *IEEE Trans. on Automatic Control*, vol. 50, no. 10, pp. 1509-1519, 2005.
- [Gom02] Gomez J. C., Baeyens E., "Subspace Identification of Multivariable Hammerstein and Wiener Models," *IFAC 15th World Congress*, Spain, 2002.
- [Haber88] Haber R., "Parametric Identification of Nonlinear Dynamic Systems Based on Nonlinear Crosscorrelation Functions," *Proc. IEE*, vol. 135, no. 6, pp. 405-420, 1988.
- [Haber99] Haber R., Keviczky L., *Nonlinear System Identification, Input-Output Modeling Approach*, Kluwer Academic Publishers, 1999.
- [IEEE92] "Special Issue on System Identification for Robust Control Design," *IEEE Trans. Automatic Control*, vol. 37, no. 7, 1992.
- [IEEE05] "Special Issue on System Identification," *IEEE Trans. Automatic Control*, vol.50, no. 10, 2005.
- [Inag84] Inagaki M., Mochizuki H., "Bilinear System Identification by Volterra Kernels Estimation," *IEEE Trans. on Automatic Control*, vol. AC-29, no. 8, pp. 746-749, 1984.
- [Jacob79] Jacobson N., *Lie Algebras*, Dover Publications Inc., 1979.
- [Kaloup03] Kalouptsidis N., Koukoulas P., Mathews V. J., "Blind Identification of Bilinear Systems," *IEEE Trans. Signal Processing*, vol. 51, no. 2, pp. 484-499, 2003.
- [Kaloup05] Kalouptsidis N., Koukoulas P., "Blind Identification of Volterra-Hammerstein Systems," *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 2777-2787, 2005.
- [Ki Baik93] Ki Baik H. et Mathews V. J., "Adaptive Lattice Bilinear Filters," *IEEE Trans. Signal Processing*, vol. 41, no. 6, pp. 2033-2046, June 1993.

- [Kiban06] Kibangou A., Favier G., "Wiener-Hammerstein Systems Modeling Using Diagonal Volterra Kernels Coefficients," *IEEE Signal Processing Letters*, vol. 13, no. 6, June 2006.
- [Kisac01] Kisacanin B. et Agarwal G. C., *Linear Control Systems with solved problems and MATLAB examples*, Kluwer Academic/Plenum Publishers, 2001.
- [Koh85] Koh T., Powers E. J., "Second-Order Volterra Filtering and Its Application to Nonlinear Systems Identification," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, no. 6, pp.1445-1455, 1985.
- [Koren91] Korenberg M. J., Paarmann L. D., "Orthogonal Approaches to Time-Series Analysis and System Identification," *IEEE SP Magazine*, pp. 29-43, 1991.
- [Kren74] Krener A. J., "Linearization and Bilinearization of Control Systems," *Proc. 1974 Allerton Conf. on Circuit and System Theory*, pp. 834-843, 1974.
- [Kren75] Krener A. J., "Bilinear and Nonlinear Realizations of Input-Output Maps," *SIAM Journal of Control*, vol. 13, no. 4, pp.827-834, 1975.
- [Lamna94] Lamnabhi-Lagarrigue F., *Analyse des Systèmes Non Linéaires*, Hermès, 1994.
- [Lar90] Larrimore W. E., "Canonical Variate Analysis in Identification, Filtering, and Adaptive Control," *Proc. 29th IEEE Conf. on Decision and Control*, Hawai, USA, pp.596-604, 1990.
- [LeGal01] N. Le Gallou, E. Ngoya, et al., "An Improved Behavioral Modeling Technique for High Power Amplifiers with Memory," *IEEE MTT-S Digest*, vol. 2, pp. 983-986, 2001.
- [Les78] Lesiak C., Krener A. J., "The Existence and Uniqueness of Volterra Series for Nonlinear Systems," *IEEE Trans. on Automatic Control*, vol. AC-23, no. 6, pp. 1090-1095, 1978.
- [Lju83] Ljung L., Söderström T., *Theory and Practice of Recursive Identification*, The MIT Press, 1983.
- [Lju99] Ljung L., *System Identification, Theory for the User*, 2nd Edition, Prentice Hall PTR, 1999.
- [Lju01] Ljung L., "Black-Box Models from Input-Output Measurements," *IEEE Instrumentation. Meas. Tech. Conf.*, Budapest, 2001, pp. 138-146.
- [Lju05] Ljung L., Vicino A., "Guest Editorial: Special Issue on System Identification," *IEEE Trans. Automatic Control*, vol. 50, no. 10, 2005, pp.1473.
- [Ln85a] Leontaritis I. J., Billings S. A., "Input-Output Parametric Models for Nonlinear Systems. Part I: Deterministic Nonlinear Systems," *Int. J. Control*, vol. 41, no.2, pp. 303-328, 1985.

- [Ln85b] Leontaritis I. J., Billings S. A., "Input-Output Parametric Models for Nonlinear Systems. Part II: Stochastic Nonlinear Systems," *Int. J. Control*, vol. 41, no. 2, pp.329-344, 1985.
- [Ln87] Leonartis I. J., Billings S. A., "Experimental Design and Identifiability for Non-linear Systems," *Int. J. Systems Sci.*, vol. 18, no. 1, pp. 189-202, 1987.
- [Luen67] Luenberger D., "Canonical Forms for Linear Multivariable Systems," *IEEE Trans. on Automatic Control*, pp. 290-293, 1967.
- [Math91] Mathews V. J., Moon T. K., "Parameter Estimation for a Bilinear Time Series Model," *Proc. IEEE*, pp. 3513-3516, 1991.
- [Math96] Mathews V. J., "Adaptive Volterra Filters Using Orthogonal Structure," *IEEE Signal Processing Lett.*, vol. 3, no. 12, pp. 307-309, 1996.
- [Medd98] Meddeb S., Tourneret J. Y., Castanie F., "Identification of Bilinear Systems Using Bayesian Inference," *Proc. IEEE ICASSP*, pp.1609-1612, 1998.
- [Medd00] Meddeb S., Tourneret J. Y., Castanie F., "Unbiased Parameter Estimation for the Identification of Bilinear Systems," *Proc. IEEE ICASSP*, pp.176-180, 2000.
- [Mitz79] Mitzel G. E., "On Transfer Function Representations for Homogeneous Nonlinear Systems," *IEEE Trans. on Automatic Control*, vol. AC-24, no. 2, pp. 242-249, 1979.
- [Mohl73] Mohler R. R., *Bilinear Control Processes*, Mathematics in Science and Engineering, vol. 106, Academic Press, 1973.
- [Mohl80] Mohler R. R., Kolodziej W. J., "An Overview of Bilinear Systems Theory and Applications," *IEEE Trans. Syst, Man, Cybern.*, vol. SMC-10, no. 10, pp. 683-688, 1980.
- [Mona87] Monaco S. et Normand-Cyrot D., "Finite Volterra-series Realizations and Input-Output Approximations of Non-linear Discrete-time Systems," *Int. J. Control*, vol. 45, no. 5, pp. 1771-1787, 1987.
- [Mon94] Monin A. Salut G., "Exact ARMA Lattice Predictors from Autocorrelation Functions", *IEEE Trans. Signal Processing*, vol. 42, no. 4, pp. 877-886, 1994.
- [Mon96a] Monin A., Salut G., "ARMA Lattice Identification: A New Hereditary Algorithm," *IEEE Trans. on Signal Processing*, vol. 44, no. 2, pp. 360-370, 1996.
- [Mon96b] Monin A. et Salut G., "IIR Volterra Filtering with Application to Bilinear Systems," *IEEE Trans. on Signal Processing*, vol. 44, no. 9, pp. 2209-2221, 1996.
- [Mon04] Monin A., "ARMAX Identification Via Hereditary Algorithm," *IEEE Trans. on Automatic Control*, vol. 49, no. 2, pp. 233-238, 2004.

- [Ogun01] Ogunfunmi T., Chang S. L., "Second Order Adaptive Volterra System Identification Based on Discrete Nonlinear Wiener Model," *IEE Proc.-Vis Image Signal Processing*, vol. 148, no. 1, pp. 21-29, 2001.
- [Park81] Parker S. R., Perry F. A., "A Discrete ARMA Model for Nonlinear System Identification," *IEEE Trans. Circuits Syst.*, vol. CAS-28, no. 2, pp. 224-233, 1981.
- [Pearl78] Pearlman J. G., "Canonical Forms for Bilinear Input/Output Maps," *IEEE Trans. on Automatic Control*, vol. AC-23, no. 4, pp. 595-602, 1978.
- [Reut94] Reuter H., "State Space Identification of Bilinear Canonical Forms," *Proc. of the IEE Control Conf.*, pp.833-838, 1994.
- [Rugh81] Rugh W. J., *Nonlinear System Theory, The Volterra/Wiener Approach*, The Johns Hopkins University Press 1981. Web version 2002: www.ece.jhu.edu/~rugh/volterra/book.pdf
- [Sche80] Schetzen M., *The Volterra and Wiener Theories of Nonlinear Systems*, Krieger Publishing Company, 1980.
- [Shynk89] Shynk J. J., "Adaptive IIR Filtering," *IEEE ASSP Magazine*, pp. 4-21, 1989.
- [Son79] Sontag E. D., "Realization Theory of Discrete-Time Nonlinear Systems: Part I The Bounded Case," *IEEE Trans. on Circuits Systems*, vol. CAS-26, no. 4, pp. 342-356, 1979.
- [Sulei06] Suleiman W., Etcheverry G., Monin A., "Nouvelle approche pour l'identification des systèmes non linéaires," Conférence Internationale Francophone d'Automatique - CIFA 2006, Bordeaux (France).
- [Tseng95] Tseng C. H., Powers E. J., "Identification of Cubic Systems Using Higher Order Moments of I.I.D. Signals," *IEEE Trans. Signal Processing*, vol. 43, no. 7, pp. 1733-1735, 1995.
- [Tsou01] Tsoukas V., Koukoulas P., Kalouptsidis N., "Identification of Input-Output Bilinear Systems Using Cumulants," *IEEE Trans. Signal Processing*, vol. 49, no. 11, pp. 2753-2761, 2001.
- [Van94] Van Overschee P., De Moor B., "N4SID: Subspace Algorithms for the Identification of Combined Deterministic-Stochastic Systems," *Automatica*, vol. 30, no. 1, pp. 75-93, 1994.
- [Van95] Van Overschee P., De Moor B., "A Unifying Theorem for Three Subspace System Identification Algorithms," *Automatica*, vol. 31, no. 12, pp. 1853-1864, 1995.
- [Van96] Van Overschee P., De Moor B., *Subspace Identification for Linear Systems*, Kluwer Academic Publishers, 1996.

- [Verd99] Verdult V., Verhaegen M., Chou C. T., "Identification of MIMO Bilinear State Space Models Using Separable Least Squares," *Proc. American Control Conf.*, pp. 838-842, 1999.
- [Verd01] Verdult V., Verhaegen M., "Identification of Multivariable Bilinear State Space Systems Based on Subspace Techniques and Separable Least Squares Optimization," *International Journal of Control*, vol. 74, no.18, pp. 1824-1836, 2001.
- [Verd02] Verdult V., *Nonlinear System Identification: A State-Space Approach*, PhD. Thesis, University of Twente, 2002.
- [Verh94] Verhaegen M. "Identification of the Deterministic Part of MIMO State Space Models given in Innovations Form from Input-Output Data," *Automatica*, vol.30, no. 1, pp. 61-74, 1994.
- [Vib93] Viberg M., et al., "Performance of Subspace Based State Space System Identification Methods," *Proc. 12th IFAC World Congress*, Sydney, Australia, vol. 7, pp. 369-372, 1993.
- [Vib95] Viberg M., "Subspace-based Methods for the Identification of Linear Time-invariant Systems," *Automatica*, vol. 31, no. 12, pp. 1835-1851, 1995.
- [Vuole01] Voulevi J., "*Analysis, Measurement and Cancellation of the Bandwidth and Amplitude Dependence of Intermodulation Distortion in RF Power Amplifiers*," PhD. Thesis, University of Oulu, Finland, 2001.
- [Wang03] Wang T., Brazil T. J., "Volterra-Mapping-Based Behavioral Modeling of Nonlinear Circuits and Systems for High Frequencies," *IEEE Trans. on Microwave Theory and Techniques*, vol. 51, no. 5, pp. 1433-1440, 2003.
- [West96] Westwick D., Verhaegen M., "Identifying MIMO Wiener Systems Using Subspace Model Identification Methods," *Signal Processing*, vol. 52, pp. 235-258, 1996.
- [Zhang98] Zhang Q., et al, "Factors Affecting Volterra Kernel Estimation: Emphasis on Lung Tissue Viscoelasticity," *Annals of Biomedical Engineering*, vol. 26, pp. 103-116, 1998.
- [Zhao98] Zhao X., Marmarelis V. Z., "Nonlinear Parametric Models from Volterra Kernels Measurements," *Mathl. Comput. Modelling*, vol. 27, no. 5, pp. 37-43, 1998.
- [Zhu99] Zhu Z., Leung H., "Adaptive Identification of Bilinear Systems," *Proc. IEEE ICASSP*, pp.1289-1292, 1999.

Resumé en français

Les systèmes bilinéaires jouent un rôle central dans la représentation des systèmes non-linéaires en tant qu'approximant des systèmes non-linéaires analytiques généraux. Une classe particulière de ces systèmes bilinéaires, dits homogènes, permet de représenter la fonctionnelle entrée-sortie comme une série de Volterra à horizon infini mais de degré fini. L'algorithme d'identification héréditaire jusqu'alors limité aux systèmes linéaires est ici étendu à l'identification de ces systèmes en utilisant une structure canonique de base pour l'approximant de degré deux (quadratique). Une forme NARX multidimensionnelle (Nonlinear AutoRegressive eXogenous input) est exploitée ici pour autoriser l'identification du système par la méthode héréditaire. On montre, sur un exemple académique quadratique, la supériorité de cette approche comparée aux algorithmes classiques.

Mots clés - Séries de Volterra, Systèmes Biliéaires Homogènes, Identification Héréditaire.

Resumé en anglais

Quadratic systems are the first kind of nonlinear systems in which we are interested in order to study polynomial nonlinearities. They can be approximated by bilinear models with nilpotent structure that approximate certain nonlinearities and generate finite degree Volterra series. The hereditary identification algorithm limited until now to linear systems is extended here for identification of nonlinear systems by implementing a canonical structure to the approximant of degree two (quadratic). A NARX (Nonlinear Autoregressive eXogenous input) multidimensional expression is employed in order to perform identification by hereditary computation.

Key words- Volterra Series, Homogenous Bilinear Systems, Hereditary Identification.