

Université de Montréal

TESTS D'INDÉPENDANCE DE DEUX SÉRIES  
MULTIVARIÉES AUTORÉGRESSIVES  
D'ORDRE INFINI

par

CHAFIK BOUHADDIOUI

Département de mathématiques et de statistique  
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures  
en vue de l'obtention du grade de  
Philosophiæ Doctor (Ph.D.)  
en Statistique

août 2001



QA

3

U54

2002

v. 008

QMS REV 5.1

Université de Montréal

Faculté des études supérieures

Cette thèse intitulée

TESTS D'INDÉPENDANCE DE DEUX SÉRIES  
MULTIVARIÉES AUTORÉGRESSIVES  
D'ORDRE INFINI

présentée par

CHAFIK BOUHADDIOUI

a été évaluée par un jury composé des personnes suivantes :

*Robert Cléroux*

---

(président-rapporteur)

*Roch Roy*

---

(directeur de recherche)

*Jean-Marie Dufour*

---

(membre du jury)

*A.Ian McLeod*

---

(examineur externe)

*Jacques Ferland*

---

(représentant du doyen)

Thèse acceptée le :

17 janvier 2002

## SOMMAIRE

---

Dans plusieurs situations, nous voulons vérifier l'existence de relations entre des séries chronologiques multivariées. Haugh (1976) a proposé un test asymptotique afin de vérifier l'hypothèse de non corrélation (ou d'indépendance dans le cas gaussien) entre deux séries univariées ARMA. El Himdi et Roy (1997) ont généralisé cette approche au cas de deux séries multivariées ARMA stationnaires et inversibles.

Hong (1996) a proposé une approche semi-paramétrique pour tester l'indépendance de deux séries univariées autorégressives d'ordre infini ( $\text{VAR}(\infty)$ ). Son approche consiste à filtrer chaque série à l'aide d'une autorégression d'ordre fini et la statistique de test proposée fait intervenir les corrélations résiduelles croisées à tous les délais. Ces derniers sont pondérés à l'aide d'une fonction de noyau et d'un point de troncature. Sous des conditions générales de régularité, il montre que sous l'hypothèse nulle, la statistique de test, centrée et réduite de façon appropriée, suit asymptotiquement une loi  $\mathcal{N}(0, 1)$ . D'une certaine façon, la statistique portemanteau de Haugh peut être vue comme un cas particulier en utilisant le noyau uniforme tronqué.

Dans cette thèse, nous généralisons l'approche d'El Himdi et Roy (1997) au cas de deux processus multivariés  $\text{VAR}(\infty)$  stationnaires et inversibles. Les statistiques de tests sont basées sur les matrices de corrélations croisées résiduelles. Sous

l'hypothèse nulle de non corrélation et en utilisant des résultats de Lewis et Reinsel (1985) sur l'approximation d'un processus VAR( $\infty$ ) par une autorégression finie, il est montré qu'un vecteur quelconque de corrélations croisées résiduelles suit la même loi asymptotique que le vecteur correspondant des corrélations croisées entre les deux séries d'innovations. Il découle du résultat précédent que les statistiques de test considérées suivent asymptotiquement une loi khi-deux. Deux statistiques de test sont décrites. La première est basée sur la matrice des corrélations croisées résiduelles à un délai particulier alors que la deuxième fait intervenir les corrélations croisées à plusieurs délais, disons de  $-M$  à  $M$  où  $M$  est un entier positif indépendant de la taille échantionnale  $N$ .

En utilisant à nouveau les résultats de Lewis et Reinsel (1985), nous montrons que l'analogie de la statistique de Hong (1996) pour des séries multivariées suit aussi une loi normale sous l'hypothèse nulle. La statistique de test proposée généralise d'une certaine façon celle d'El Himdi et Roy (1997) pour le cas de deux processus VAR( $\infty$ ). Pour une certaine classe de noyaux, la puissance locale et la puissance globale asymptotiques sont étudiées de façon analytique et le noyau optimal est identifié. Le contrôle du niveau et la puissance sont aussi étudiés par simulation en échantillons finis. Une application à des données réelles est aussi présentée.

## REMERCIEMENTS

---

Je voudrais exprimer mes remerciements à mon directeur de recherche, le professeur Roch Roy, pour l'intérêt, les conseils judicieux et ses critiques éclairés. Ses commentaires et suggestions ont été d'une grande utilité pour l'avancement de ce travail.

Cette recherche a bénéficié d'une bourse de la Faculté des études supérieures, d'un appui financier du Fonds FCAR, du Conseil de Recherche en Science Naturelles et en Génie du Canada (CRNSG) et du réseau de centre d'excellence sur les mathématiques des technologies de l'information et des systèmes complexes (MATISC), par le biais de subventions à mon directeur de recherche.

Je voudrais aussi exprimer mes remerciements les plus vifs à mes parents qui m'ont soutenu et encouragé durant toutes mes études. Aussi, mon frère et mes soeurs pour leurs encouragements. Je remercie ma fiancée Mélanie pour son soutien durant les moments les plus difficiles ainsi que tous mes amis, spécialement Mr M. Halmi et Dr. A. Merzouki. Je tiens également à remercier mon ami P.Duchesne pour les discussions statistiques durant notre scolarité et de tous les bons moments que nous avons passé au département. Finalement, je remercie le support informatique du département de mathématiques et de statistique pour leur soutien technique, et spécialement Mr. N. Beauchemin.

# Table des matières

---

<b>Sommaire</b> .....	iii
<b>remerciements</b> .....	v
<b>Liste des tableaux</b> .....	ix
<b>Liste des figures</b> .....	xi
<b>INTRODUCTION</b> .....	1
<b>Chapitre 1. PROCESSUS MULTIVARIÉS STATIONNAIRES..</b>	<b>7</b>
1.1. Processus stationnaires .....	8
1.2. Processus linéaires .....	13
1.2.1. Processus moyennes mobiles d'ordre infini .....	13
1.2.2. Processus autorégressifs d'ordre infini .....	14
1.2.3. Processus ARMA multivariés .....	15
1.3. Approximation d'un processus autorégressif d'ordre infini .....	16
<b>Chapitre 2. TESTS D'INDÉPENDANCE BASÉS SUR UN SOUS-ENSEMBLE DE DÉLAIS</b> .....	<b>20</b>
2.1. Définitions et généralités .....	22
2.2. Distribution asymptotique de $\mathbf{r}_{\hat{\mathbf{a}}}^{(12)}$ pour deux processus ARMA multivariés .....	23

2.3.	Généralisation au cas de deux processus multivariés $VAR(\infty)$ .....	26
2.3.1.	Procédures des tests d'indépendance .....	35
<b>Chapitre 3.</b>	<b>TESTS D'INDÉPENDANCE BASÉS SUR TOUS</b>	
	<b>LES DÉLAIS.....</b>	<b>38</b>
3.1.	Approche de Hong pour deux processus univariés $AR(\infty)$ .....	39
3.2.	Généralisation au cas de deux processus multivariés $VAR(\infty)$ .....	42
3.2.1.	Définitions et généralités.....	42
3.2.2.	Loi asymptotique sous l'hypothèse nulle.....	46
<b>Chapitre 4.</b>	<b>PUISSANCE ASYMPTOTIQUE.....</b>	<b>78</b>
4.1.	Introduction.....	78
4.2.	Puissance asymptotique locale.....	79
4.3.	Puissance asymptotique globale.....	88
<b>Chapitre 5.</b>	<b>ÉTUDE PAR SIMULATION.....</b>	<b>96</b>
5.1.	Méthodologie de la simulation.....	97
5.1.1.	Choix des modèles de processus.....	97
5.1.2.	Description de la simulation.....	97
5.1.3.	Procédure des tests.....	98
5.2.	Analyse des résultats de simulation.....	102
5.2.1.	Discussion de l'étude de niveau.....	102
5.2.2.	Discussion de l'étude de puissance.....	103
<b>Chapitre 6.</b>	<b>APPLICATION.....</b>	<b>116</b>
6.1.	Analyse et modélisation des séries.....	119

6.2. Tests de non corrélation.....	127
<b>CONCLUSION.....</b>	<b>131</b>
<b>Bibliographie.....</b>	<b>134</b>

## Liste des tableaux

---

5.1	Zones de non signification à 5% et 1% du pourcentage de rejets basé sur 5000 et 2000 réalisations respectivement pour chacun des trois niveaux nominaux. ....	101
5.2	Modèles et paramètres des processus utilisés dans l'étude par simulation	105
5.3	Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle AR(1) avec $\mathcal{Q}_N$ et $\mathcal{Q}_N^*$ basés sur 5000 réalisations pour $N = 100$ et 2000 réalisations pour $N = 200$ en utilisant les noyaux DAN, PAR, BP et TR. ....	106
5.4	Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle AR(1) avec $\mathcal{Q}_N, \mathcal{Q}_N^*$ et $\mathcal{Q}_{\hat{\alpha}, M}^*$ basés sur 5000 réalisations pour $N = 100$ et 2000 réalisations pour $N = 200$ en utilisant les noyaux TH et BAR. ....	107
5.5	Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle MA(1) avec $\mathcal{Q}_N$ et $\mathcal{Q}_N^*$ basés sur 5000 réalisations pour $N = 100$ et 2000 réalisations pour $N = 200$ en utilisant les noyaux DAN, PAR, BP et TR. ....	108
5.6	Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle MA(1) avec $\mathcal{Q}_N, \mathcal{Q}_N^*$ et $\mathcal{Q}_{\hat{\alpha}, M}^*$ basés sur 5000 réalisations pour $N = 100$ et 2000 réalisations pour $N = 200$ en utilisant les noyaux TH et BAR. ....	109

5.7	Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle $AR(1)_\delta$ , $\delta = 1$ avec $\mathcal{Q}_N$ et $\mathcal{Q}_N^*$ basés sur 2000 réalisations en utilisant les noyaux DAN, PAR, BP et TR. ....	110
5.8	Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle $AR(1)_\delta$ , $\delta = 1$ avec $\mathcal{Q}_N$ , $\mathcal{Q}_N^*$ et $Q_{\hat{a},M}^*$ basés sur 2000 réalisations en utilisant les noyaux TH et BAR. ....	111
5.9	Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle $AR(1)_\delta$ , $\delta = 1.5$ avec $\mathcal{Q}_N$ et $\mathcal{Q}_N^*$ basés sur 2000 réalisations en utilisant les noyaux DAN, PAR, BP et TR. ....	112
5.10	Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle $AR(1)_\delta$ , $\delta = 1.5$ avec $\mathcal{Q}_N$ , $\mathcal{Q}_N^*$ et $Q_{\hat{a},M}^*$ basés sur 2000 réalisations en utilisant les noyaux TH et BAR. ....	113
5.11	Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle $AR(1)_\delta$ , $\delta = 2$ avec $\mathcal{Q}_N$ et $\mathcal{Q}_N^*$ basés sur 2000 réalisations en utilisant les noyaux DAN, PAR, BP et TR. ....	114
5.12	Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle $AR(1)_\delta$ , $\delta = 2$ avec $\mathcal{Q}_N$ , $\mathcal{Q}_N^*$ et $Q_{\hat{a},M}^*$ basés sur 2000 réalisations en utilisant les noyaux TH et BAR. ....	115
6.1	Valeurs de la statistique $Q_{\hat{a},M}^*$ et niveaux empiriques en fonction de $M$ pour la série canadienne et américaine différenciées. ....	130
6.2	Valeur des statistiques $\mathcal{Q}_N$ et $\mathcal{Q}_N^*$ avec $M = 4, 7, 11$ pour la série canadienne et américaine différenciées. ....	130

## Liste des figures

---

- 5.1 Comparaison des différents noyaux utilisés ..... 99
- 6.1 Séries d'indicateurs économiques canadiens et américains..... 118
- 6.2 Composantes de la série résiduelle correspondant à la série canadienne différenciée..... 122
- 6.3 Autocorrélations et corrélations croisées résiduelles de la série canadienne différenciée..... 123
- 6.4 Composantes de la série résiduelle correspondant à la série américaine différenciée..... 125
- 6.5 Autocorrélations et corrélations croisées résiduelles de la série américaine différenciée..... 126
- 6.6 Valeurs de la statistique  $Q_{\hat{a}}^{(12)}(j)^*$  à différents délais pour les séries canadienne et américaine différenciées. La ligne horizontale pointillée représente la valeur critique marginale au niveau  $\alpha = 0.05$ . Les autres lignes représentent les valeurs critiques des tests simultanés aux délais  $|j| \leq M$  au niveau global  $\alpha = 0.05$  pour  $M = 1(-o-)$  et  $M = 3(-+-)$ ... 129

## INTRODUCTION

---

Ces dernières années, plusieurs recherches ont été réalisées dans le domaine des séries chronologiques. Particulièrement, l'étude de relations et de liaisons qui peuvent exister entre deux séries chronologiques multivariées a intéressé plusieurs chercheurs dans différents disciplines, surtout en économétrie, où elle permet de répondre à des questions importantes sur la compréhension des mécanismes économiques sous-jacents. L'étude de liaisons entre deux séries chronologiques multivariées est souvent la première étape à l'identification des modèles régissant le système à l'étude. D'où la nécessité d'avoir des procédures simples à appliquer et à interpréter afin de vérifier l'indépendance entre deux séries chronologiques multivariées. Dans le cas univarié, Haugh et Box (1977) ont utilisé une procédure de filtrage par des modèles ARMA dans le but d'identifier une liaison de type causal entre deux séries chronologiques.

Pour étudier l'indépendance de deux séries chronologiques univariées, Haugh (1976) a proposé une approche basée sur les corrélations résiduelles. Notons par  $\{X_t^{(1)}\}$  et  $\{X_t^{(2)}\}$  les deux séries d'intérêt. Dans l'approche de Haugh, elles sont préalablement filtrées à l'aide de modèles du type ARMA stationnaires et inversibles. La statistique de test est alors basée sur les corrélations croisées entre les deux séries résiduelles résultantes  $\{\hat{a}_t^{(1)}\}$  et  $\{\hat{a}_t^{(2)}\}$ , notées  $r_{\hat{a}}^{(12)}(j)$  et définies par

la relation

$$r_{\hat{a}}^{(12)}(j) = \frac{\sum_{t=j+1}^N \hat{a}_t^{(1)} \hat{a}_{t-j}^{(2)}}{\left\{ \sum_{t=1}^N \hat{a}_t^{(1)2} \sum_{t=1}^N \hat{a}_t^{(2)2} \right\}^{1/2}}, \quad |j| \leq N-1,$$

où  $N$  représente la longueur des séries. Sous l'hypothèse nulle d'indépendance entre les deux séries  $\{X_t^{(1)}\}$  et  $\{X_t^{(2)}\}$ , Haugh (1976) montre que pour toute suite quelconque de délais distincts  $j_1, \dots, j_m$ , qui ne dépendent pas de  $N$ , le vecteur de corrélations croisées résiduelles  $\sqrt{N}(r_{\hat{a}}^{(12)}(j_1), \dots, r_{\hat{a}}^{(12)}(j_m)) = \sqrt{N}\mathbf{r}_{\hat{a}}^{(12)}$ , suit asymptotiquement une distribution multivariée normale  $\mathcal{N}(\mathbf{0}, \mathbf{I})$ . Ce qui a conduit à la définition de la statistique de type portemanteau

$$Q_{\hat{a}, M}^{(12)} = N \sum_{j=-M}^M r_{\hat{a}}^{(12)}(j)^2,$$

pour tester l'hypothèse nulle d'indépendance contre l'hypothèse alternative de corrélation sérielle à au moins un délai  $j$ ,  $|j| \leq M \leq N-1$ , où  $M$  est fixe par rapport à  $N$ . Cette statistique suit asymptotiquement sous l'hypothèse nulle une loi khi-deux à  $2M+1$  degrés de liberté et l'hypothèse nulle est rejetée pour les grandes valeurs.

Hong (1996) a proposé une approche semi-paramétrique pour tester l'indépendance de deux séries univariées autorégressives d'ordre infini (AR( $\infty$ )). Notons en passant que sous certaines conditions, la classe des processus AR( $\infty$ ) englobent la classe des processus autorégressifs-moyennes mobiles (ARMA). Son approche consiste à filtrer chaque série à l'aide d'une autorégression d'ordre fini, et dont l'ordre croît de façon adéquate avec la longueur de la série, voir Berk (1974). La statistique de test  $Q_N$  proposée fait intervenir les corrélations croisées à tous les

délais, et elle prend la forme suivante

$$\mathcal{Q}_N = \frac{N \sum_{j=1-N}^{N-1} k^2(j/M) r_{\hat{\mathbf{a}}}^{(12)}(j)^2 - S_N(k)}{\{2D_N(k)\}^{1/2}},$$

où  $k(\cdot)$  est un noyau qui vérifie quelques hypothèses de régularité et  $M$  est un paramètre de lissage tel que  $M = M(N) \rightarrow \infty$  et  $M/N \rightarrow 0$  quand  $N \rightarrow \infty$ . Les deux quantités  $S_N(k)$  et  $D_N(k)$  sont définies au chapitre 3 par les relations (3.1.8) et (3.1.9) et sont reliées à la moyenne et à la variance asymptotique de la somme pondérée des corrélations croisées résiduelles au carré apparaissant dans la définition de  $\mathcal{Q}_N$ . En utilisant le noyau uniforme tronqué, nous retrouvons une version modifiée de la statistique de Haugh. Sous des conditions générales de régularité, Hong (1996) a montré que sous l'hypothèse nulle d'indépendance, la statistique de test  $\mathcal{Q}_N$  suit asymptotiquement une loi  $\mathcal{N}(0, 1)$  et le test rejette pour les grandes valeurs de la statistique.

El Himdi et Roy (1997) ont généralisé l'approche de Haugh dans le cas de processus multivariés stationnaires. Ainsi, en considérant deux processus ARMA vectoriels, ils déduisent de Roy (1989) la distribution asymptotique du vecteur  $\mathbf{r}_{\hat{\mathbf{a}}}^{(12)}$  des corrélations croisées entre les deux processus innovations  $\{\mathbf{a}_t^{(1)}\}$  et  $\{\mathbf{a}_t^{(2)}\}$  associés respectivement aux deux processus ARMA multivariés  $\{\mathbf{X}_t^{(1)}\}$  et  $\{\mathbf{X}_t^{(2)}\}$  supposés indépendants ou non corrélés. Pour déduire un test d'indépendance entre les deux séries, ils ont montré que le vecteur  $\mathbf{r}_{\hat{\mathbf{a}}}^{(12)}$  des corrélations croisées résiduelles admet la même distribution asymptotique que le vecteur  $\mathbf{r}_{\hat{\mathbf{a}}}^{(12)}$ . Et ainsi, ils ont obtenu une généralisation au cas multivarié de la statistique portemanteau  $\mathcal{Q}_{\hat{\mathbf{a}}, M}$ , ainsi qu'une version modifiée notée  $\mathcal{Q}_{\hat{\mathbf{a}}, M}^*$ .

L'objet de ce travail est de proposer deux approches différentes pour tester l'indépendance de deux séries multivariées stationnaires. La première est une généralisation du test d'El Himdi et Roy (1997) au cas de processus autorégressifs d'ordre infini et ne fait intervenir qu'un nombre fixé de délais. La deuxième est une généralisation de l'approche de Hong (1996) au cas multivarié et fait intervenir tous les délais.

Le premier chapitre est consacré à la présentation de quelques notations, définitions et résultats concernant les séries chronologiques multivariées qui seront utilisés dans les autres chapitres. Nous rappelons aussi la définition et certaines propriétés des processus multivariés autorégressifs d'ordre infini dont des résultats de Lewis et Reinsel (1985) sur l'approximation d'un processus autorégressif d'ordre infini par un processus d'ordre fini. Ces résultats constituent une généralisation au cas multivarié de résultats univariés de Berk (1974) qui ont été utilisés par Hong (1996).

Au chapitre 2, nous généralisons l'approche d'El Himdi et Roy (1997) au cas de deux processus multivariés autorégressifs d'ordre infini VAR( $\infty$ ). Notons  $\{\hat{\mathbf{a}}_t^{(1)}\}$  et  $\{\hat{\mathbf{a}}_t^{(2)}\}$  les deux séries résiduelles, issues de l'approximation d'un processus VAR( $\infty$ ) par un processus autorégressif d'ordre fini. La matrice  $\mathbf{R}_{\hat{\mathbf{a}}}^{(12)}(j_1)$  désigne la matrice de corrélation croisée de délai  $j$  entre les deux séries  $\{\hat{\mathbf{a}}_t^{(1)}\}$  et  $\{\hat{\mathbf{a}}_t^{(2)}\}$ . Le vecteur des corrélations croisées aux délais  $j_1, \dots, j_m$  est alors défini par  $\mathbf{r}_{\hat{\mathbf{a}}}^{(12)} = (\text{vec}(\mathbf{R}_{\hat{\mathbf{a}}}^{(12)}(j_1)^T), \dots, \text{vec}(\mathbf{R}_{\hat{\mathbf{a}}}^{(12)}(j_m)^T))^T$  où  $\text{vec}(A)$  est l'opérateur usuel qui transforme la matrice  $A$  en un vecteur colonne. Nous montrons dans la section 2.1, qu'un vecteur quelconque de corrélations croisées  $\mathbf{r}_{\hat{\mathbf{a}}}^{(12)}$  possède la même

distribution asymptotique que celle du vecteur correspondant de corrélations croisées entre les deux séries d'innovations  $\mathbf{r}_a^{(12)}$ . L'ordre fini de cette autorégression dépend de la taille échantillonnale  $N$  mais croît moins rapidement que  $N$  vers l'infini. La distribution de  $\mathbf{r}_a^{(12)}$  nous permet de déduire un test d'indépendance à un délai fixe  $j$ , basée sur une statistique notée  $Q_a^{(12)}(j)^*$ . Aussi, nous généralisons dans ce cas la statistique portemanteau  $Q_{a,M}^{(12)*}$ .

Au chapitre 3, nous proposons une généralisation au cas multivarié de la statistique de Hong (1996). En utilisant les résultats de Lewis et Reinsel (1985) sur l'approximation d'un processus VAR( $\infty$ ) par une autorégression multivariée d'ordre fini, nous montrons que l'analogue de la statistique de Hong (1996) pour des séries multivariées, notée  $Q_N$ , suit aussi une loi  $\mathcal{N}(0, 1)$  sous l'hypothèse nulle d'indépendance. Le test est unilatéral et rejette pour les grandes valeurs. La démonstration de ce résultat est longue et repose sur plusieurs lemmes techniques dont les preuves sont présentées en annexe au chapitre 3.

Au chapitre 4, nous traitons la question de la puissance asymptotique du test de Hong généralisé. Deux méthodes de comparaisons sont utilisées. L'une due à Pitman (1979) pour la puissance locale et l'autre due à Bahadur (1960) pour la puissance globale. Dans le cas de la puissance locale, nous utilisons l'approche de Pitman. Ainsi, sous une classe d'alternatives locales, nous présentons la distribution asymptotique de la statistique du test. Et sous une classe spécifique de noyaux, nous trouvons le noyau optimal, c'est à dire qui maximise la puissance asymptotique locale. Dans le cas de la puissance asymptotique globale, l'approche de Bahadur est utilisée. On montre la convergence de la statistique  $Q_N$  sous une hypothèse générale de dépendance des deux processus. La pente asymptotique de

Bahadur est aussi obtenue. Dans les deux approches, l'efficacité relative asymptotique d'un noyau par rapport à un autre est calculée.

Au chapitre 5, nous rapportons les résultats d'une expérience de simulation afin de vérifier si le niveau des tests est bien contrôlé aux niveaux asymptotiques nominaux usuels. Aussi, nous comparons la puissance de  $Q_N$  pour différents noyaux et  $Q_N$  est aussi comparée au test basé sur la statistique portemanteau  $Q_{\hat{a},M}^{(12)*}$  d'El Himdi et Roy (1997).

Finalement, au chapitre 6, nous illustrons l'applicabilité des nouveaux tests proposés en utilisant des données réelles provenant de Racette et Raynauld (1992). Il s'agit de données trimestrielles portant sur des indicateurs économiques canadiens et américains et on s'intéresse aux relations existant entre les deux économies.

# Chapitre 1

---

## PROCESSUS MULTIVARIÉS STATIONNAIRES

Souvent, nous étudions des modèles qui décrivent des relations entre des variables multivariées qui dépendent du temps. De tels modèles multivariés surgissent quand on observe simultanément plusieurs séries chronologiques au lieu d'en observer une seule, comme dans le cas de l'analyse des séries chronologiques univariées. Les séries chronologiques multivariées sont d'intérêt considérable pour une variété de disciplines telles qu'en ingénierie, sciences économiques et finance. Dans l'étude des processus multivariés, un cadre est nécessaire pour décrire, non seulement les différentes propriétés de chaque composante du processus individuellement, mais également des liaisons pouvant exister entre les diverses composantes. Le but d'analyser et modéliser une série multivariée est de comprendre les relations dynamiques, par rapport au temps, entre les différentes séries et améliorer la précision des prévisions pour chaque série en utilisant non seulement l'information contenue dans le passé de la série considérée mais aussi l'information fournie par l'historique des autres séries.

Dans ce chapitre, nous introduisons les différentes définitions et notations qui seront utilisées tout au long de ce travail. Nous rappelons quelques définitions et

résultats importants sur les processus linéaires. Nous décrivons aussi l'approximation d'un processus linéaire autorégressif d'ordre infini par un processus autorégressif d'ordre fini et nous énonçons les principaux résultats asymptotiques correspondants.

### 1.1. PROCESSUS STATIONNAIRES

Considérons le processus  $\mathbf{X} = \{\mathbf{X}_t, t \in \mathbb{Z}\}$  de dimension  $m$ , stationnaire au second ordre, avec  $\mathbf{X}_t = (X_{1,t}, \dots, X_{m,t})^T$ , et  $\mu_{\mathbf{X}} = \mathbb{E}(\mathbf{X}_t)$  le vecteur moyenne du processus  $\mathbf{X}$ , qui est constant pour tout  $t \in \mathbb{Z}$ .

La fonction matricielle d'autocovariance  $\{\Gamma_{\mathbf{X}}(j), j \in \mathbb{Z}\}$  est définie par

$$\Gamma_{\mathbf{X}}(j) = \mathbb{E}(\mathbf{X}_t - \mu_{\mathbf{X}})(\mathbf{X}_{t-j} - \mu_{\mathbf{X}})^T; \quad t, j \in \mathbb{Z}.$$

Soit  $\mathbf{A}(j) = (A_{ls}(j))_{n \times n}$  une suite matricielle. On dit que la suite  $\mathbf{A}(j)$  est absolument sommable si elle vérifie la condition

$$\sum_{j=-\infty}^{\infty} |A_{ls}(j)| < \infty, \quad l, s = 1, \dots, m, \quad j \in \mathbb{Z}. \quad (1.1.1)$$

La suite  $\mathbf{A}(j)$  est dite de carré sommable si

$$\sum_{j=-\infty}^{\infty} \|\mathbf{A}(j)\|_2^2 < \infty, \quad j \in \mathbb{Z}, \quad (1.1.2)$$

où  $\|\cdot\|_2$  représente la norme euclidienne définie par  $\|\mathbf{A}\|_2^2 = \text{tr}(\mathbf{A}^T \mathbf{A})$  et  $\text{tr}(\mathbf{A})$  désigne l'opérateur matriciel trace de  $\mathbf{A}$ .

La suite matricielle  $\Gamma_{\mathbf{X}}(j) = (\gamma_{ls}(j))_{m \times m}$  est supposée absolument sommable.

De même, la matrice d'autocorrélation  $\rho_{\mathbf{X}}(j) = (\rho_{uv}(j))_{m \times m}$  du processus  $\mathbf{X}$  au délai  $j$  est définie par

$$\rho_{uv}(j) = \frac{\gamma_{uv}(j)}{\{\gamma_{uu}(0)\gamma_{vv}(0)\}^{1/2}}, \quad 1 \leq u, v \leq m, \quad j \in \mathbb{Z}. \quad (1.1.3)$$

Si on définit la matrice diagonale  $\mathbf{D}\{b_i\}$  de dimension  $m$  dont les éléments diagonaux sont  $b_1, \dots, b_m$ , on peut écrire la matrice  $\boldsymbol{\rho}_{\mathbf{X}}(j)$  sous la forme matricielle suivante:

$$\boldsymbol{\rho}_{\mathbf{X}}(j) = \mathbf{D}\{\gamma_{ii}(0)\}^{-1/2} \boldsymbol{\Gamma}_{\mathbf{X}}(j) \mathbf{D}\{\gamma_{ii}(0)\}^{-1/2}, \quad j \in \mathbb{Z}. \quad (1.1.4)$$

Notons que pour tout délai  $j$  entier, on a :

$$\boldsymbol{\rho}_{\mathbf{X}}(j) = \boldsymbol{\rho}_{\mathbf{X}}(-j)^T \text{ et } \boldsymbol{\Gamma}_{\mathbf{X}}(j) = \boldsymbol{\Gamma}_{\mathbf{X}}(-j)^T. \quad (1.1.5)$$

La densité spectrale du processus  $\mathbf{X}$ ,  $\mathbf{f}_{\mathbf{X}}(\lambda) = (f_{uv}(\lambda))_{m \times m}$ , en termes de la fonction d'autocovariance, est définie par

$$\mathbf{f}_{\mathbf{X}}(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \boldsymbol{\Gamma}_{\mathbf{X}}(j) e^{-i\lambda j}, \quad -\pi \leq \lambda \leq \pi,$$

et les éléments  $f_{uv}(\lambda)$  s'écrivent

$$f_{uv}(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_{uv}(j) e^{-i\lambda j}, \quad -\pi \leq \lambda \leq \pi. \quad (1.1.6)$$

La condition de sommabilité absolue de la suite  $\boldsymbol{\Gamma}_{\mathbf{X}}(j)$  assure l'existence de la densité spectrale. Pour toute fréquence  $\lambda$ , la densité spectrale  $\mathbf{f}_{\mathbf{X}}(\lambda)$  est une matrice hermitienne et définie non négative, voir Fuller (1996, p. 170). De même, on peut exprimer la matrice d'autocovariance à l'aide de la densité spectrale comme suit:

$$\boldsymbol{\Gamma}_{\mathbf{X}}(j) = \int_{-\pi}^{\pi} e^{ij\lambda} \mathbf{f}_{\mathbf{X}}(\lambda) d\lambda, \quad j \in \mathbb{Z}.$$

La fonction de cohérence complexe du processus  $\mathbf{X}$  est définie par la matrice  $\mathbf{S}_{\mathbf{X}}(\lambda) = (\mathcal{S}_{uv}(\lambda))_{m \times m}$  dont les éléments sont donnés par

$$\mathcal{S}_{uv}(\lambda) = \frac{f_{uv}(\lambda)}{\sqrt{f_{uu}(\lambda)f_{vv}(\lambda)}}, \quad -\pi \leq \lambda \leq \pi. \quad (1.1.7)$$

La fonction de cohérence du processus  $\mathbf{X}$  est défini par la matrice  $|\mathcal{S}_{\mathbf{X}}(\lambda)|$  dont l'élément  $(u, v)$  est le module de la cohérence complexe, i.e.

$$|\mathcal{S}_{uv}(\lambda)| = \frac{|f_{uv}(\lambda)|}{\sqrt{f_{uu}(\lambda)f_{vv}(\lambda)}}, \quad -\pi \leq \lambda \leq \pi, \quad (1.1.8)$$

où  $|\cdot|$  représente le module d'un nombre complexe.

La densité spectrale peut être nulle en quelques fréquences, et dans cette situation on assigne à la fonction de cohérence la valeur nulle pour ces fréquences. Notons que  $\mathcal{S}_{uv}(\lambda)$  peut être interprétée comme une mesure de corrélation entre  $X_{t,u}$  et  $X_{t,v}$  à la fréquence  $\lambda$ . D'après Priestley (1981, p. 661), la fonction de cohérence entre deux processus stationnaires dont la densité spectrale existe en toutes fréquences, reste invariante sous une transformation linéaire.

Supposons maintenant que le processus  $\mathbf{X}$  est séparé en deux sous processus  $\mathbf{X}^{(1)} = \{\mathbf{X}_t^{(1)}, t \in \mathbb{Z}\}$  et  $\mathbf{X}^{(2)} = \{\mathbf{X}_t^{(2)}, t \in \mathbb{Z}\}$  de dimension  $m_1$  et  $m_2$  respectivement avec  $m = m_1 + m_2$ . La matrice d'autocovariance  $\Gamma_{\mathbf{X}}(j)$  va être partitionnée en blocs matriciels comme suit:

$$\begin{pmatrix} \Gamma_{\mathbf{X}}^{(11)}(j) & \Gamma_{\mathbf{X}}^{(12)}(j) \\ \Gamma_{\mathbf{X}}^{(21)}(j) & \Gamma_{\mathbf{X}}^{(22)}(j) \end{pmatrix}, \quad j \in \mathbb{Z},$$

où les deux blocs diagonaux  $\Gamma_{\mathbf{X}}^{(hh)}(j)$ ;  $h = 1, 2$ , sont les matrices d'autocovariances au délai  $j$  associées à chacun des deux sous processus  $\mathbf{X}^{(h)}$ ,  $\Gamma_{\mathbf{X}}^{(12)}(j)$  est la matrice de covariance croisée de délai  $j$  entre ces deux sous processus. Pour tout délai positif  $j$ , on a  $\Gamma_{\mathbf{X}}^{(21)}(j) = \Gamma_{\mathbf{X}}^{(12)}(-j)^T$ .

La matrice de corrélation  $\boldsymbol{\rho}_{\mathbf{X}}(j)$  est partitionnée de façon similaire à  $\boldsymbol{\Gamma}_{\mathbf{X}}(j)$ , soit

$$\begin{pmatrix} \boldsymbol{\rho}_{\mathbf{X}}^{(11)}(j) & \boldsymbol{\rho}_{\mathbf{X}}^{(12)}(j) \\ \boldsymbol{\rho}_{\mathbf{X}}^{(21)}(j) & \boldsymbol{\rho}_{\mathbf{X}}^{(22)}(j) \end{pmatrix}, \quad j \in \mathbb{Z}.$$

La matrice  $\boldsymbol{\rho}_{\mathbf{X}}^{(12)}(j) = (\rho_{ls}^{(12)}(j))_{m_1 \times m_2}$  représente la matrice de corrélation croisée au délai  $j$  dont les éléments sont définis par:

$$\rho_{ls}^{(12)}(j) = \frac{\gamma_{ls}^{(12)}(j)}{\{\gamma_{lu}^{(11)}(0)\gamma_{ss}^{(22)}(0)\}^{1/2}}; \quad 1 \leq l \leq m_1; \quad 1 \leq s \leq m_2. \quad (1.1.9)$$

Aussi, notons que la densité spectrale possède la même répartition matricielle que la matrice de covariance. Ainsi, on peut représenter la fonction de cohérence  $\mathbf{S}_{\mathbf{X}}$  sous la forme matricielle suivante:

$$\mathbf{S}_{\mathbf{X}}(\lambda) = \begin{pmatrix} \mathbf{S}_{\mathbf{X}}^{(11)}(\lambda) & \mathbf{S}_{\mathbf{X}}^{(12)}(\lambda) \\ \mathbf{S}_{\mathbf{X}}^{(21)}(\lambda) & \mathbf{S}_{\mathbf{X}}^{(22)}(\lambda) \end{pmatrix}, \quad -\pi \leq \lambda \leq \pi, \quad (1.1.10)$$

où  $\mathbf{S}_{\mathbf{X}}^{(12)}(\lambda) = (\mathcal{S}_{uv}^{(12)}(\lambda))_{m_1 \times m_2}$  représente la matrice de cohérence croisée dont les éléments s'écrivent en fonction des éléments des matrices de densités spectrales des deux sous processus et de ceux des éléments de la matrice de densité spectrale croisée, *i.e.*

$$\mathcal{S}_{uv}^{(12)}(\lambda) = \frac{|f_{uv}^{(12)}(\lambda)|}{\sqrt{f_{uu}^{(11)}(\lambda)f_{vv}^{(22)}(\lambda)}}, \quad -\pi \leq \lambda \leq \pi. \quad (1.1.11)$$

Si on considère une réalisation de longueur  $N$  du processus  $\mathbf{X}$ , la matrice de covariance échantillonnale croisée au délai  $j$  est définie par

$$\mathbf{C}_{\mathbf{X}}^{(12)}(j) = N^{-1} \sum_{t=j+1}^N (\mathbf{X}_t^{(1)} - \bar{\mathbf{X}}^{(1)}) (\mathbf{X}_{t-j}^{(2)} - \bar{\mathbf{X}}^{(2)})^T, \quad (1.1.12)$$

pour  $0 \leq j \leq N-1$ . Aussi, on pose pour  $1-N \leq j \leq 0$ ,  $\mathbf{C}_{\mathbf{X}}^{(12)}(-j) = \mathbf{C}_{\mathbf{X}}^{(21)}(j)^T$  et  $\mathbf{C}_{\mathbf{X}}^{(12)}(j) = 0$  pour  $|j| > N$ . Ainsi, les éléments de la matrice de covariance croisée

$\mathbf{C}_{\mathbf{X}}^{(12)}(j) = (c_{ls}^{(12)}(j))_{m_1 \times m_2}$  s'écrivent sous la forme

$$c_{ls}^{(12)}(j) = \begin{cases} N^{-1} \sum_{t=j+1}^N (X_{l,t}^{(1)} - \bar{X}_l^{(1)}) (X_{s,t-j}^{(2)} - \bar{X}_s^{(2)}) & \text{si } 0 \leq j \leq N-1, \\ N^{-1} \sum_{t=-j+1}^N (X_{l,t+j}^{(1)} - \bar{X}_l^{(1)}) (X_{s,t}^{(2)} - \bar{X}_s^{(2)}) & \text{si } -N+1 \leq j \leq 0, \end{cases} \quad (1.1.13)$$

où  $\bar{\mathbf{X}}^{(h)} = N^{-1} \sum_{t=1}^N \mathbf{X}_t^{(h)} = (\bar{X}_1^{(h)}, \dots, \bar{X}_{m_h}^{(h)})^T$  est le vecteur moyenne échantionnale de la série multivariée  $\{\mathbf{X}_t^{(h)}\}$ ,  $h = 1, 2$ .

De même, on définit la matrice de corrélation croisée échantionnale de délai  $j$ ,  $0 \leq |j| \leq N-1$ , par  $\mathbf{R}_{\mathbf{X}}^{(12)}(j) = (r_{ls}^{(12)}(j))_{m_1 \times m_2}$ , où

$$r_{ls}^{(12)}(j) = \frac{c_{ls}^{(12)}(j)}{\sqrt{c_{ll}^{(11)}(0)c_{ss}^{(22)}(0)}}; \quad 1 \leq l \leq m_1; \quad 1 \leq s \leq m_2. \quad (1.1.14)$$

Sous une forme matricielle, on peut écrire

$$\mathbf{R}_{\mathbf{X}}^{(12)}(j) = \mathbf{D}_1 \{c_{ii}^{-1/2}(0)\} \mathbf{C}_{\mathbf{X}}^{(12)}(j) \mathbf{D}_2 \{c_{ii}^{-1/2}(0)\} \quad (1.1.15)$$

où la matrice  $\mathbf{D}_h$ ;  $h = 1, 2$  est définie par (1.1.4).

La matrice densité spectrale croisée échantionnale  $\hat{\mathbf{f}}_{\mathbf{X}}^{(12)}(\lambda) = (\hat{f}_{uv}^{(12)}(\lambda))_{m_1 \times m_2}$  est définie par

$$\hat{f}_{uv}^{(12)}(\lambda) = \frac{N^{-1}}{2\pi} \sum_{j=1-N}^{N-1} c_{uv}^{(12)}(j) e^{-ij\lambda}, \quad -\pi \leq \lambda \leq \pi.$$

Finalement, on définit la matrice de cohérence complexe croisée échantionnale

$\hat{\mathbf{S}}_{\mathbf{X}}^{(12)}(\lambda) = (\hat{S}_{uv}^{(12)}(\lambda))_{m_1 \times m_2}$  par

$$\hat{S}_{uv}^{(12)}(\lambda) = \frac{\hat{f}_{uv}^{(12)}(\lambda)}{\sqrt{\hat{f}_{uu}^{(11)}(\lambda)\hat{f}_{vv}^{(22)}(\lambda)}}, \quad -\pi \leq \lambda \leq \pi. \quad (1.1.16)$$

## 1.2. PROCESSUS LINÉAIRES

### 1.2.1. Processus moyennes mobiles d'ordre infini

Par le théorème de Wold, un processus  $\mathbf{X} = \{\mathbf{X}_t, t \in \mathbb{Z}\}$  stationnaire au sens large purement non-déterministe admet une représentation de la forme suivante:

$$\mathbf{X}_t = \boldsymbol{\mu} + \sum_{j=0}^{\infty} \boldsymbol{\Psi}_j \mathbf{a}_{t-j} = \boldsymbol{\mu} + \boldsymbol{\Psi}(B) \mathbf{a}_t, \quad t \in \mathbb{Z}, \quad (1.2.1)$$

où  $\boldsymbol{\Psi}_0 = \mathbf{I}_m$ ,  $\boldsymbol{\Psi}(z) = \mathbf{I}_m + \sum_{l=1}^{\infty} \boldsymbol{\Psi}_l z^l$ , les  $\boldsymbol{\Psi}_l$  représentent les coefficients matriciels du processus et  $B$  est l'opérateur retard tel que  $B^l \mathbf{a}_t = \mathbf{a}_{t-l}$ . La matrice identité de dimension  $m$  est notée  $\mathbf{I}_m$ . Pour assurer la convergence en moyenne quadratique, on suppose que la suite  $\boldsymbol{\Psi}_l$  est de carré sommable. Le processus  $\{\mathbf{a}_t\}$  de dimension  $m$  est un bruit blanc faible, c'est-à-dire  $\mathbb{E}(\mathbf{a}_t) = \mathbf{0}$ ,  $\mathbb{E}(\mathbf{a}_t \mathbf{a}_t^T) = \boldsymbol{\Sigma}$  et  $\mathbb{E}(\mathbf{a}_t \mathbf{a}_{t+l}^T) = \mathbf{0}$  pour  $l \neq 0$ . Aussi, pour satisfaire la condition d'inversibilité, on suppose que

$$\det \{\boldsymbol{\Psi}(z)\} \neq 0 \text{ pour } |z| \leq 1. \quad (1.2.2)$$

La représentation (1.2.1) définit un processus linéaire et est aussi appelée représentation moyenne mobile d'ordre infini, qu'on note  $\text{MA}(\infty)$ .

Dans plusieurs résultats concernant des propriétés des processus moyennes mobiles, on suppose que la série  $(\boldsymbol{\Psi}_l)_{l=0}^{\infty}$  est absolument convergente.

La fonction d'autocovariance d'un processus  $\text{MA}(\infty)$  peut être définie en termes des coefficients moyenne mobile de la façon suivante

$$\boldsymbol{\Gamma}(j) = \sum_{l=0}^{\infty} \boldsymbol{\Psi}_l \boldsymbol{\Sigma} \boldsymbol{\Psi}_{l+j}^T.$$

Et ainsi, d'après Fuller (1996, p. 180), on peut déduire la matrice densité spectrale du processus MA( $\infty$ ) et qui est donnée par

$$\mathbf{f}(\lambda) = \frac{1}{2\pi} \mathbf{\Psi}(e^{i\lambda}) \mathbf{\Sigma} \mathbf{\Psi}(e^{-i\lambda})^T, \quad -\pi \leq \lambda \leq \pi.$$

Finalement, on peut noter que la représentation (1.2.1) dite, *représentation de Wold*, est définie quand  $\mathbf{a}_t$  décrit l'erreur de prévision du meilleur prédicteur linéaire de  $\mathbf{X}_t$ , qu'on note par  $\hat{\mathbf{X}}_{t-1}(1)$ , basé sur une infinité de valeurs passées  $\mathbf{X}_{t-1}, \mathbf{X}_{t-2}, \dots$ , c'est-à-dire  $\mathbf{a}_t = \mathbf{X}_t - \hat{\mathbf{X}}_{t-1}(1)$ .

### 1.2.2. Processus autorégressifs d'ordre infini

Un processus multivarié  $\mathbf{X} = \{\mathbf{X}_t, t \in \mathbb{Z}\}$  stationnaire au sens large de dimension  $m$  admet une représentation autorégressive d'ordre infini VAR( $\infty$ ) si on peut l'écrire sous la forme

$$\mathbf{X}_t - \sum_{j=1}^{\infty} \mathbf{\Phi}_j \mathbf{X}_{t-j} = \mathbf{\Phi}(B) \mathbf{X}_t = \mathbf{a}_t, \quad (1.2.3)$$

où  $\mathbf{\Phi}(z) = \mathbf{I}_m - \sum_{l=1}^{\infty} \mathbf{\Phi}_l z^l$ ,  $\sum_{l=1}^{\infty} \|\mathbf{\Phi}_l\|_2^2 < \infty$ , et les  $\mathbf{\Phi}_l$  sont les coefficients matriciels autorégressifs du processus. En termes des coefficients, la condition de stationnarité s'écrit

$$\det \{\mathbf{\Phi}(z)\} \neq 0 \text{ pour } |z| \leq 1. \quad (1.2.4)$$

Un processus VAR( $\infty$ ) stationnaire admet une représentation MA( $\infty$ ) où  $\mathbf{\Psi}(z) = \mathbf{\Phi}(z)^{-1}$ . Le processus d'innovation  $\{\mathbf{a}_t\}$  est le même que celui apparaissant dans la représentation MA( $\infty$ ) définie par (1.2.1).

### 1.2.3. Processus ARMA multivariés

Un processus multivarié  $\mathbf{X} = \{\mathbf{X}_t, t \in \mathbb{Z}\}$  admet une représentation autorégressive-moyenne mobile d'ordre  $(p, q)$ , qu'on note VARMA( $p, q$ ), si on peut l'écrire sous la forme suivante:

$$(\mathbf{X}_t - \boldsymbol{\mu}) - \sum_{l=1}^p \boldsymbol{\Phi}_l (\mathbf{X}_{t-l} - \boldsymbol{\mu}) = \mathbf{a}_t - \sum_{l=1}^q \boldsymbol{\Theta}_l \mathbf{a}_{t-l}, \quad (1.2.5)$$

ou bien en utilisant l'opérateur retard  $B$

$$\boldsymbol{\Phi}(B) (\mathbf{X}_t - \boldsymbol{\mu}) = \boldsymbol{\Theta}(B) \mathbf{a}_t \quad (1.2.6)$$

où  $\boldsymbol{\Phi}(z) = \mathbf{I}_m - \sum_{l=1}^p \boldsymbol{\Phi}_l z^l$ ,  $\boldsymbol{\Theta}(z) = \mathbf{I}_m - \sum_{l=1}^q \boldsymbol{\Theta}_l z^l$  et les  $\boldsymbol{\Phi}_l$ ,  $\boldsymbol{\Theta}_l$  sont les coefficients réels matriciels  $m \times m$  du processus  $\mathbf{X}$ . Le processus  $\{\mathbf{a}_t\}$  est un bruit blanc faible de moyenne nulle et de matrice de covariance  $\boldsymbol{\Sigma}$ .

Brockwell et Davis (1991, p. 418-19) ont montré que si la condition de stationnarité (1.2.4) est vérifiée, le processus (1.2.5) sera stationnaire et ainsi, peut être représenté sous une forme moyenne mobile d'ordre infini, comme dans (1.2.1), et  $\boldsymbol{\Psi}(z) = \boldsymbol{\Phi}(z)^{-1} \boldsymbol{\Theta}(z)$  représente la série des coefficients matriciels du processus qui est convergente pour  $|z| \leq 1$ . Aussi, si on suppose que la condition d'inversibilité (1.2.2) est satisfaite, le processus (1.2.5) sera alors inversible avec  $\boldsymbol{\Pi}(z) = \boldsymbol{\Theta}(z)^{-1} \boldsymbol{\Phi}(z) = \mathbf{I}_m - \sum_{l=1}^{\infty} \boldsymbol{\Pi}_l z^l$  qui représente la série matricielle convergente pour  $|z| \leq 1$  des paramètres du processus  $\{\mathbf{X}_t\}$  de la représentation autorégressive d'ordre infini  $\boldsymbol{\Pi}(B) (\mathbf{X}_t - \boldsymbol{\mu}) = \mathbf{a}_t$ .

Ainsi, dans tout ce qui suit et sans perte de généralité, on ne va considérer que les processus autorégressifs d'ordre infini  $\boldsymbol{\Pi}(B) (\mathbf{X}_t - \boldsymbol{\mu}) = \mathbf{a}_t$ .

### 1.3. APPROXIMATION D'UN PROCESSUS AUTORÉGRESSIF D'ORDRE INFINI

Considérons un processus multivarié  $\mathbf{X} = \{\mathbf{X}_t, t \in \mathbb{Z}\}$  solution de l'équation aux différences (1.2.3). Supposons que les coefficients matriciels du processus  $(\Phi_l)_{l=0}^{\infty}$  satisfont la condition de stationnarité (1.2.4) et qu'ils sont absolument sommables.

À partir d'une réalisation de  $N$  observations, on veut approximer le processus  $\mathbf{X}$  par un processus autorégressif VAR( $p$ ), où  $p$  désigne l'ordre autorégressif. Pour les développements asymptotiques, nous supposons que  $p$  dépend de  $N$  et nous écrivons  $p = p(N)$ . Considérons  $\Phi(p) = (\Phi_{1,p}, \dots, \Phi_{p,p})$  la matrice formée des  $p$  coefficients autorégressifs. L'estimateur de Yule-Walker  $\hat{\Phi}(p)$  correspondant est donné par

$$\hat{\Phi}(p) = (\hat{\Phi}_{1,p}, \dots, \hat{\Phi}_{p,p}) = \hat{\mathbf{A}}_{1,p}^T \hat{\mathbf{A}}_p^{-1}, \quad (1.3.1)$$

avec

$$\hat{\mathbf{A}}_{1,p} = (N-p)^{-1} \sum_{t=p+1}^N \mathbf{X}_t(p) \mathbf{X}_t^T \quad (1.3.2)$$

et

$$\hat{\mathbf{A}}_p = (N-p)^{-1} \sum_{t=p+1}^N \mathbf{X}_t(p) \mathbf{X}_t^T(p). \quad (1.3.3)$$

où

$$\mathbf{X}_t(p) = (\mathbf{X}_{t-1}^T, \mathbf{X}_{t-2}^T, \dots, \mathbf{X}_{t-p}^T)^T. \quad (1.3.4)$$

Ainsi les résidus résultants sont donnés par

$$\hat{\mathbf{a}}_{p,t} = \mathbf{X}_t - \sum_{l=1}^p \hat{\Phi}_{l,p} \mathbf{X}_{t-l}, \quad t = p+1, \dots, N.$$

Soit  $\hat{\Sigma}_p$  l'estimateur de  $\Sigma$  défini par

$$\hat{\Sigma}_p = (N-p)^{-1} \sum_{t=p+1}^N \tilde{\mathbf{a}}_{p,t} \tilde{\mathbf{a}}_{p,t}^T \quad (1.3.5)$$

où  $\tilde{\mathbf{a}}_{p,t} = \hat{\mathbf{a}}_{p,t} - \bar{\mathbf{a}}_p$  et  $\bar{\mathbf{a}}_p = (N-p)^{-1} \sum_{t=p+1}^N \hat{\mathbf{a}}_{p,t}$ .

Dans la suite,  $\text{vec}(\mathbf{B})$  désigne l'opérateur usuel qui transforme la matrice  $\mathbf{B}$  en un vecteur colonne. Le symbole  $\otimes$  représente le produit de Kronecker de deux matrices.

Considérons  $\{l(s)\}_{s \in N}$  une séquence arbitraire de vecteurs de dimension  $sm^2 \times 1$  satisfaisant la condition  $0 < M_1 \leq \|l(s)\|_2^2 \leq M_2 < \infty$ , et notons par  $\Phi(s) = \text{vec}(\Phi(s))$  et  $\hat{\Phi}(s) = \text{vec}(\hat{\Phi}(s))$ . Pour  $s$  et  $N \rightarrow \infty$ , les propriétés asymptotiques de  $\hat{\Phi}(s)$  ont été étudiées par Lewis et Reinsel (1985). D'abord, voici le théorème qui nous donne la convergence de l'estimateur  $\hat{\Phi}(p)$ .

**Théorème 1.3.1 (Lewis et Reinsel, 1985).**

Soit  $\{\mathbf{X}_t\}$  un processus  $\text{VAR}(\infty)$  défini par la relation (1.2.3). Aussi, supposons que

- (i)  $\mathbb{E}|a_{i,t} a_{j,t} a_{k,t} a_{l,t}| \leq \gamma_4 < \infty$ ,  $1 \leq i, j, k, l \leq m$ ;
- (ii)  $p$  est une fonction de  $N$  tel que  $p(N) \rightarrow \infty$  et  $p(N)^2/N \rightarrow 0$ , quand  $N \rightarrow \infty$ ;
- (iii)  $\sqrt{p(N)} \sum_{j=p+1}^{\infty} \|\Phi_j\|_2 \rightarrow 0$ , quand  $N \rightarrow \infty$ .

Alors, nous avons

$$\|\hat{\Phi}(p) - \Phi(p)\|_2 = O_p\left(\frac{p(N)^{1/2}}{N^{1/2}}\right). \quad (1.3.6)$$

La démonstration du résultat (1.3.6) est implicite dans la preuve du théorème 1 de Lewis et Reinsel (1985). Voir aussi le théorème 2.1 de Paparoditis (1996). Pour la normalité asymptotique, nous aurons besoin de la notation suivante:  $\Gamma_p = E(\mathbf{X}_t(p)\mathbf{X}_t(p)^T)$ , la matrice de covariance du vecteur  $\mathbf{X}_t(p)$ .

**Théorème 1.3.2 (Reinsel et Lewis, 1985).**

Si on considère  $\{\mathbf{X}_t\}$  un processus  $VAR(\infty)$  qui satisfait la relation (1.2.3), et sous les conditions suivantes:

- (i)  $\mathbb{E}|a_{i,t}a_{j,t}a_{k,t}a_{l,t}| \leq \gamma_4 < \infty$ ,  $1 \leq i, j, k, l \leq m$ ;
- (ii)  $p$  est une fonction de  $N$  tel que  $p(N) \rightarrow \infty$  et  $p(N)^3/N \rightarrow 0$ , quand  $N \rightarrow \infty$ ;
- (iii)  $\sqrt{N} \sum_{j=p+1}^{\infty} \|\Phi_j\|_2 \rightarrow 0$ , quand  $N \rightarrow \infty$
- (iv)  $\{l(s)\}_{s \in N}$  une séquence arbitraire de vecteurs de dimension  $sm^2 \times 1$  satisfaisant la condition  $0 < M_1 \leq \|l(s)\|_2^2 = l(s)^T l(s) \leq M_2 < \infty$ .

Alors, il s'en suit que

$$(N-p)^{1/2} l(p)^T (\hat{\Phi}(p) - \Phi(p)) / v_N \xrightarrow{d} N(0, 1) \text{ quand } N \rightarrow \infty, \quad (1.3.7)$$

où  $v_N^2 = l(p)^T (\Gamma_p^{-1} \otimes \Sigma) l(p)$ .

De même, comme on a vu à la section 1.2.2, le processus  $\mathbf{X}$  autorégressif d'ordre infini peut être exprimé comme un processus  $MA(\infty)$  tel que défini par (1.2.1) où la suite des coefficients moyenne mobile  $\Psi_j$  est absolument sommable.

Les coefficients  $\Psi_j$  sont donnés en termes des matrices  $\Phi_j$  par  $\Psi(z) = \Phi(z)^{-1} = \mathbb{I}_m + \sum_{j=1}^{\infty} \Psi_j z^j$ . Pour utiliser les résultats déjà existants sur les paramètres autorégressifs  $\Phi_j$ , Lütkepohl (1989, p.372) a exprimé les coefficients matriciels  $\Psi_j$  en termes des coefficients  $\Phi_j$ . Ils sont donnés par

$$\Psi_j = \mathbf{J}_p \Pi_p^j \mathbf{J}_p^T$$

où  $\mathbf{J}_p$  est la matrice  $m \times mp$  définie par  $\mathbf{J}_p = (\mathbf{I}_m \mathbf{0} \dots \mathbf{0})$ ,  $\mathbf{0}$  est la matrice  $m \times m$  nulle, et  $\mathbf{\Pi}_p$  est la matrice bloc  $mp \times mp$  définie comme suit:

$$\mathbf{\Pi}_p = \begin{pmatrix} \mathbf{\Phi}_1 & \mathbf{\Phi}_2 & \cdots & \cdot & \mathbf{\Phi}_p \\ \mathbf{I}_m & \mathbf{0} & \cdots & \cdot & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_m & \cdots & \cdot & \mathbf{0} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I}_m & \mathbf{0} \end{pmatrix}$$

Dans son livre, Lütkepohl (1991, p.314) a étudié sous les mêmes conditions que celles des théorèmes 1.3.1 et 1.3.2, les propriétés asymptotiques du vecteur  $\hat{\Psi}(p)$  où  $\hat{\Psi}(p) = \text{vec}(\hat{\Psi}_{1,p}, \dots, \hat{\Psi}_{p,p})$  avec  $\hat{\Psi}_{j,p} = \mathbf{J}_p \hat{\Pi}_p^j \mathbf{J}_p^T$  et  $\hat{\Pi}_p$  est obtenu en remplaçant dans la matrice  $\mathbf{\Pi}_p$  les matrices  $\mathbf{\Phi}_j$  par leurs estimateurs  $\hat{\Phi}_{j,p}$ .

## Chapitre 2

---

### TESTS D'INDÉPENDANCE BASÉS SUR UN SOUS-ENSEMBLE DE DÉLAIS

Dans l'étude des séries temporelles, les processus linéaires stationnaires ont été utilisés fréquemment au cours des dernières années. Ces processus incluent en particulier les processus autorégressifs, moyennes mobiles et autorégressifs-moyennes mobiles qu'on note souvent, respectivement, par AR, MA et ARMA. Ces derniers ont fait l'objet de plusieurs études concernant l'identification, l'estimation et la validation ou diagnostic. Un des outils les plus utilisés dans la dernière étape de validation du modèle est la fonction de corrélation des séries résiduelles. La distribution asymptotique de ces corrélations résiduelles a été étudiée dans le but de l'utiliser dans la construction de tests pour chercher la liaison existante entre deux processus linéaires stationnaires. L'étude des relations qui peuvent exister entre deux séries chronologiques univariées ou multivariées intervient dans plusieurs domaines telles que l'économétrie ou la finance.

Dans le cas univarié, Haugh (1976) a utilisé une approche de modélisation par des processus ARMA pour tester l'indépendance de deux séries univariées. Les corrélations croisées entre les deux séries résiduelles issues de cette modélisation sont ensuite utilisées dans le calcul d'une statistique de type portemanteau pour tester l'hypothèse d'indépendance. McLeod (1979) s'est intéressé à des tests

de corrélations croisées à des délais positifs ou des délais négatifs lorsqu'il y a corrélation instantanée entre les deux séries.

Dans le cas multivarié, El Himdi et Roy (1997) ont généralisé l'approche univariée de Haugh (1976) au cas de deux séries multivariées stationnaires au second ordre. Ils ont établi que sous l'hypothèse de non corrélation entre les séries originales, la distribution asymptotique d'un ensemble fini de corrélations croisées entre les deux séries résiduelles issues de la procédure de filtrage est la même que celle du vecteur de corrélations croisées entre les deux séries d'innovations correspondantes. La distribution asymptotique de ces corrélations croisées résiduelles a permis de déduire une procédure asymptotique pour tester la non corrélation entre deux séries multivariées.

Dans ce chapitre, nous nous proposons de généraliser l'approche d'El Himdi et Roy (1997) au cas de deux processus linéaires multivariés stationnaires. Nous allons tout d'abord rappeler quelques résultats concernant la distribution asymptotique d'ensembles finis de corrélations et covariances croisées sérielles entre deux séries chronologiques multivariées stationnaires au second ordre. Nous généralisons la procédure de tests d'El Himdi et Roy (1997) au cas de deux séries multivariées d'ordre infini  $\text{VAR}(\infty)$ . Nous établissons ainsi que sous l'hypothèse de non corrélation entre les deux séries originales, les distributions asymptotiques d'un ensemble fini de corrélations croisées entre les séries résiduelles issues de la procédure d'approximation des deux séries par des autorégressions d'ordres finis est la même que celle des corrélations croisées correspondantes entre les deux séries d'innovations. La distribution asymptotique de ces corrélations croisées

résiduelles permettra de déduire une procédure asymptotique pour tester l'indépendance entre deux séries multivariées.

## 2.1. DÉFINITIONS ET GÉNÉRALITÉS

Considérons une réalisation de longueur  $N$  d'un processus multivarié stationnaire  $\mathbf{X} = \left\{ \left( \mathbf{X}_t^{(1)T}, \mathbf{X}_t^{(2)T} \right)^T, t \in \mathbb{Z} \right\}$  de dimension  $m = m_1 + m_2$  où  $m_1$  et  $m_2$  sont les dimensions des sous processus  $\mathbf{X}^{(1)}$  et  $\mathbf{X}^{(2)}$  respectivement. Soit  $d_1, \dots, d_n$  une suite de  $n$  entiers distincts tels que  $|d_i| < N, i = 1, \dots, n$ . Nous définissons les vecteurs  $\mathbf{c}_\mathbf{X}$  et  $\mathbf{r}_\mathbf{X}$  des covariances et des corrélations échantionnales, de dimensions  $nm^2$  chacun par:

$$\mathbf{c}_\mathbf{X} = \left( \text{vec}(\mathbf{C}_\mathbf{X}(d_1))^T, \dots, \text{vec}(\mathbf{C}_\mathbf{X}(d_n))^T \right)^T,$$

$$\mathbf{r}_\mathbf{X} = \left( \text{vec}(\mathbf{R}_\mathbf{X}(d_1))^T, \dots, \text{vec}(\mathbf{R}_\mathbf{X}(d_n))^T \right)^T.$$

De même on introduit les vecteurs  $\mathbf{c}_\mathbf{X}^{(12)}$  et  $\mathbf{r}_\mathbf{X}^{(12)}$  de dimensions  $nm_1m_2$  des covariances et corrélations sérielles croisées entre les deux processus  $\mathbf{X}^{(1)}$  et  $\mathbf{X}^{(2)}$ :

$$\mathbf{c}_\mathbf{X}^{(12)} = \left( \text{vec}(\mathbf{C}_\mathbf{X}^{(12)}(d_1))^T, \dots, \text{vec}(\mathbf{C}_\mathbf{X}^{(12)}(d_n))^T \right)^T, \quad (2.1.1)$$

$$\mathbf{r}_\mathbf{X}^{(12)} = \left( \text{vec}(\mathbf{R}_\mathbf{X}^{(12)}(d_1))^T, \dots, \text{vec}(\mathbf{R}_\mathbf{X}^{(12)}(d_n))^T \right)^T. \quad (2.1.2)$$

Par un choix approprié des délais  $d_1, \dots, d_n$ , ces deux vecteurs ont été souvent utilisés dans la littérature pour analyser la structure de dépendance pouvant exister entre les deux processus multivariés  $\mathbf{X}^{(1)}$  et  $\mathbf{X}^{(2)}$ . L'étude de leurs distributions est nécessaire pour l'élaboration de tests d'indépendance entre deux séries vectorielles stationnaires.

Dans un contexte de séries chronologiques, Roy et Cléroux (1993) ont généralisé la concept de corrélation vectorielle introduit par Escoufier (1973). Les deux auteurs ont proposé le coefficient suivant de corrélation vectorielle croisée au délai  $j$  entre deux processus vectoriels  $\mathbf{X}^{(1)}$  et  $\mathbf{X}^{(2)}$

$$\lambda_{\mathbf{X}}^{(12)}(j) = \frac{\text{tr} \left( \boldsymbol{\rho}_{\mathbf{X}}^{(12)}(j) \boldsymbol{\rho}_{\mathbf{X}}^{(12)}(j)^T \right)}{\left\{ \text{tr} \left( \boldsymbol{\rho}_{\mathbf{X}}^{(11)}(0)^2 \right) \text{tr} \left( \boldsymbol{\rho}_{\mathbf{X}}^{(22)}(0)^2 \right) \right\}^{1/2}} \quad j \in \mathbb{Z}. \quad (2.1.3)$$

Ce coefficient de corrélation vectoriel est tel que  $0 \leq \lambda_{\mathbf{X}}^{(12)}(j) \leq 1$  et, une condition nécessaire et suffisante de non corrélation au délai  $j$  entre  $\mathbf{X}^{(1)}$  et  $\mathbf{X}^{(2)}$ , i.e.  $\boldsymbol{\rho}_{\mathbf{X}}^{(12)}(j) = 0$ , est que  $\lambda_{\mathbf{X}}^{(12)}(j) = 0$ . Les mêmes auteurs, voir Cléroux et Roy (1988), ont introduit un autre indicateur de corrélation de type scalaire en remplaçant, dans la définition (2.1.3), les vecteurs de corrélations croisées par les matrices de covariances croisées et qu'ils ont noté  $\tilde{\lambda}_{\mathbf{X}}^{(12)}(j)$ . Ces deux coefficients de corrélation vectorielle ont les mêmes propriétés sauf pour l'invariance qui est une propriété propre à  $\lambda_{\mathbf{X}}^{(12)}(j)$ . Des estimateurs convergents, notés  $\hat{\lambda}_{\mathbf{X}}^{(12)}(j)$  et  $\hat{\tilde{\lambda}}_{\mathbf{X}}^{(12)}(j)$ , des coefficients  $\lambda_{\mathbf{X}}^{(12)}(j)$  et  $\tilde{\lambda}_{\mathbf{X}}^{(12)}(j)$  sont obtenus en remplaçant les matrices de corrélations et de covariances théoriques par leurs estimations. De plus, Roy et Cléroux (1993) ont étudié la distribution asymptotique de l'estimateur  $\hat{\lambda}_{\mathbf{X}}^{(12)}(j)$  sous l'hypothèse nulle de non corrélation au délai  $j$ ,  $\lambda_{\mathbf{X}}^{(12)}(j) = 0$ , et aussi sous l'alternative  $0 < \lambda_{\mathbf{X}}^{(12)}(j) < 1$ . Ce qui permet de déduire un test de non corrélation pour différents délais  $j$  entre les deux séries considérées.

## 2.2. DISTRIBUTION ASYMPTOTIQUE DE $\mathbf{r}_{\hat{\mathbf{a}}}^{(12)}$ POUR DEUX PROCESSUS ARMA MULTIVARIÉS

Les deux vecteurs définis par (2.1.1) et (2.1.2) vont être utilisés pour l'élaboration du test d'indépendance des deux sous processus  $\mathbf{X}^{(h)}$  pour  $h = 1, 2$ .

El Himdi et Roy (1997) ont étudié la distribution asymptotique de ces vecteurs, pour des délais quelconques  $d_1, \dots, d_n$ , dans le cas où les deux sous processus sont autorégressifs-moyennes mobiles multivariés d'ordres finis (ARMA( $p_h, q_h$ )), c'est à dire que  $\mathbf{X}_t^{(h)}$  est solution de

$$\Phi^{(h)}(B)\mathbf{X}_t^{(h)} = \Theta^{(h)}(B)\mathbf{a}_t^{(h)} \quad (2.2.1)$$

où  $\Theta^{(h)}(z) = \mathbf{I}_{m_h} - \sum_{l=1}^{q_h} \Theta_l^{(h)} z^l$  et  $\Phi^{(h)}(z) = \mathbf{I}_{m_h} - \sum_{l=1}^{p_h} \Phi_l^{(h)} z^l$ .  $\{\mathbf{a}_t^{(h)}\}$  est un bruit blanc dont les matrices de covariance et de corrélation sont notées respectivement  $\Sigma_h$  et  $\rho_h$ . Ces deux sous processus vérifient les conditions de stationnarité (1.2.4) et d'inversibilité (1.2.2).

Dans tout ce qui suit, on suppose que les deux sous processus  $\mathbf{X}^{(1)}$  et  $\mathbf{X}^{(2)}$  sont non corrélés. Ainsi, on peut remarquer que le processus  $\mathbf{X} = \left\{ \left( \mathbf{X}_t^{(1)T}, \mathbf{X}_t^{(2)T} \right)^T ; t \in \mathbb{Z} \right\}$  qui est un processus résultant multivarié de dimension  $m = m_1 + m_2$  est décrit par le modèle ARMA( $p, q$ ) stationnaire et inversible

$$\Phi(B)\mathbf{X}_t = \Theta(B)\mathbf{a}_t$$

de polynôme autorégressif  $\Phi(B) = \mathbf{I}_m - \sum_{l=1}^p \Phi_l B^l$  de degré  $p = \max(p_1, p_2)$  et de polynôme moyenne mobile  $\Theta(B) = \mathbf{I}_m - \sum_{r=1}^q \Theta_r B^r$  de degré  $q = \max(q_1, q_2)$ .

Les coefficients  $\Phi_l$ ,  $l = 1, \dots, p$  et  $\Theta_r$ ,  $r = 1, \dots, q$ , sont donnés par:

$$\Phi_l = \begin{pmatrix} \Phi_l^{(1)} & \mathbf{0} \\ \mathbf{0} & \Phi_l^{(2)} \end{pmatrix}, \quad \Theta_r = \begin{pmatrix} \Theta_r^{(1)} & \mathbf{0} \\ \mathbf{0} & \Theta_r^{(2)} \end{pmatrix},$$

avec  $\Phi_l^{(h)} = \mathbf{0}$  si  $p_h < l \leq p$  et  $\Theta_r^{(h)} = \mathbf{0}$  si  $q_h < r \leq q$ . Puisque  $\det\{\Phi(z)\} = \det\{\Phi^{(1)}(z)\} \det\{\Phi^{(2)}(z)\}$  et  $\det\{\Theta(z)\} = \det\{\Theta^{(1)}(z)\} \det\{\Theta^{(2)}(z)\}$ , la stationnarité et l'inversibilité du processus  $\{\mathbf{X}_t\}$  se déduisent directement de celles de ses deux composantes  $\{\mathbf{X}_t^{(1)}\}$  et  $\{\mathbf{X}_t^{(2)}\}$ . Aussi,  $\{\mathbf{a}_t\}$  est un bruit blanc de matrice

de covariance définie positive

$$\Sigma = \begin{pmatrix} \Sigma_1 & \mathbf{0} \\ \mathbf{0} & \Sigma_2 \end{pmatrix}. \quad (2.2.2)$$

La matrice de corrélation est donnée par

$$\rho = \begin{pmatrix} \rho_1 & \mathbf{0} \\ \mathbf{0} & \rho_2 \end{pmatrix}. \quad (2.2.3)$$

Dans le but d'étudier l'inter-dépendance entre deux processus ARMA multivariés  $\mathbf{X}^{(1)}$  et  $\mathbf{X}^{(2)}$ , El Himdi et Roy (1997) ont utilisé la fonction matricielle de corrélation croisée entre les innovations associées à ces deux processus, notée  $\rho_{\mathbf{a}}^{(12)}(j)$ . De l'inversibilité et de la stationnarité des sous processus, il s'en suit l'équivalence des deux hypothèses suivantes, soit

$$\rho_{\mathbf{X}}^{(12)}(j) = \mathbf{0}, \forall j \Leftrightarrow \rho_{\mathbf{a}}^{(12)}(j) = \mathbf{0}, \forall j. \quad (2.2.4)$$

En disposant des séries résiduelles  $\{\hat{\mathbf{a}}_t^{(h)}\}$ , obtenues d'un ajustement du modèle (2.2.1) aux deux séries, les auteurs ont estimé les matrices  $\Sigma$  et  $\rho_{\mathbf{a}}^{(12)}(j)$  par les caractéristiques échantillonales résiduelles correspondantes  $\mathbf{C}_{\hat{\mathbf{a}}}^{(hh)}(0)$  et  $\mathbf{R}_{\hat{\mathbf{a}}}^{(12)}(j)$ . Ainsi, l'étude de la distribution asymptotique de ces estimateurs a constitué l'étape principale dans la construction du test d'indépendance généralisant celui utilisé par Haugh (1976) dans le cas univarié. Les deux auteurs ont montré le théorème suivant.

### **Théorème 2.2.1.**

*Soient  $\mathbf{X}^{(1)}$  et  $\mathbf{X}^{(2)}$  deux processus stationnaires au second ordre vérifiant le modèle ARMA multivarié (2.2.1) et satisfaisant certaines hypothèses de régularité. Si ces deux processus sont non corrélés, alors les deux vecteurs  $\sqrt{N}\mathbf{r}_{\hat{\mathbf{a}}}^{(12)}$  et  $\sqrt{N}\mathbf{r}_{\hat{\mathbf{a}}}^{(12)}$  définis par (2.1.2) possèdent la même distribution asymptotique  $\mathbf{N}(\mathbf{0}, \mathbf{I}_n \otimes (\rho_2 \otimes \rho_1))$ .*

La distribution asymptotique de  $\sqrt{N}\mathbf{r}_a^{(12)}$  découle d'un résultat général de Roy (1989), voir aussi Chitturi (1976) et Li et McLeod (1981). El Himdi et Roy (1997) ont construit un test de non corrélation entre deux processus multivariés ARMA stationnaire et inversible. Ce test consiste à trancher entre l'hypothèse nulle de non corrélation  $\mathcal{H}_0$  et l'hypothèse alternative  $\mathcal{H}_1$  données par

$$\mathcal{H}_0 : \rho_a^{(12)}(j) = \mathbf{0} \quad \forall j \in \mathbb{Z} \quad \text{vs} \quad \mathcal{H}_1 : \exists j \in \mathbb{Z}, |j| \leq M \text{ et } \rho_a^{(12)}(j) \neq \mathbf{0}.$$

Pour un ensemble fini de délais  $\{d_1, \dots, d_n\}$ , la distribution asymptotique du vecteur  $\mathbf{r}_a^{(12)}$  donnée par le théorème ci dessus, est particulièrement simple à utiliser pour construire des tests pour  $\mathcal{H}_0$ .

### 2.3. GÉNÉRALISATION AU CAS DE DEUX PROCESSUS MULTIVARIÉS $VAR(\infty)$

Dans cette section, nous généraliserons le théorème 2.3.1 à des processus autorégressifs multivariés d'ordre infini tels que définis par (1.2.3). Ainsi, nous supposons, pour  $h = 1, 2$ , que  $\{\mathbf{X}_t^{(h)}\}$  sont deux sous processus autorégressifs multivariés d'ordre infini  $VAR(\infty)$  qui vérifient l'équation aux différences (1.2.3). La série des coefficients autorégressifs  $(\Phi_t^{(h)})_{t=0}^{\infty}$  vérifie la condition de stationnarité (1.2.4) et est supposée absolument sommable. Le processus  $\mathbf{a} = \left\{ \left( \mathbf{a}_t^{(1)T}, \mathbf{a}_t^{(2)T} \right)^T, t \in \mathbb{Z} \right\}$  est un bruit blanc de matrice de covariance  $\Sigma$  est de la forme (2.2.2). En plus, les deux sous processus  $\{\mathbf{a}_t^{(h)}\}$  vérifient l'hypothèse suivante.

#### Hypothèse 2.3.1.

(i)  $\{\mathbf{a}_t^{(h)}\}$  est un bruit blanc de dimension  $m_h$ , avec une matrice de covariance  $\Sigma_h$  i.e.  $E(\mathbf{a}_t^{(h)}) = \mathbf{0}$  et  $E(\mathbf{a}_t^{(h)} \mathbf{a}_t^{(h)T}) = \Sigma_h$ .

(ii) Les moments d'ordre quatre des composantes  $\{a_{i,t}^{(h)}, i = 1, \dots, m_h\}$  du vecteur  $\mathbf{a}_t^{(h)}$  existent i.e. pour tout  $t$ ,

$$E|a_{i,t}^{(h)} a_{j,t}^{(h)} a_{k,t}^{(h)} a_{l,t}^{(h)}| < \infty, \quad i, j, k, l = 1, \dots, m_h.$$

La condition (ii) est équivalente à la condition d'existence de tous les cumulants d'ordre quatre, notés  $\kappa_{ijkl}$ , du sous processus  $\{\mathbf{a}_t^{(h)}\}$ , avec

$$\kappa_{ijkl} = \text{cum}\{a_{i,t}^{(h)}, a_{j,t}^{(h)}, a_{k,t}^{(h)}, a_{l,t}^{(h)}\}$$

dont l'existence s'avère nécessaire pour établir plusieurs résultats asymptotiques des estimateurs des moindres carrés des paramètres autorégressifs, des covariances échantillonnales et aussi concernant l'approximation d'un processus VAR( $\infty$ ) par un modèle autorégressif d'ordre fini.

La statistique du test d'indépendance entre les deux processus multivariés  $\mathbf{X}^{(1)}$  et  $\mathbf{X}^{(2)}$  est basée sur les séries résiduelles  $\{\hat{\mathbf{a}}_t^{(h)}\}$  obtenues en approximant les deux sous processus VAR( $\infty$ ) par un processus autorégressif d'ordre fini. Ainsi, à partir de réalisations de longueur  $N$ , on ajuste un modèle AR( $p_h$ ) à chaque série. L'ordre autorégressif  $p_h$  varie avec  $N$  et au besoin, nous écrivons  $p_h(N)$ . Pour  $h = 1, 2$ , les résidus résultants sont donnés par l'équation

$$\hat{\mathbf{a}}_t^{(h)} = \begin{cases} \mathbf{X}_t^{(h)} - \sum_{l=1}^{p_h} \hat{\Phi}_{l,p_h} \mathbf{X}_{t-l}^{(h)} & \text{si } t = p_h + 1, \dots, N, \\ \mathbf{0} & \text{si } t \leq p_h, \end{cases} \quad (2.3.1)$$

où les  $\hat{\Phi}_{l,p_h}$  sont les estimateurs de Yule-Walker des coefficients autorégressifs donnés par (1.3.1) ou des estimateurs asymptotiquement équivalents. La matrice

de covariance croisée résiduelle est définie par  $\mathbf{C}_{\hat{\mathbf{a}}}^{(12)}(j) = (\hat{c}_{ls}^{(12)}(j))_{m_1 \times m_2}$ , où

$$\hat{c}_{ls}^{(12)}(j) = \begin{cases} N^{-1} \sum_{t=j+1}^N \hat{a}_{l,t}^{(1)} \hat{a}_{s,t-j}^{(2)} & \text{si } 0 \leq j \leq N-1, \\ N^{-1} \sum_{t=-j+1}^N \hat{a}_{l,t+j}^{(1)} \hat{a}_{s,t}^{(2)} & \text{si } -N+1 \leq j \leq 0. \end{cases} \quad (2.3.2)$$

La matrice de corrélation croisée résiduelle  $\mathbf{R}_{\hat{\mathbf{a}}}^{(12)}(j) = (\hat{r}_{ls}^{(12)}(j))_{m_1 \times m_2}$  est donnée par  $\hat{r}_{ls}^{(12)}(j) = \frac{\hat{c}_{ls}^{(12)}(j)}{\sqrt{\hat{c}_{ll}^{(11)}(0) \hat{c}_{ss}^{(22)}(0)}}$  pour  $1-N \leq j \leq N-1$ .

Dans ce qui suit, nous énonçons et montrons le théorème qui nous donne la distribution asymptotique du vecteur  $\mathbf{r}_{\hat{\mathbf{a}}}^{(12)}$  sous l'hypothèse de non corrélation des deux processus VAR( $\infty$ ).

### Théorème 2.3.1.

Soient  $\mathbf{X}^{(1)}$  et  $\mathbf{X}^{(2)}$  deux processus linéaires stationnaires au second ordre vérifiant le modèle VAR( $\infty$ ) décrit par l'équation (1.2.3). Les cumulants d'ordre quatre des processus d'innovations correspondants sont supposés nuls. Si ces deux processus sont indépendants et vérifient les hypothèses suivantes:

- (i)  $\mathbb{E}|a_{i,t} a_{j,t} a_{k,t} a_{l,t}| \leq \gamma_4 < \infty$ ,  $1 \leq i, j, k, l \leq m$ ;
- (ii)  $p_h$ ,  $h = 1, 2$ , sont deux suites d'entiers positifs telles que  $p_h$  est une fonction de  $N$  vérifiant  $p_h(N) \rightarrow \infty$  et  $p_h(N)^2/N \rightarrow 0$ , quand  $N \rightarrow \infty$ ;
- (iii)  $\sqrt{N} \sum_{j=p_h+1}^{\infty} \|\Phi_j^{(h)}\|_2 \rightarrow 0$ , quand  $N \rightarrow \infty$ .

Alors les deux vecteurs  $\sqrt{N} \mathbf{r}_{\hat{\mathbf{a}}}^{(12)}$  et  $\sqrt{N} \mathbf{r}_{\mathbf{a}}^{(12)}$  définis par (2.1.2) possèdent la même distribution asymptotique  $\mathbf{N}(\mathbf{0}, \mathbb{I}_n \otimes (\boldsymbol{\rho}_2 \otimes \boldsymbol{\rho}_1))$ .

#### PREUVE DU THÉORÈME 2.3.1.

Pour montrer le théorème 2.3.1, il suffit de montrer que  $\sqrt{N} \left( \mathbf{r}_{\hat{\mathbf{a}}}^{(12)} - \mathbf{r}_{\mathbf{a}}^{(12)} \right) \rightarrow \mathbf{0}$  en probabilité, ou encore, quand  $N \rightarrow \infty$ ,

$$\forall j \in \{j_1, \dots, j_m\}, \quad \sqrt{N} \left[ \mathbf{R}_{\hat{\mathbf{a}}}^{(12)}(j) - \mathbf{R}_{\mathbf{a}}^{(12)}(j) \right] \xrightarrow{p} \mathbf{0}. \quad (2.3.3)$$

D'après (1.1.15), on peut écrire la matrice de corrélation croisée en terme des matrices d'autocovariances et de covariance croisée, par

$$\mathbf{R}_{\mathbf{a}}^{(12)}(j) = \mathbf{D}_1\{c_{ii}^{-1/2}(0)\}\mathbf{C}_{\mathbf{a}}^{(12)}(j)\mathbf{D}_2\{c_{ii}^{-1/2}(0)\},$$

où  $\mathbf{D}_h$ ,  $h = 1, 2$ , est définie par (1.1.4). Pour simplifier la notation, on écrira  $\mathbf{D}_h\{c_{ii}^{-1/2}(0)\} = \mathbf{D}_h^{-1}$ . La matrice  $\mathbf{R}_{\hat{\mathbf{a}}}^{(12)}(j)$  admet une représentation similaire. Ainsi, on obtient

$$\begin{aligned} \mathbf{R}_{\hat{\mathbf{a}}}^{(12)}(j) - \mathbf{R}_{\mathbf{a}}^{(12)}(j) &= \hat{\mathbf{D}}_1^{-1} \left[ \mathbf{C}_{\hat{\mathbf{a}}}^{(12)}(j) - \mathbf{C}_{\mathbf{a}}^{(12)}(j) \right] \hat{\mathbf{D}}_2^{-1} \\ &+ \hat{\mathbf{D}}_1^{-1} \mathbf{D}_1 \mathbf{R}_{\mathbf{a}}^{(12)}(j) \mathbf{D}_2 \hat{\mathbf{D}}_2^{-1} - \mathbf{R}_{\mathbf{a}}^{(12)}(j). \end{aligned} \quad (2.3.4)$$

On sait que, sous les conditions du théorème 2.3.1,  $\sqrt{N}\mathbf{R}_{\mathbf{a}}^{(12)}(j)$  converge en distribution. Aussi, d'après Lütkepohl (1989), on a que  $\hat{\mathbf{D}}_i^{-1} - \mathbf{D}_i = O_p(N^{-1/2})$ ,  $i = 1, 2$ , et on peut donc écrire que  $\hat{\mathbf{D}}_1^{-1}\mathbf{D}_1 = \mathbb{I} + O_p(N^{-1/2})$ . Donc,  $\hat{\mathbf{D}}_1^{-1}\mathbf{D}_1$  et  $\hat{\mathbf{D}}_2^{-1}\mathbf{D}_2$  converge en probabilité vers la matrice identité. Ainsi, on déduit que

$$\sqrt{N} \left[ \hat{\mathbf{D}}_1^{-1} \mathbf{D}_1 \mathbf{R}_{\mathbf{a}}^{(12)}(j) \mathbf{D}_2 \hat{\mathbf{D}}_2^{-1} - \mathbf{R}_{\mathbf{a}}^{(12)}(j) \right] \xrightarrow{p} \mathbf{0}, \text{ quand } N \rightarrow \infty. \quad (2.3.5)$$

Puisque les matrices  $\hat{\mathbf{D}}_1^{-1}$  et  $\hat{\mathbf{D}}_2^{-1}$  convergent en probabilité, alors pour montrer la relation (2.3.3), il suffit de vérifier que  $\sqrt{N} \left( \mathbf{C}_{\hat{\mathbf{a}}}^{(12)}(j) - \mathbf{C}_{\mathbf{a}}^{(12)}(j) \right) \rightarrow \mathbf{0}$  en probabilité quand  $N \rightarrow \infty$ . Si on note par  $\hat{\boldsymbol{\eta}}_t = \hat{\mathbf{a}}_t^{(1)} - \mathbf{a}_t^{(1)}$  et  $\hat{\boldsymbol{\delta}}_t = \hat{\mathbf{a}}_t^{(2)} - \mathbf{a}_t^{(2)}$ , on écrit alors

$$\sqrt{N} \left( \mathbf{C}_{\hat{\mathbf{a}}}^{(12)}(j) - \mathbf{C}_{\mathbf{a}}^{(12)}(j) \right) = \mathcal{C}_1 + \mathcal{C}_2 + \mathcal{C}_3, \quad (2.3.6)$$

avec

$$\mathcal{C}_1 = N^{-1/2} \sum_{t=j+1}^N \hat{\boldsymbol{\eta}}_t \mathbf{a}_{t-j}^{(2)T},$$

$$\mathcal{C}_2 = N^{-1/2} \sum_{t=j+1}^N \mathbf{a}_t^{(1)} \hat{\boldsymbol{\delta}}_{t-j}^T,$$

$$\mathcal{C}_3 = N^{-1/2} \sum_{t=j+1}^N \hat{\boldsymbol{\eta}}_t \hat{\boldsymbol{\delta}}_{t-j}^T$$

Ainsi, il suffit de montrer que  $\mathcal{C}_i \rightarrow \mathbf{0}$  en probabilité,  $i = 1, 2, 3$ , quand  $N \rightarrow \infty$ .

En écrivant  $\hat{\boldsymbol{\eta}}_t$  en terme des paramètres du modèle et d'après les équations (1.3.1) à (1.3.4), on obtient

$$\hat{\boldsymbol{\eta}}_t = \left\{ \boldsymbol{\Phi}(p_1) - \hat{\boldsymbol{\Phi}}(p_1) \right\} \mathbf{X}_t^{(1)}(p_1) + \boldsymbol{\xi}_{t,p_1}, \quad (2.3.7)$$

avec

$$\boldsymbol{\xi}_{t,p_1} = \sum_{l=p_1+1}^{\infty} \boldsymbol{\Phi}_l^{(1)} \mathbf{X}_{t-l}^{(1)}, \quad (2.3.8)$$

qui représente le biais de l'approximation du processus  $\{\mathbf{X}_t^{(1)}\}$  par un processus autorégressif d'ordre fini  $p_1$ .

Par la suite, on remarque que  $\hat{\boldsymbol{\eta}}_t$  ainsi que la différence (2.3.6) dépendent de la différence entre les paramètres du modèle et leurs estimateurs. Aussi, de l'équation (1.3.6), nous avons

$$\|\hat{\boldsymbol{\Phi}}(p_h) - \boldsymbol{\Phi}(p_h)\|_2 = O_p \left( \frac{p_h^{1/2}}{N^{1/2}} \right). \quad (2.3.9)$$

$h = 1, 2$ . Ainsi, si on considère le premier terme  $\mathcal{C}_1$  de la somme dans l'équation (2.3.6), et par définition de  $\hat{\boldsymbol{\eta}}_t$  donnée par 2.3.7, on peut écrire

$$\begin{aligned} \|\mathcal{C}_1\|_2 &= \left\| N^{-1/2} \sum_{t=j+1}^N \hat{\boldsymbol{\eta}}_t \mathbf{a}_{t-j}^{(2)T} \right\|_2 \\ &= \left\| N^{-1/2} \sum_{t=j+1}^N \left\{ \boldsymbol{\Phi}(p_1) - \hat{\boldsymbol{\Phi}}(p_1) \right\} \mathbf{X}_t^{(1)}(p_1) \mathbf{a}_{t-j}^{(2)T} + N^{-1/2} \sum_{t=j+1}^N \boldsymbol{\xi}_t(p_1) \mathbf{a}_{t-j}^{(2)T} \right\|_2 \end{aligned}$$

En utilisant l'inégalité triangulaire et l'inégalité pour la norme d'un produit, on obtient

$$\|\mathcal{C}_1\|_2 \leq \|\Phi(p_1) - \hat{\Phi}(p_1)\|_2 \|N^{-1/2} \sum_{t=j+1}^N \mathbf{X}_t^{(1)}(p_1) \mathbf{a}_{t-j}^{(2)T}\|_2 + \|N^{-1/2} \sum_{t=j+1}^N \xi_t(p_1) \mathbf{a}_{t-j}^{(2)T}\|_2.$$

Notons par

$$\mathbb{A} = \mathbb{E} \|N^{-1/2} \sum_{t=j+1}^N \mathbf{X}_t^{(1)}(p_1) \mathbf{a}_{t-j}^{(2)T}\|_2^2. \quad (2.3.10)$$

Par définition de la norme euclidienne, on a

$$\begin{aligned} \mathbb{A} &= N^{-1} \mathbb{E} \{ \text{tr} [ (\sum_{t=j+1}^N \mathbf{X}_t^{(1)}(p_1) \mathbf{a}_{t-j}^{(2)T}) (\sum_{t=j+1}^N \mathbf{X}_t^{(1)}(p_1) \mathbf{a}_{t-j}^{(2)T})^T ] \} \\ &= N^{-1} \mathbb{E} \{ \text{tr} [ \sum_{t=j+1}^N \mathbf{X}_t^{(1)}(p_1) \mathbf{a}_{t-j}^{(2)T} \mathbf{a}_{t-j}^{(2)} \mathbf{X}_t^{(1)}(p_1)^T ] \} \\ &+ N^{-1} \mathbb{E} \{ \text{tr} [ \sum_{t=j+1}^N \sum_{\substack{t_1=j+1 \\ t_1 \neq t}}^N \mathbf{X}_t^{(1)}(p_1) \mathbf{a}_{t-j}^{(2)T} \mathbf{a}_{t_1-j}^{(2)} \mathbf{X}_{t_1}^{(1)}(p_1)^T ] \} \end{aligned}$$

En utilisant des propriétés usuelles de la trace d'une matrice et sous l'hypothèse d'indépendance des sous processus, on obtient

$$\begin{aligned} \mathbb{A} &= N^{-1} \text{tr} \{ \sum_{t=j+1}^N \mathbb{E}(\mathbf{a}_{t-j}^{(2)T} \mathbf{a}_{t-j}^{(2)}) \mathbb{E}(\mathbf{X}_{t-j}^{(1)}(p_1)^T \mathbf{X}_{t-j}^{(1)}(p_1)) \} \\ &+ N^{-1} \text{tr} \{ \sum_{t=j+1}^N \sum_{\substack{t_1=j+1 \\ t_1 \neq t}}^N \mathbb{E}(\mathbf{a}_{t-j}^{(2)T} \mathbf{a}_{t_1-j}^{(2)}) \mathbb{E}(\mathbf{X}_{t_1}^{(1)}(p_1)^T \mathbf{X}_t^{(1)}(p_1)) \} \end{aligned}$$

Puisque  $\mathbb{E}(\mathbf{a}_{t-j}^{(2)T} \mathbf{a}_{t_1-j}^{(2)}) = 0$  et

$$\mathbb{E}(\mathbf{X}_t^{(1)}(p_1)^T \mathbf{X}_t^{(1)}(p_1)) = \mathbb{E} \left( \sum_{i=1}^{p_1} \mathbf{X}_{t-i}^{(1)T} \mathbf{X}_{t-i}^{(1)} \right) \leq \Delta p_1, \quad (2.3.11)$$

nous obtenons alors

$$\mathbb{E} \|N^{-1/2} \sum_{t=j+1}^N \mathbf{X}_t^{(1)}(p_1) \mathbf{a}_{t-j}^{(2)T}\|_2^2 \leq \Delta p_1. \quad (2.3.12)$$

Pour la deuxième partie de  $\|\mathcal{C}_1\|_2$ , en effectuant le même développement de la norme que dans la première partie, on a

$$\mathbb{E}\|N^{-1/2} \sum_{t=j+1}^N \boldsymbol{\xi}_t(p_1) \mathbf{a}_{t-j}^{(2)T}\|_2^2 = N^{-1} \mathbb{E}[\text{tr}\{\sum_{t=j+1}^N \boldsymbol{\xi}_t(p_1)^T \boldsymbol{\xi}_t(p_1) \mathbf{a}_{t-j}^{(2)T} \mathbf{a}_{t-j}^{(2)}\}].$$

Aussi, et par définition de  $\boldsymbol{\xi}_t(p_1)$  donnée par (2.3.8), on note que

$$\begin{aligned} \mathbb{E}\|\boldsymbol{\xi}_t(p_1)\|_2^2 &= \mathbb{E}\left\|\sum_{l=p_1+1}^{\infty} \boldsymbol{\Phi}_l^{(1)} \mathbf{X}_{t-l}^{(1)}\right\|_2^2 \\ &= \mathbb{E}[\text{tr}\{(\sum_{l=p_1+1}^{\infty} \boldsymbol{\Phi}_l^{(1)} \mathbf{X}_{t-l}^{(1)}) (\sum_{l=p_1+1}^{\infty} \boldsymbol{\Phi}_l^{(1)} \mathbf{X}_{t-l}^{(1)})^T\}] \\ &= \sum_{l=p_1+1}^{\infty} \sum_{l_1=p_1+1}^{\infty} \text{tr}(\boldsymbol{\Phi}_l^{(1)} \mathbb{E}(\mathbf{X}_{t-l}^{(1)} \mathbf{X}_{t-l_1}^{(1)T}) \boldsymbol{\Phi}_{l_1}^{(1)T}). \end{aligned} \quad (2.3.13)$$

Puisque la suite matricielle  $\boldsymbol{\Gamma}_{\mathbf{X}}^{(11)}(l)$  est supposée absolument sommable, alors  $\mathbb{E}(\mathbf{X}_{t-l}^{(1)} \mathbf{X}_{t-l_1}^{(1)T}) \leq \Delta \mathbb{I}_{m_1}$ . Aussi, d'après l'inégalité de Cauchy-Schwarz, on a

$$|\text{tr}(\boldsymbol{\Phi}_l^{(1)} \boldsymbol{\Phi}_{l_1}^{(1)T})| = |\langle \boldsymbol{\Phi}_l^{(1)}, \boldsymbol{\Phi}_{l_1}^{(1)} \rangle| \leq \|\boldsymbol{\Phi}_l^{(1)}\|_2 \|\boldsymbol{\Phi}_{l_1}^{(1)}\|_2.$$

On obtient donc

$$\begin{aligned} \mathbb{E}\|\boldsymbol{\xi}_t(p_1)\|_2^2 &\leq \Delta \sum_{l=p_1+1}^{\infty} \sum_{l_1=p_1+1}^{\infty} \|\boldsymbol{\Phi}_l^{(1)}\|_2 \|\boldsymbol{\Phi}_{l_1}^{(1)}\|_2 \\ &\leq \Delta \left(\sum_{l=p_1+1}^{\infty} \|\boldsymbol{\Phi}_l^{(1)}\|_2\right)^2. \end{aligned} \quad (2.3.14)$$

Ainsi, sous l'hypothèse d'indépendance des deux sous processus, on déduit que

$$\begin{aligned} \mathbb{E}\|N^{-1/2} \sum_{t=j+1}^N \boldsymbol{\xi}_t(p_1) \mathbf{a}_{t-j}^{(2)T}\|_2^2 &\leq \Delta \left(\sum_{l=p_1+1}^{\infty} \|\boldsymbol{\Phi}_l^{(1)}\|_2\right)^2 \\ &= O\left(\frac{1}{N}\right). \end{aligned} \quad (2.3.15)$$

Les résultats (2.3.9), (2.3.12) et (2.3.15), ainsi que les hypothèses (ii) et (iii) nous permettent de déduire que

$$\|\mathcal{C}_1\|_2 = O_p\left(\frac{p_1^{1/2}}{N^{1/2}}\right)O_p(p_1^{1/2}) + O_p(N^{-1/2}) = O_p\left(\frac{p_1}{N^{1/2}}\right), \quad (2.3.16)$$

et donc,  $\|\mathcal{C}_1\|_2$  converge en probabilité vers 0. Par symétrie, on peut obtenir un résultat similaire pour le deuxième terme  $\mathcal{C}_2$  de la somme (2.3.6), et on a que  $\|\mathcal{C}_2\|_2$  converge en probabilité vers 0 quand  $N \rightarrow \infty$ .

Finalement, pour le troisième terme  $\mathcal{C}_3$  de la somme (2.3.6), compte tenu des hypothèses d'indépendance des deux processus, des hypothèses sur les ordres d'autorégressions  $p_h$ ,  $h = 1, 2$ , et les hypothèses sur les paramètres des deux processus, on peut écrire

$$\begin{aligned} \|\mathcal{C}_3\|_2 &= \left\| N^{-1/2} \sum_{t=j+1}^N \hat{\eta}_t \hat{\delta}_{t-j}^T \right\|_2 \\ &\leq \|\Phi(p_1) - \hat{\Phi}(p_1)\|_2 \|\Phi(p_2) - \hat{\Phi}(p_2)\|_2 \left\| N^{-1/2} \sum_{t=j+1}^N \mathbf{X}_t^{(1)}(p_1) \mathbf{X}_{t-j}^{(2)}(p_2)^T \right\|_2 \\ &+ \|\Phi(p_1) - \hat{\Phi}(p_1)\|_2 \left\| N^{-1/2} \sum_{t=j+1}^N \mathbf{X}_t^{(1)}(p_1) \boldsymbol{\xi}_{t-j}(p_2)^T \right\|_2 \\ &+ \|\Phi(p_2) - \hat{\Phi}(p_2)\|_2 \left\| N^{-1/2} \sum_{t=j+1}^N \boldsymbol{\xi}_t(p_1) \mathbf{X}_{t-j}^{(2)}(p_2)^T \right\|_2 \\ &+ \left\| N^{-1/2} \sum_{t=j+1}^N \boldsymbol{\xi}_t(p_1) \boldsymbol{\xi}_{t-j}(p_2)^T \right\|_2. \end{aligned} \quad (2.3.17)$$

Pour la première partie de cette inégalité, notons

$$\mathbf{V}_1 = \mathbb{E} \left\| N^{-1/2} \sum_{t=j+1}^N \mathbf{X}_t^{(1)}(p_1) \mathbf{X}_{t-j}^{(2)}(p_2)^T \right\|_2^2.$$

En utilisant l'équation (2.3.11), on obtient que  $\mathbb{E}(\mathbf{X}_t^{(h)}(p_h)^T \mathbf{X}_t^{(h)}(p_h)) \leq \Delta p_h$  pour  $h = 1, 2$ , et en utilisant des résultats de la preuve du théorème 1 de Lewis et

Reinsel (1985), on déduit que

$$\mathbf{V}_1 = \mathbb{E} \left\| N^{-1/2} \sum_{t=j+1}^N \mathbf{X}_t^{(1)}(p_1) \mathbf{X}_{t-j}^{(2)}(p_2)^T \right\|_2^2 \leq \Delta p_1 p_2. \quad (2.3.18)$$

Pour le deuxième terme dans le membre de droite de (2.3.17), notons

$$\mathbf{V}_2 = \mathbb{E} \left\| N^{-1/2} \sum_{t=j+1}^N \mathbf{X}_t^{(1)}(p_1) \boldsymbol{\xi}_{t-j}(p_2)^T \right\|_2^2.$$

Si on note par  $\mathbf{Y}_t = \mathbf{X}_t^{(1)}(p_1) \boldsymbol{\xi}_{t-j}(p_2)^T$ , on remarque que sous l'hypothèse d'indépendance des deux sous processus, on a

$$\begin{aligned} \text{tr}(\mathbb{E}(\mathbf{Y}_t \mathbf{Y}_t^T)) &= \text{tr}(\mathbb{E}(\mathbf{X}_t^{(1)}(p_1) \boldsymbol{\xi}_{t-j}(p_2)^T \boldsymbol{\xi}_{t-j}(p_2) \mathbf{X}_t^{(1)}(p_1)^T)) \\ &= \mathbb{E}(\|\mathbf{X}_t^{(1)}(p_1)\|_2^2) \mathbb{E}(\|\boldsymbol{\xi}_{t-j}(p_2)\|_2^2). \end{aligned}$$

Et d'après (2.3.14) et (2.3.11), on a

$$\mathbb{E}(\mathbf{Y}_t \mathbf{Y}_t^T) \leq \Delta p_1 \left( \sum_{l=p_2+1}^{\infty} \|\Phi_l^{(2)}\|_2 \right)^2. \quad (2.3.19)$$

En effectuant des développements similaires à ceux pour  $\mathbf{A}$  dans (2.3.10), et en appliquant la relation (2.3.19) et l'inégalité de Cauchy-Schwarz, on obtient alors

$$\begin{aligned} \mathbf{V}_2 &= N^{-1} \mathbb{E} \left\{ \text{tr} \left[ \left( \sum_{t=j+1}^N \mathbf{X}_t^{(1)}(p_1) \boldsymbol{\xi}_{t-j}^{(2)}(p_2)^T \right) \left( \sum_{t=j+1}^N \mathbf{X}_t^{(1)}(p_1) \boldsymbol{\xi}_{t-j}^{(2)}(p_2)^T \right)^T \right] \right\} \\ &\leq \Delta p_1 \left( \sum_{l=p_2+1}^{\infty} \|\Phi_l^{(2)}\|_2 \right)^2. \end{aligned} \quad (2.3.20)$$

Par symétrie, on peut montrer aussi que

$$\mathbb{E} \left\| N^{-1/2} \sum_{t=j+1}^N \boldsymbol{\xi}_t(p_1) \mathbf{X}_{t-j}^{(2)}(p_2)^T \right\|_2^2 \leq \Delta p_2 \left( \sum_{l=p_1+1}^{\infty} \|\Phi_l^{(1)}\|_2 \right)^2. \quad (2.3.21)$$

Notons par  $\mathbf{V}_3$  le dernier terme du membre de droite de (2.3.17).

$$\mathbf{V}_3 = \mathbb{E} \left\| N^{-1/2} \sum_{t=j+1}^N \boldsymbol{\xi}_t(p_1) \boldsymbol{\xi}_{t-j}(p_2)^T \right\|_2^2.$$

Sous l'hypothèse d'indépendance, en utilisant les mêmes techniques que celles de  $V_2$ , et appliquant le résultat (2.3.13), on obtient

$$\begin{aligned} V_3 &= N^{-1} \mathbb{E} \left\{ \text{tr} \left[ \left( \sum_{t=j+1}^N \boldsymbol{\xi}_t(p_1) \boldsymbol{\xi}_{t-j}(p_2)^T \right) \left( \sum_{t=j+1}^N \boldsymbol{\xi}_t(p_1) \boldsymbol{\xi}_{t-j}(p_2)^T \right)^T \right] \right\} \\ &\leq \Delta \left( \sum_{l=p_1+1}^{\infty} \|\boldsymbol{\Phi}_l^{(1)}\|_2 \right)^2 \left( \sum_{l=p_2+1}^{\infty} \|\boldsymbol{\Phi}_l^{(2)}\|_2 \right)^2. \end{aligned} \quad (2.3.22)$$

Et ainsi, en combinant les résultats (2.3.18), (2.3.20), (2.3.21) et (2.3.22), on obtient

$$\begin{aligned} \|\mathcal{C}_3\|_2 &= O_p\left(\frac{p_1^{1/2}}{N^{1/2}}\right) O_p\left(\frac{p_2^{1/2}}{N^{1/2}}\right) O_p(p_1^{1/2} p_2^{1/2}) \\ &+ O_p\left(\frac{p_1^{1/2}}{N^{1/2}}\right) O_p(p_1^{1/2}) \sum_{l=p_2+1}^{\infty} \|\boldsymbol{\Phi}_l^{(2)}\|_2 + O_p\left(\frac{p_2^{1/2}}{N^{1/2}}\right) O_p(p_2^{1/2}) \sum_{l=p_1+1}^{\infty} \|\boldsymbol{\Phi}_l^{(1)}\|_2 \\ &+ O_p\left(\sum_{l=p_1+1}^{\infty} \|\boldsymbol{\Phi}_l^{(1)}\|_2\right) \sum_{l=p_2+1}^{\infty} \|\boldsymbol{\Phi}_l^{(2)}\|_2, \end{aligned}$$

et en utilisant les hypothèses (ii) et (iii), on a

$$\|\mathcal{C}_3\|_2 \xrightarrow{P} 0.$$

Ainsi, on a montré que  $\sqrt{N} \left( \mathbf{C}_{\hat{\mathbf{a}}}^{(12)}(j) - \mathbf{C}_{\mathbf{a}}^{(12)}(j) \right) \rightarrow \mathbf{0}$  en probabilité quand  $N \rightarrow \infty$ , et par suite on a montré (2.3.3).  $\square$

### 2.3.1. Procédures des tests d'indépendance

Le théorème 2.3.1 nous permet, dans le cas de deux séries multivariées issues de la classe des processus  $\text{VAR}(\infty)$ , de définir une statistique de test de non corrélation entre deux processus  $\mathbf{a}^{(1)}$  et  $\mathbf{a}^{(2)}$ . Puisque  $\hat{\boldsymbol{\rho}}_h = \mathbf{R}_{\hat{\mathbf{a}}}^{(hh)}(0)$ ,  $h = 1, 2$ , est un estimateur convergeant en probabilité de  $\boldsymbol{\rho}_h$ , nous définissons la statistique suivante

$$Q_{\hat{\mathbf{a}}}^{(12)}(j) = N \mathbf{r}_{\hat{\mathbf{a}}}^{(12)}(j)^T \left( \hat{\boldsymbol{\rho}}_2^{-1} \otimes \hat{\boldsymbol{\rho}}_1^{-1} \right) \mathbf{r}_{\hat{\mathbf{a}}}^{(12)}(j), \quad (2.3.23)$$

qui est asymptotiquement distribuée comme une variable du khi-deux à  $m_1 m_2$  degrés de liberté. Ainsi, pour tester, au niveau de signification  $\alpha$  l'hypothèse de non corrélation à un délai  $j$  fixé quelconque,  $\boldsymbol{\rho}_{\mathbf{a}}^{(12)}(j) = \mathbf{0}$ , entre les deux processus  $\mathbf{a}^{(1)}$  et  $\mathbf{a}^{(2)}$ , la statistique  $Q_{\hat{\mathbf{a}}}^{(12)}(j)$  peut être comparée au  $\alpha$ -fractile de la distribution asymptotique précédente.

De même, ce théorème 2.3.1 peut nous permettre aussi de déduire un test global de non corrélation dans le cas de deux séries VAR( $\infty$ ). Comme a été mentionné dans El Himdi et Roy (1997), le choix de la suite  $j = -M, \dots, M$ , où  $M$  prend une valeur indépendante de  $N$ , dans la définition du vecteur  $\mathbf{r}_{\hat{\mathbf{a}}}^{(12)}$ , semble raisonnable dans le cas où aucune direction de causalité n'est à priori soupçonnée. Pour un tel choix de délais et pour une réalisation de longueur  $N$  des deux séries multivariées, la statistique  $Q_{\hat{\mathbf{a}},M}^{(12)}$  définie par:

$$\begin{aligned} Q_{\hat{\mathbf{a}},M}^{(12)} &= N \mathbf{r}_{\hat{\mathbf{a}}}^{(12)T} (\mathbf{I}_{2M+1} \otimes \hat{\boldsymbol{\rho}}_2^{-1} \otimes \hat{\boldsymbol{\rho}}_1^{-1}) \mathbf{r}_{\hat{\mathbf{a}}}^{(12)} \\ &= \sum_{j=-M}^M Q_{\hat{\mathbf{a}}}^{(12)}(j), \end{aligned} \quad (2.3.24)$$

suit asymptotiquement une loi khi-deux à  $(2M + 1)m_1 m_2$  degrés de liberté.

La statistique  $Q_{\hat{\mathbf{a}},M}^{(12)}$  peut aussi s'exprimer en termes des deux autocovariances  $\mathbf{C}_{\hat{\mathbf{a}}}^{(hh)}(j)$ ,  $h = 1, 2$ , et des covariances croisées  $\mathbf{C}_{\hat{\mathbf{a}}}^{(12)}(j)$  des séries résiduelles  $\{\hat{\mathbf{a}}_t^{(1)}\}$  et  $\{\hat{\mathbf{a}}_t^{(2)}\}$ . D'après le lemme 4.1 d'El Himdi et Roy (1997), on peut écrire

$$Q_{\hat{\mathbf{a}},M}^{(12)} = N \sum_{j=-M}^M \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)^T \left( \mathbf{C}_{\hat{\mathbf{a}}}^{(22)}(0)^{-1} \otimes \mathbf{C}_{\hat{\mathbf{a}}}^{(11)}(0)^{-1} \right) \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j). \quad (2.3.25)$$

L'hypothèse d'indépendance des deux séries VAR( $\infty$ ) sera ainsi rejetée pour les grandes valeurs de  $Q_{\hat{\mathbf{a}},M}^{(12)}$  comparées aux valeurs critiques issues de la distribution du khi-deux. Aussi, dans le but de corriger la variance asymptotique de

L'hypothèse d'indépendance des deux séries VAR( $\infty$ ) sera ainsi rejetée pour les grandes valeurs de  $Q_{\hat{\mathbf{a}},M}^{(12)}$  comparées aux valeurs critiques issues de la distribution du khi-deux. Aussi, dans le but de corriger la variance asymptotique de chaque composante  $\hat{c}_{uv}^{(12)}(j)$  du vecteur  $\mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)$ , la statistique  $Q_{\hat{\mathbf{a}},M}^{(12)}$  peut être modifiée en utilisant une correction analogue à celle introduite par Haugh (1976), El Himdi et Roy (1997) et aussi par Hosking (1980) concernant le test portemanteau de bruit blanc. La statistique modifiée  $Q_{\hat{\mathbf{a}},M}^{(12)*}$  est définie par:

$$Q_{\hat{\mathbf{a}},M}^{(12)*} = \sum_{j=-M}^M Q_{\hat{\mathbf{a}}}^{(12)}(j)^*, \quad (2.3.26)$$

avec

$$Q_{\hat{\mathbf{a}}}^{(12)}(j)^* = \frac{N}{N - |j|} Q_{\hat{\mathbf{a}}}^{(12)}(j).$$

Cette statistique modifiée suit, sous l'hypothèse de non corrélation, la même distribution que  $Q_{\hat{\mathbf{a}},M}^{(12)}$ . El Himdi et Roy (1997) ont illustré par simulation que la statistique modifiée  $Q_{\hat{\mathbf{a}},M}^{(12)*}$  est mieux approximée par la distribution de khi-deux pour de courtes séries que la statistique  $Q_{\hat{\mathbf{a}},M}^{(12)}$ , et particulièrement pour les valeurs de  $M$  relativement grandes. Ainsi, dans notre étude de simulation au chapitre 5, nous n'étudions que la distribution de la statistique modifiée  $Q_{\hat{\mathbf{a}},M}^{(12)*}$ .

## Chapitre 3

---

### TESTS D'INDÉPENDANCE BASÉS SUR TOUS LES DÉLAIS

Hong (1996) a généralisé l'approche de Haugh en considérant deux processus linéaires univariés stationnaires, qui admettent une représentation  $AR(\infty)$  stationnaire. Il approxime ces derniers par des processus autorégressifs d'ordres finis dont l'ordre dépend de la taille échantionnale, voir Berk (1974). La statistique du test, basée sur tous les délais, est définie comme une combinaison linéaire pondérée par des corrélations croisées résiduelles. Les poids sont déterminés par un noyau et une suite de points de troncature. La distribution asymptotique de cette statistique suit une loi normale sous l'hypothèse nulle contrairement à celle de Haugh qui suit une loi khi-carré.

Dans ce chapitre, nous nous proposons de généraliser l'approche de Hong (1996) au cas de deux processus linéaires multivariés stationnaires qui ont une représentation  $VAR(\infty)$ . Nous allons tout d'abord rappeler quelques résultats univariés concernant la distribution asymptotique de la statistique de test. Ensuite, dans le cas multivarié, nous montrons que la distribution de la statistique du test, basée sur tous les délais, sous l'hypothèse d'indépendance entre les deux processus multivariés  $VAR(\infty)$  suit aussi une loi normale.

### 3.1. APPROCHE DE HONG POUR DEUX PROCESSUS UNIVARIÉS AR( $\infty$ )

Considérons les deux processus linéaires suivants

$$X_t^{(h)} = \sum_{j=0}^{\infty} \phi_j^{(h)} a_{t-j}^{(h)}, \quad t \in \mathbb{Z}; \quad h = 1, 2, \quad (3.1.1)$$

où  $a_t^{(h)}$  sont deux bruits blancs de moyennes 0 et de variances  $\sigma_h^2$  et tels que  $\mathbb{E}(a_t^{(h)4}) < \infty$ . Les coefficients  $\phi_j^{(h)}$  sont tels que  $\sum_{j=0}^{\infty} |\phi_j^{(h)}| < \infty$  avec  $\phi_0^{(h)} = 1$  et satisfont les conditions d'inversibilité. Pour  $h = 1, 2$ , Ces deux processus ont une représentation autorégressive d'ordre infini AR( $\infty$ ), qu'on écrit

$$\Phi^{(h)}(B)X_t^{(h)} = a_t^{(h)} \quad ; \quad t \in \mathbb{Z},$$

avec

$$\Phi^{(h)}(B) = 1 - \sum_{j=1}^{\infty} \phi_j^{(h)} B^j = \left( \sum_{j=0}^{\infty} \phi_j^{(h)} B^j \right)^{-1},$$

où  $B$  désigne l'opérateur retard. À partir de réalisations de longueur  $N$ , nous approximations chacun des deux processus par une autorégression d'ordre fini  $p_h$  pour le processus  $X^{(h)}$ . Les séries résiduelles basées sur les estimateurs des moindres carrés des paramètres sont données par

$$\hat{a}_t^{(h)} = X_t^{(h)} - \hat{\phi}^{(h)}(p_h)^T X_t^{(h)}(p_h),$$

avec  $X_t^{(h)}(p_h) = (X_{t-1}^{(h)}, \dots, X_{t-p_h}^{(h)})^T$  et

$$\hat{\phi}^{(h)}(p_h) = \left\{ \sum_{t=p_h+1}^N X_t^{(h)}(p_h) X_t^{(h)}(p_h)^T \right\}^{-1} \sum_{t=p_h+1}^N X_t^{(h)}(p_h) X_t^{(h)}.$$

La fonction de corrélation croisée résiduelle est définie par

$$r_{\hat{a}}^{(12)}(j) = \frac{c_{\hat{a}}^{(12)}(j)}{\{c_{\hat{a}}^{(11)}(0)\hat{c}_{\hat{a}}^{(22)}(0)\}^{1/2}}$$

La fonction de corrélation croisée résiduelle est définie par

$$r_{\hat{a}}^{(12)}(j) = \frac{c_{\hat{a}}^{(12)}(j)}{\{c_{\hat{a}}^{(11)}(0)c_{\hat{a}}^{(22)}(0)\}^{1/2}}$$

où  $c_{\hat{a}}^{(12)}(j)$  désigne la fonction de covariance croisée résiduelle donnée par

$$c_{\hat{a}}^{(12)}(j) = \begin{cases} N^{-1} \sum_{t=j+1}^N \hat{a}_t^{(1)} \hat{a}_{t-j}^{(2)} & \text{si } j \geq 0, \\ N^{-1} \sum_{t=-j+1}^N \hat{a}_{t+j}^{(1)} \hat{a}_t^{(2)} & \text{si } j < 0, \end{cases}$$

$$c_{\hat{a}}^{(hh)}(0) = N^{-1} \sum_{t=1}^N \hat{a}_t^{(h)2}.$$

Hong (1996) a proposé, sous une approche semi-paramétrique, un test unilatéral d'indépendance entre deux séries univariées d'ordre infini. La statistique de test  $Q_N$  proposée fait intervenir les corrélations croisées à tous les délais, pondérées par un noyau qui vérifie l'hypothèse suivante. On désigne par  $\mathbb{L}^2(\mathbb{R})$  l'espace des fonctions réelles et qui sont de carrés intégrables.

### Hypothèse 3.1.1.

Les noyaux  $k$  sont dans la classe  $\mathcal{K}_1$  suivante

$$\mathcal{K}_1 = \{k : \mathbb{R} \rightarrow [-1, 1]; k(0) = 1, k(x) = k(-x) \forall x \in \mathbb{R}, k \in \mathbb{L}^2(\mathbb{R}),$$

$k(\cdot)$  est continue au point 0 et admet au plus un ensemble fini de points de discontinuité}.

Des exemples de noyaux appartenant à la classe  $\mathcal{K}_1$  qu'on utilisera pour des fins de simulation sont:

Noyau uniforme tronqué (TR):

$$k(z) = \begin{cases} 1 & |z| \leq 1, \\ 0 & \text{sinon;} \end{cases} \quad (3.1.2)$$

Noyau de Bartlett (BAR):

$$k(z) = \begin{cases} 1 - |z| & |z| \leq 1, \\ 0 & \text{sinon;} \end{cases} \quad (3.1.3)$$

Noyau de Daniell (DAN):

$$k(z) = \frac{\sin(\pi z)}{\pi z}, \quad z \in \mathbb{R}; \quad (3.1.4)$$

Noyau de Parzen (PAR):

$$k(z) = \begin{cases} 1 - 6z^2 + 6|z|^3 & \text{si } |z| \leq 0.5, \\ 2(1 - |z|)^3 & \text{si } 0.5 \leq |z| \leq 1, \\ 0 & \text{sinon;} \end{cases} \quad (3.1.5)$$

Noyau de Bartlett-Priestley (BP):

$$k(z) = \frac{3}{(\pi z)^2} \left\{ \frac{\sin(\pi z)}{\pi z} - \cos(\pi z) \right\}, \quad z \in \mathbb{R}; \quad (3.1.6)$$

Noyau de Tukey-Hamming (TH):

$$k(z) = \begin{cases} \frac{1}{2}(1 + \cos(\pi z)) & \text{si } |z| \leq 1, \\ 0 & \text{sinon.} \end{cases} \quad (3.1.7)$$

Pour d'autres noyaux appartenant aussi à  $\mathcal{K}_1$  et leurs propriétés, voir Priestley (1981, p. 441).

On définit aussi les deux quantités suivantes:

$$S_N(k) = \sum_{j=1-N}^{N-1} (1 - |j|/N)k^2(j/M), \quad (3.1.8)$$

$$D_N(k) = \sum_{j=2-N}^{N-2} (1 - |j|/N)(1 - (|j| + 1)/N)k^4(j/M). \quad (3.1.9)$$

La statistique  $\mathcal{Q}_N$  est donnée par:

$$\mathcal{Q}_N = \frac{N \sum_{j=1-N}^{N-1} k^2 (j/M) r_{\hat{a}}^{(12)}(j)^2 - S_N(k)}{\{2D_N(k)\}^{1/2}},$$

$M$  est un paramètre de troncature ou de lissage tel que  $M = M(N) \rightarrow \infty$  et  $M/N \rightarrow 0$  quand  $N \rightarrow \infty$ .  $S_N(k)$  et  $D_N(k)$  sont données par les relations (3.1.8) et (3.1.9). Ainsi, Hong (1996) a établi la distribution asymptotique de la statistique  $\mathcal{Q}_N$  sous l'hypothèse nulle d'indépendance, sous des conditions de régularités sur les paramètres des processus et sur les ordres des deux autorégressions. On a le résultat suivant.

**Théorème 3.1.1.**

Soient  $X^{(h)}$ ,  $h = 1, 2$ , deux processus linéaires vérifiant l'équation (3.1.1). Soient  $M \rightarrow \infty$  tel que  $M/N \rightarrow 0$ ,  $p_h = p_h(N)$  deux suites d'entiers positifs qui vérifient

$$p_h = o\left(\frac{N^{1/2}}{M^{1/4}}\right), \quad N \sum_{j=p_h+1}^{\infty} \phi_j^{(h)2} = o\left(\frac{N^{1/2}}{M^{1/4}}\right).$$

Supposons que l'hypothèse 3.1.1 est satisfaite. Si les deux processus  $X^{(h)}$  sont indépendants alors  $\mathcal{Q}_N \rightarrow \mathcal{N}(0, 1)$  en distribution.

## 3.2. GÉNÉRALISATION AU CAS DE DEUX PROCESSUS MULTIVARIÉS

### VAR( $\infty$ )

#### 3.2.1. Définitions et généralités

Dans ce qui suit, on suppose que pour  $h = 1, 2$ ,  $\{\mathbf{X}_t^{(h)}\}$  sont deux sous processus autorégressifs multivariés d'ordre infini qui vérifient l'équation aux différences (1.2.3). La série des coefficients autorégressifs  $(\Phi_t^{(h)})_{t=0}^{\infty}$  vérifie la condition de stationnarité (1.2.4) et est supposée absolument sommable. Le processus  $\mathbf{a} = \left\{ \left( \mathbf{a}_t^{(1)T}, \mathbf{a}_t^{(2)T} \right)^T, t \in \mathbb{Z} \right\}$  est un bruit blanc de matrice de covariance  $\Sigma$

qui admet la même partition que (2.2.2). En plus, les deux sous processus  $\{\mathbf{a}_t^{(h)}\}$  vérifient l'hypothèse 2.3.1.

Dans cette version de Hong généralisée, la statistique du test d'indépendance entre les deux processus multivariés  $\mathbf{X}^{(1)}$  et  $\mathbf{X}^{(2)}$  est basée sur les corrélations croisées entre les séries résiduelles  $\{\hat{\mathbf{a}}_t^{(h)}\}$  obtenues en approximant les deux sous processus VAR( $\infty$ ) par des autorégressions d'ordre fini. Les délais sont pondérés à l'aide d'un noyau qui vérifie l'hypothèse 3.1.1

Si on note par  $\mathbf{r}_{\hat{\mathbf{a}}}^{(12)}(j) = \text{vec}(\mathbf{R}_{\hat{\mathbf{a}}}^{(12)}(j))$ , la version multivariée de la statistique de Hong pour tester la non corrélation des deux processus  $\mathbf{X}^{(1)}$  et  $\mathbf{X}^{(2)}$  est basée sur la forme quadratique suivante:

$$\mathcal{T}(\hat{\mathbf{a}}, \hat{\Sigma}) = \sum_{j=1-N}^{N-1} k^2(j/M) Q_{\hat{\mathbf{a}}}^{(12)}(j) \quad (3.2.1)$$

où

$$Q_{\hat{\mathbf{a}}}^{(12)}(j) = N \mathbf{r}_{\hat{\mathbf{a}}}^{(12)}(j)^T \left( \mathbf{R}_{\hat{\mathbf{a}}}^{(22)}(0)^{-1} \otimes \mathbf{R}_{\hat{\mathbf{a}}}^{(11)}(0)^{-1} \right) \mathbf{r}_{\hat{\mathbf{a}}}^{(12)}(j). \quad (3.2.2)$$

La statistique centrée et réduite correspondante est la suivante:

$$\mathcal{Q}_N = \frac{\mathcal{T}(\hat{\mathbf{a}}, \hat{\Sigma}) - m_1 m_2 S_N(k)}{\sqrt{2m_1 m_2 D_N(k)}}. \quad (3.2.3)$$

Comme a été mentioné dans la section 2.3.1, cette statistique peut aussi s'exprimer en terme des autocovariances  $\mathbf{C}_{\hat{\mathbf{a}}}^{(hh)}(0)$  et des covariances croisées  $\mathbf{C}_{\hat{\mathbf{a}}}^{(12)}(j)$  des mêmes séries résiduelles. Ainsi, d'après le lemme 4.1 d'El Himdi et Roy (1997), la forme quadratique  $\mathcal{T}(\hat{\mathbf{a}}, \hat{\Sigma})$  définie par 3.2.1 s'écrira aussi sous la forme suivante:

$$\mathcal{T}(\hat{\mathbf{a}}, \hat{\Sigma}) = N \sum_{j=1-N}^{N-1} k^2(j/M) \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)^T \left( \mathbf{C}_{\hat{\mathbf{a}}}^{(22)}(0)^{-1} \otimes \mathbf{C}_{\hat{\mathbf{a}}}^{(11)}(0)^{-1} \right) \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)$$

avec  $\mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j) = \text{vec} \left( \mathbf{C}_{\hat{\mathbf{a}}}^{(12)}(j) \right)$ .

Aussi, on définit par

$$\mathcal{T}(\mathbf{a}, \boldsymbol{\Sigma}) = N \sum_{j=1-N}^{N-1} k^2(j/M) \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)^T (\boldsymbol{\Sigma}_2^{-1} \otimes \boldsymbol{\Sigma}_1^{-1}) \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j),$$

où  $\mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j) = \text{vec} \left( \mathbf{C}_{\hat{\mathbf{a}}}^{(12)}(j) \right)$ ,  $\mathbf{C}_{\hat{\mathbf{a}}}^{(12)}(j) = \left( c_{uv}^{(12)}(j) \right)_{m_1 \times m_2}$  est défini de la même manière que  $\mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)$  en remplaçant les séries résiduelles  $\left( \hat{\mathbf{a}}_t^{(1)}, \hat{\mathbf{a}}_t^{(2)} \right)_{t=1}^N$  par les séries innovations  $\left( \mathbf{a}_t^{(1)}, \mathbf{a}_t^{(2)} \right)_{t=1}^N$ .

Dans la classe de noyaux  $\mathcal{K}_1$ , les conditions  $k(0) = 1$  et que le noyau  $k(\cdot)$  est continue au point 0 reflètent que pour un  $j$  petit par rapport à  $M$ , le poids attribué à  $\mathbf{r}_{\hat{\mathbf{a}}}^{(12)}(j)$  sera près de 1. Les deux quantités  $S_N(k)$  et  $D_N(k)$ , définies par (3.1.8) et (3.1.9), sont approximativement proportionnelles à la moyenne et la variance respectivement de  $\mathcal{T}(\hat{\mathbf{a}}, \hat{\boldsymbol{\Sigma}})$ . Les facteurs  $(1 - |j|/N)$  et  $(1 - (|j| + 1)/N)$  peuvent être considérés comme des corrections échantillonnales finies et ils sont asymptotiquement négligeables. Le paramètre de troncature  $M$  est tel que:  $M = M(N) \rightarrow \infty$ ,  $M/N \rightarrow 0$ .

Sous ces conditions sur le paramètre  $M$  et pour un noyau  $k$  dans  $\mathcal{K}_1$ , on peut vérifier que

$$M^{-1} S_N(k) \rightarrow S(k) \text{ et } M^{-1} D_N(k) \rightarrow D(k), \quad (3.2.4)$$

avec

$$S(k) = \int_{-\infty}^{+\infty} k^2(z) dz, \quad D(k) = \int_{-\infty}^{+\infty} k^4(z) dz. \quad (3.2.5)$$

Ainsi, on peut définir une version alternative de la statistique  $\mathcal{Q}_N$ , qu'on note  $\mathcal{Q}_N^*$ , qui prend en considération les convergences ci-dessus des facteurs de correction

$S_N(k)$  et  $D_N(k)$  vers  $S(k)$  et  $D(k)$  respectivement. On définit alors

$$\mathcal{Q}_N^* = \frac{\mathcal{T}(\hat{\mathbf{a}}, \hat{\Sigma}) - Mm_1m_2S(k)}{\sqrt{2Mm_1m_2D(k)}}.$$

Les deux statistiques  $\mathcal{Q}_N$  et  $\mathcal{Q}_N^*$  ont les mêmes distributions asymptotiques, et les mêmes propriétés de puissance.

El Himdi et Roy (1997) ont proposé la statistique

$$Q_{\hat{\mathbf{a}}, M}^{(12)} = \sum_{j=-M}^M Q_{\hat{\mathbf{a}}}^{(12)}(j),$$

où  $M$  est fixé à priori, pour tester la non corrélation entre les deux processus  $\{\mathbf{X}_t^{(1)}\}$  et  $\{\mathbf{X}_t^{(2)}\}$ , généralisant ainsi, au cas multivarié, celle de Haugh (1976). Asymptotiquement, cette statistique suit une loi  $\chi_{(2M+1)m_1m_2}^2$ .

En éliminant les facteurs de standardisation  $S_N(k)$  et  $D_N(k)$ , la statistique  $QH_M$  peut être vue comme un cas particulier de la statistique  $\mathcal{Q}_N$  définie par (3.2.3), avec le choix du noyau uniforme tronqué défini par (3.1.2). Comme on va le voir plus loin, plusieurs choix de noyaux donneront des résultats meilleurs que ceux du test d'El Himdi et Roy (1997). D'autre part, les résidus  $\{\hat{\mathbf{a}}_t^{(1)}\}$  et  $\{\hat{\mathbf{a}}_t^{(2)}\}$  utilisés par El Himdi et Roy (1997) sont obtenus en estimant les vrais modèles ARMA( $p, q$ ). Comme il a été souligné par Haugh (1976), cette approche est de nature quelque peu paramétrique puisqu'en pratique les vrais modèles ne sont pas connus et des erreurs de spécification sont possibles. Aussi, une spécification erronée des modèles servant à décrire chacune des deux séries peut mener à de fausses conclusions puisqu'elle risque d'infirmar la distribution asymptotique de la statistique du test.

Dans notre cas, on approxime  $\{\mathbf{X}_t^{(1)}\}$  et  $\{\mathbf{X}_t^{(2)}\}$  par deux autorégressions d'ordres finis  $p_h$ ,  $h = 1, 2$ . Les ordres  $p_h$  croissent adéquatement en fonction de la taille échantillonnale, voir Lewis et Reinsel (1985). Ce qui permet, sous des conditions additionnelles sur les coefficients des processus VAR( $\infty$ ) de montrer que  $\hat{\mathbf{a}}_t^{(1)}$  et  $\hat{\mathbf{a}}_t^{(2)}$  sont des estimateurs convergents de  $\mathbf{a}_t^{(1)}$  et  $\mathbf{a}_t^{(2)}$ . Aussi, la statistique  $\mathcal{Q}_N$  fait intervenir les corrélations croisées à tous les délais. Ce qui permettra d'avoir une convergence de  $\mathcal{Q}_N$  vers une loi normale.

### 3.2.2. Loi asymptotique sous l'hypothèse nulle

Dans ce qui suit, la notion de différence de martingales sera souvent utilisée. La définition suivante provient de White (1984).

#### Définition 3.2.1.

*Soit  $\{X_t, t \in \mathbb{N}\}$  une suite de variables aléatoires définies sur un espace de probabilité  $(\Omega, \mathcal{F}, P)$ . Soit  $\{\mathcal{F}_t\}$  une suite croissante de sous  $\sigma$ -algèbres de  $\mathcal{F}$  telle que  $X_t$  est  $\mathcal{F}_t$ -mesurable. On dit que  $\{X_t, \mathcal{F}_t\}$  est une différence de martingales si  $\mathbb{E}(X_t | \mathcal{F}_{t-1}) = 0$  presque partout, pour  $t \geq 2$ .*

Avant d'énoncer le théorème de convergence de la statistique  $\mathcal{Q}_N$ , voici un théorème central limite pour une différence de martingales établi par Brown (1971) qui nous servira dans la preuve de notre résultat. Ce théorème se retrouve en particulier dans Taniguchi et Kakizawa (2000, p.22) et une version plus générale pour un tableau triangulaire de variables aléatoires est présenté dans Fuller (1996, p.235). Dans ce qui suit, nous notons par  $\mathbb{I}$  la fonction indicatrice.

#### Théorème 3.2.1.

*Soient  $\{X_n, n = 1, 2, \dots\}$  une suite de variables aléatoires définies sur un espace de probabilité  $(\Omega, \mathfrak{F}, P)$  et  $\{\mathfrak{F}_n, n = 1, 2, \dots\}$  une suite de sous  $\sigma$ -algèbres de  $\mathfrak{F}$  telle*

que  $\mathfrak{F}_n$  est la sous  $\sigma$ -algèbre engendrée par  $\{X_1, \dots, X_n\}$ . Supposons que  $\{X_n, \mathcal{F}_n\}$  est une différence de martingales et notons par:

$$S_0 = 0, S_n = \sum_{i=1}^n X_i, V_n^2 = \sum_{i=1}^n E(X_i^2 | \mathfrak{F}_{i-1}) \text{ et } s_n^2 = E(V_n^2).$$

Si on suppose que

$$s_n^{-2} V_n^2 \xrightarrow{p} 1,$$

et si la condition de Lindeberg suivante est satisfaite

$$s_n^{-2} \sum_{i=1}^n E(X_i^2 \mathbb{I}(|X_i| \geq \varepsilon s_n)) \rightarrow 0, \quad \forall \varepsilon > 0,$$

quand  $n \rightarrow \infty$ , alors on a

$$s_n^{-1} S_n \xrightarrow{d} \mathcal{N}(0, 1).$$

En pratique, la condition de Lyapounov suivante est plus facile à vérifier et implique la condition de Lindeberg

$$s_n^{-2-\delta} \sum_{i=1}^n E(|X_i|^{2+\delta}) \rightarrow 0, \quad \delta > 0.$$

Sous l'hypothèse d'indépendance entre les deux sous processus  $\mathbf{X}^{(1)}$  et  $\mathbf{X}^{(2)}$ , on obtiendra la distribution asymptotique de  $\mathcal{Q}_N$  sous les hypothèses faites sur les processus  $\mathbf{X}^{(1)}$  et  $\mathbf{X}^{(2)}$  et sous les hypothèses 2.3.1 et 3.1.1.

### **Théorème 3.2.2.**

Supposons que les hypothèses 2.3.1 et 3.1.1 sont vérifiées. Soient  $M = M(N) \rightarrow \infty$ ,  $M/N \rightarrow 0$  quand  $N \rightarrow \infty$  et  $p_h$  avec  $h = 1, 2$ , deux suites d'entiers positifs qui satisfont les conditions suivantes:

$$p_h = o\left(\frac{N^{1/2}}{M^{1/4}}\right), \quad N \sum_{j=p_h+1}^{\infty} \|\Phi_j^{(h)}\|_2^2 = o\left(\frac{N^{1/2}}{M^{1/4}}\right).$$

Si les deux processus innovations  $\{\mathbf{a}_t^{(1)}\}$  et  $\{\mathbf{a}_s^{(2)}\}$  sont indépendants, alors

$$\mathcal{Q}_N \xrightarrow{d} \mathcal{N}(0, 1).$$

PREUVE DU THÉORÈME 3.2.2.

Pour démontrer ce théorème, on fera intervenir les matrices de covariances des deux sous processus  $\mathbf{a}^{(1)}$  et  $\mathbf{a}^{(2)}$  dans la statistique  $\mathcal{Q}_N$ . Notons d'abord que d'après Lütkepohl (1991, p.309), on a pour  $h = 1, 2$ ,

$$\mathbf{C}_{\hat{\mathbf{a}}}^{(hh)}(0) - \boldsymbol{\Sigma}_h = \mathcal{O}_p(N^{-1/2}) \quad (3.2.6)$$

et il en découle que

$$\mathbf{C}_{\hat{\mathbf{a}}}^{(22)}(0)^{-1} \otimes \mathbf{C}_{\hat{\mathbf{a}}}^{(11)}(0)^{-1} - \boldsymbol{\Sigma}_2^{-1} \otimes \boldsymbol{\Sigma}_1^{-1} = \mathcal{O}_p(N^{-1/2}). \quad (3.2.7)$$

Ainsi, on peut introduire directement dans notre statistique la matrice de covariance des innovations et on aura:

$$\mathcal{T}(\hat{\mathbf{a}}, \hat{\boldsymbol{\Sigma}}) = \mathcal{T}(\hat{\mathbf{a}}, \boldsymbol{\Sigma}) + (\mathcal{T}(\hat{\mathbf{a}}, \hat{\boldsymbol{\Sigma}}) - \mathcal{T}(\hat{\mathbf{a}}, \boldsymbol{\Sigma})),$$

où

$$\mathcal{T}(\hat{\mathbf{a}}, \boldsymbol{\Sigma}) = N \sum_{j=1-N}^{N-1} k^2(j/M) \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)^T (\boldsymbol{\Sigma}_2^{-1} \otimes \boldsymbol{\Sigma}_1^{-1}) \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j).$$

Ainsi, la statistique  $\mathcal{Q}_N$  devient

$$\begin{aligned} \mathcal{Q}_N &= \frac{\mathcal{T}(\hat{\mathbf{a}}, \hat{\boldsymbol{\Sigma}}) - m_1 m_2 S_N(k)}{\sqrt{2m_1 m_2 D_N(k)}} \\ &= \frac{\mathcal{T}(\hat{\mathbf{a}}, \boldsymbol{\Sigma}) + (\mathcal{T}(\hat{\mathbf{a}}, \hat{\boldsymbol{\Sigma}}) - \mathcal{T}(\hat{\mathbf{a}}, \boldsymbol{\Sigma})) - m_1 m_2 S_N(k)}{\sqrt{2m_1 m_2 D_N(k)}} \\ &= \frac{\mathcal{T}(\hat{\mathbf{a}}, \boldsymbol{\Sigma}) - m_1 m_2 S_N(k)}{\sqrt{2m_1 m_2 D_N(k)}} + \frac{\mathcal{T}(\hat{\mathbf{a}}, \hat{\boldsymbol{\Sigma}}) - \mathcal{T}(\hat{\mathbf{a}}, \boldsymbol{\Sigma})}{\sqrt{2m_1 m_2 D_N(k)}}. \end{aligned} \quad (3.2.8)$$

Ainsi, il suffit de montrer que la première partie de  $\mathcal{Q}_N$  converge en distribution vers la loi normale et que la deuxième partie converge en probabilité vers 0. On

énoncera une proposition qui sera le résultat principal correspondant à la première partie.

**Proposition 3.2.1.**

Supposons que les hypothèses 2.3.1 et 3.1.1 sont vérifiées. Soit  $M = M(N) \rightarrow \infty$ ,  $M/N \rightarrow 0$ . Si on suppose que les processus  $\{\mathbf{a}_t^{(1)}\}$  et  $\{\mathbf{a}_s^{(2)}\}$  sont indépendants, alors

$$\frac{\mathcal{T}(\hat{\mathbf{a}}, \Sigma) - m_1 m_2 S_N(k)}{\sqrt{2m_1 m_2 D_N(k)}} \xrightarrow{d} \mathcal{N}(0, 1).$$

Cette proposition découle des deux lemmes suivants dont les preuves techniques et longues sont en appendice à la fin de ce chapitre. En effet, de façon similaire à (3.2.8), nous pouvons écrire

$$\frac{\mathcal{T}(\hat{\mathbf{a}}, \Sigma) - m_1 m_2 S_N(k)}{\sqrt{2m_1 m_2 D_N(k)}} = \frac{\mathcal{T}(\mathbf{a}, \Sigma) - m_1 m_2 S_N(k)}{\sqrt{2m_1 m_2 D_N(k)}} + \frac{\mathcal{T}(\hat{\mathbf{a}}, \Sigma) - \mathcal{T}(\mathbf{a}, \Sigma)}{\sqrt{2m_1 m_2 D_N(k)}}.$$

**Lemme 3.2.1.**

Supposons que les hypothèses 2.3.1 et 3.1.1 sont réalisées. Soit  $M = M(N) \rightarrow \infty$ ,  $M/N \rightarrow 0$ . Si on suppose que les processus  $\{\mathbf{a}_t^{(1)}\}$  et  $\{\mathbf{a}_s^{(2)}\}$  sont indépendants, alors

$$\frac{\mathcal{T}(\mathbf{a}, \Sigma) - m_1 m_2 S_N(k)}{\sqrt{2m_1 m_2 D_N(k)}} \xrightarrow{d} \mathcal{N}(0, 1).$$

**Lemme 3.2.2.**

Supposons que les hypothèses 2.3.1 et 3.1.1 sont réalisées. Soit  $M = M(N) \rightarrow \infty$ ,  $M/N \rightarrow 0$ . Si on suppose que les processus  $\{\mathbf{a}_t^{(1)}\}$  et  $\{\mathbf{a}_s^{(2)}\}$  sont indépendants, alors

$$\mathcal{T}(\hat{\mathbf{a}}, \Sigma) - \mathcal{T}(\mathbf{a}, \Sigma) = o_p(M^{1/2}).$$

Pour la deuxième partie de la statistique  $\mathcal{Q}_N$ , nous avons le résultat suivant.

**Proposition 3.2.2.**

Supposons que les conditions du théorème 3.2.2 sont satisfaites, alors

$$\frac{\mathcal{T}(\hat{\mathbf{a}}, \hat{\Sigma}) - \mathcal{T}(\hat{\mathbf{a}}, \Sigma)}{\sqrt{2m_1 m_2 D_N(k)}} \xrightarrow{p} 0.$$

PREUVE DE LA PROPOSITION 3.2.2.

Puisque  $D_N(k) = MD(k)\{1 + o(1)\}$ , il suffit alors de montrer que

$$\mathcal{T}(\hat{\mathbf{a}}, \hat{\Sigma}) - \mathcal{T}(\hat{\mathbf{a}}, \Sigma) = O_p(M/N^{1/2}).$$

Or, d'après la relation (3.2.7), on peut déduire que

$$\begin{aligned} \mathcal{T}(\hat{\mathbf{a}}, \hat{\Sigma}) - \mathcal{T}(\hat{\mathbf{a}}, \Sigma) &= N \sum_{j=1-N}^{N-1} k^2(j/M) \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)^T \mathbf{O}_p(N^{-1/2}) \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j) \\ &= O_p(N^{1/2}) \sum_{j=1-N}^{N-1} k^2(j/M) \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)^T \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j). \end{aligned}$$

Or, on peut voir que

$$\mathcal{B}(N) = \sum_{j=1-N}^{N-1} k^2(j/M) \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)^T \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j) = O_p(M/N). \quad (3.2.9)$$

En effet,  $\mathcal{B}(N)$  peut être écrite comme

$$\begin{aligned} \mathcal{B}(N) &= \sum_{j=1-N}^{N-1} k^2(j/M) \{ \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)^T \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j) - \mathbf{c}_{\mathbf{a}}^{(12)}(j)^T \mathbf{c}_{\mathbf{a}}^{(12)}(j) \} \\ &\quad + \sum_{j=1-N}^{N-1} k^2(j/M) \mathbf{c}_{\mathbf{a}}^{(12)}(j)^T \mathbf{c}_{\mathbf{a}}^{(12)}(j). \end{aligned}$$

Or d'après les relations (3.2.43) et (3.2.44), on a

$$\sum_{j=1-N}^{N-1} k^2(j/M) \{ \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)^T \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j) - \mathbf{c}_{\mathbf{a}}^{(12)}(j)^T \mathbf{c}_{\mathbf{a}}^{(12)}(j) \} = o_p(M^{1/2}/N).$$

Ainsi il suffit de vérifier que

$$\mathcal{B}_1(N) = \sum_{j=1-N}^{N-1} k^2(j/M) \mathbf{c}_a^{(12)}(j)^T \mathbf{c}_a^{(12)}(j) = O_p(M/N).$$

Nous pouvons écrire

$$\begin{aligned} \mathcal{B}_1(N) &= \sum_{j=1-N}^{N-1} k^2(j/M) \mathbf{c}_a^{(12)}(j)^T \mathbf{c}_a^{(12)}(j) \\ &= \sum_{j=1-N}^{N-1} k^2(j/M) \text{tr}[\mathbf{C}_a^{(12)}(j)^T \mathbf{C}_a^{(12)}(j)] \\ &= \sum_{j=1-N}^{-1} k^2(j/M) \text{tr}[\mathbf{C}_a^{(12)}(j)^T \mathbf{C}_a^{(12)}(j)] + \sum_{j=0}^{N-1} k^2(j/M) \text{tr}[\mathbf{C}_a^{(12)}(j)^T \mathbf{C}_a^{(12)}(j)] \\ &= \mathcal{B}_2(N) + \mathcal{B}_3(N) \end{aligned}$$

où

$$\mathcal{B}_2(N) = \sum_{j=1}^{N-1} k^2(j/M) \text{tr}[\mathbf{C}_a^{(12)}(-j)^T \mathbf{C}_a^{(12)}(-j)],$$

$$\mathcal{B}_3(N) = \sum_{j=0}^{N-1} k^2(j/M) \text{tr}[\mathbf{C}_a^{(12)}(j)^T \mathbf{C}_a^{(12)}(j)].$$

Ainsi, en considérant chaque terme séparément et en utilisant la relation  $\mathbf{C}_a^{(12)}(-j) = \mathbf{C}_a^{(21)}(j)^T$ , on obtient

$$\begin{aligned} \mathcal{B}_2(N) &= N^{-2} \sum_{j=1}^{N-1} k^2(j/M) \text{tr} \left[ \left( \sum_{t=j+1}^N \mathbf{a}_t^{(2)} \mathbf{a}_{t-j}^{(1)T} \right) \left( \sum_{t=j+1}^N \mathbf{a}_{t-j}^{(1)} \mathbf{a}_t^{(2)T} \right) \right] \\ &= N^{-2} \sum_{j=1}^{N-1} k^2(j/M) \text{tr} \left[ \sum_{t=j+1}^N \mathbf{a}_t^{(2)} \mathbf{a}_{t-j}^{(1)T} \mathbf{a}_{t-j}^{(1)} \mathbf{a}_t^{(2)T} \right. \\ &\quad \left. + \sum_{t=j+1}^N \sum_{\substack{t_1=j+1 \\ t_1 \neq t}}^N \mathbf{a}_t^{(2)} \mathbf{a}_{t-j}^{(1)T} \mathbf{a}_{t_1-j}^{(1)} \mathbf{a}_{t_1}^{(2)T} \right]. \end{aligned} \tag{3.2.10}$$

Comme  $\mathbf{a}^{(1)}$  et  $\mathbf{a}^{(2)}$  sont deux bruits blancs forts indépendants,  $\mathbb{E}(\mathbf{a}_t^{(2)} \mathbf{a}_{t-j}^{(1)T} \mathbf{a}_{t_1-j}^{(1)} \mathbf{a}_{t_1}^{(2)T}) = 0$  lorsque  $t \neq t_1$ , le deuxième terme de (3.2.10) est nul. De plus, en utilisant la propriété  $\text{tr}(AB) = \text{tr}(BA)$  et le fait que  $\mathbb{E}[\text{tr}(\mathbf{a}_t^{(h)T} \mathbf{a}_t^{(h)})]$  est bornée, on obtient

$$\begin{aligned} \mathbb{E}(\mathcal{B}_2(N)) &\leq \Delta N^{-2} \sum_{j=1}^{N-1} (N-j) k^2 (j/M) \\ &= \Delta M N^{-1} \left\{ \frac{1}{M} \sum_{j=1}^{N-1} \left(1 - \frac{j}{N}\right) k^2 (j/M) \right\} \\ &= O(M/N). \end{aligned}$$

Par symétrie, on montre aussi que

$$\mathbb{E}(\mathcal{B}_3(N)) = O(M/N).$$

Ainsi,

$$\mathcal{T}(\hat{\mathbf{a}}, \hat{\Sigma}) - \mathcal{T}(\hat{\mathbf{a}}, \Sigma) = O_p(N^{1/2}) O_p(M/N) = O_p(M/N^{1/2}).$$

D'où le résultat. □

Finalement, d'après les deux propositions et les deux lemmes, le théorème de convergence de la statistique du test en découle. □

Sous les conditions sur  $p_h$ , les effets échantillonnaires des  $\hat{\Phi}_l^{(h)}$  n'influencent pas la distribution asymptotique de  $\mathcal{Q}_N$ . La condition  $p_h = o\left(\frac{N^{1/2}}{M^{1/4}}\right)$  contraint  $p_h$  à ne pas croître trop rapidement, plus particulièrement,  $p_h$  ne va pas croître plus vite que  $N^{1/2}$ , quand  $M$  tend vers l'infini. Ceci nous assure que la variance asymptotique des  $\hat{\Phi}_l^{(h)}$  sera négligeable. D'autre part, la condition  $N \sum_{l=p_h+1}^{\infty} \|\Phi_l^{(h)}\|_2^2 = o\left(\frac{N^{1/2}}{M^{1/4}}\right)$  pousse aussi  $p_h$  à ne pas croître trop rapidement et  $\|\Phi_l^{(h)}\|_2$  à décroître assez rapidement vers 0 et ainsi le biais causé par l'approximation des processus VAR( $\infty$ ) par des autorégressions finies est négligeable.

## Appendice

Avant de commencer la preuve de cette proposition, et étant donnée l'utilisation fréquente de l'inégalité de Minkowski, on va la rappeler brièvement.

Si on considère la norme  $\|\cdot\|_p$  définie sur l'espace des fonctions  $p$ -intégrables  $\mathbb{L}_p$ , et si  $(f_i)_{i=1}^n$ , une suite de fonctions dans  $\mathbb{L}_p$ , on a,

$$\left\| \sum_{i=1}^n f_i \right\|_p \leq \left[ \sum_{i=1}^n \|f_i\|_p^{1/p} \right]^p,$$

qu'on va pouvoir appliquer sur l'espace  $\mathbb{X}_p$  des variables aléatoires d'ordre  $p$ , telles que  $\mathbb{E}(|X|^p) < \infty$  quand  $X \in \mathbb{X}_p$ , ce qui donne

$$\mathbb{E} \left( \sum_{i=1}^N X_i \right)^p \leq \left[ \sum_{i=1}^N \mathbb{E}^{1/p} (X_i^p) \right]^p. \quad (3.2.11)$$

Aussi, notons par  $\langle \mathbf{x}_t, \mathbf{x}_s \rangle = \mathbf{x}_t^T \mathbf{x}_s$  le produit scalaire des deux vecteurs et la norme euclidienne d'un vecteur est donnée par  $\|\mathbf{x}_t\| = \sqrt{\langle \mathbf{x}_t, \mathbf{x}_t \rangle}$ .

PREUVE DU LEMME 3.2.1:

Considérons la transformation linéaire suivante:

$$\mathbf{b}_t = \Sigma^{-1/2} \mathbf{a}_t.$$

On note d'abord que le processus  $\mathbf{b} = \{\mathbf{b}_t, t \in \mathbb{Z}\}$  est un bruit blanc de matrice de covariance l'identité. Aussi, on peut vérifier que  $\mathcal{T}_a^{(12)} = \mathcal{T}_b^{(12)}$ . En effet, on remarque d'abord que la matrice de covariance croisée  $\mathbf{C}_b^{(12)}(j)$  du bruit blanc  $\mathbf{b}$  s'écrit en termes de la matrice de covariance croisée du bruit blanc  $\mathbf{a}$ , et on a

$$\mathbf{C}_b^{(12)}(j) = \Sigma_1^{-1/2} \mathbf{C}_a^{(12)}(j) \Sigma_2^{-1/2}.$$

Par la suite, si on applique la propriété suivante:  $\text{vec}(ABC) = (C^T \otimes A)\text{vec}(B)$ , on obtient que

$$\mathbf{c}_b^{(12)}(j) = (\Sigma_2 \otimes \Sigma_1)^{-1/2} \mathbf{c}_a^{(12)}(j) \quad (3.2.12)$$

et donc

$$\begin{aligned} \mathcal{T}_a^{(12)} &= N \sum_{j=1-N}^{N-1} k^2(j/M) \mathbf{c}_a^{(12)}(j)^T (\Sigma_2^{-1} \otimes \Sigma_1^{-1}) \mathbf{c}_a^{(12)}(j) \\ &= N \sum_{j=1-N}^{N-1} k^2(j/M) \mathbf{c}_b^{(12)}(j)^T \mathbf{c}_b^{(12)}(j) \\ &= \mathcal{T}_b^{(12)}. \end{aligned}$$

Ainsi, pour montrer le lemme 3.2.1, on va considérer la quantité  $\mathcal{T}_b^{(12)}$ , où le processus  $\mathbf{b}$  a pour matrice de covariance l'identité. Par définition de  $\mathcal{T}_b^{(12)}$ , et d'après (3.2.3), (2.3.2), et sachant que  $\|\text{vec}(A)\| = \|A\|_2$ , on écrit

$$\begin{aligned} \mathcal{T}_b^{(12)} &= N \sum_{j=1-N}^{N-1} k^2(j/M) \mathbf{c}_b^{(12)}(j)^T \mathbf{c}_b^{(12)}(j) \\ &= N \sum_{j=1-N}^{N-1} k^2(j/M) \text{tr}[\mathbf{C}_b^{(12)}(j)^T \mathbf{C}_b^{(12)}(j)] \\ &= N \sum_{j=1}^{N-1} k^2(j/M) \text{tr}[\mathbf{C}_b^{(12)}(-j)^T \mathbf{C}_b^{(12)}(-j)] \\ &\quad + N \sum_{j=0}^{N-1} k^2(j/M) \text{tr}[\mathbf{C}_b^{(12)}(j)^T \mathbf{C}_b^{(12)}(j)]. \end{aligned}$$

Or, pour  $j > 0$ , puisque  $\mathbf{C}_b^{(12)}(-j) = \mathbf{C}_b^{(21)}(j)^T$ , on a

$$\begin{aligned} \text{tr}[\mathbf{C}_b^{(12)}(-j)^T \mathbf{C}_b^{(12)}(-j)] &= N^{-2} \text{tr}[(\sum_{t=j+1}^N \mathbf{b}_{t-j}^{(1)} \mathbf{b}_t^{(2)T})^T (\sum_{t=j+1}^N \mathbf{b}_{t-j}^{(1)} \mathbf{b}_t^{(2)T})] \\ &= N^{-2} [\sum_{t=j+1}^N \|\mathbf{b}_{t-j}^{(1)}\|^2 \|\mathbf{b}_t^{(2)}\|^2 \\ &\quad + 2 \sum_{t=j+2}^N \sum_{s=j+1}^{t-1} \langle \mathbf{b}_{t-j}^{(1)}, \mathbf{b}_{s-j}^{(1)} \rangle \langle \mathbf{b}_t^{(2)}, \mathbf{b}_s^{(2)} \rangle]. \end{aligned}$$

De même, on a aussi pour  $j \geq 0$

$$\begin{aligned} \text{tr}[\mathbf{C}_b^{(12)}(j)^T \mathbf{C}_b^{(12)}(j)] &= N^{-2} [\sum_{t=j+1}^N \|\mathbf{b}_t^{(1)}\|^2 \|\mathbf{b}_{t-j}^{(2)}\|^2 \\ &\quad + 2 \sum_{t=j+2}^N \sum_{s=j+1}^{t-1} \langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle \langle \mathbf{b}_{t-j}^{(2)}, \mathbf{b}_{s-j}^{(2)} \rangle]. \end{aligned}$$

Notons alors par

$$H_{1N} = N^{-1} \sum_{j=0}^{N-1} k^2(j/M) \sum_{t=j+1}^N \|\mathbf{b}_t^{(1)}\|^2 \|\mathbf{b}_{t-j}^{(2)}\|^2, \quad (3.2.13)$$

$$H_{2N} = N^{-1} \sum_{j=1}^{N-1} k^2(j/M) \sum_{t=j+1}^N \|\mathbf{b}_{t-j}^{(1)}\|^2 \|\mathbf{b}_t^{(2)}\|^2, \quad (3.2.14)$$

$$H_N = H_{1N} + H_{2N}, \quad (3.2.15)$$

$$W_{1N}^* = 2N^{-1} \sum_{j=0}^{N-2} k^2(j/M) \sum_{t=j+2}^N \sum_{s=j+1}^{t-1} \langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle \langle \mathbf{b}_{t-j}^{(2)}, \mathbf{b}_{s-j}^{(2)} \rangle, \quad (3.2.16)$$

$$W_{2N}^* = 2N^{-1} \sum_{j=1}^{N-2} k^2(j/M) \sum_{t=j+2}^N \sum_{s=j+1}^{t-1} \langle \mathbf{b}_{t-j}^{(1)}, \mathbf{b}_{s-j}^{(1)} \rangle \langle \mathbf{b}_t^{(2)}, \mathbf{b}_s^{(2)} \rangle, \quad (3.2.17)$$

$$W_N^* = W_{1N}^* + W_{2N}^*. \quad (3.2.18)$$

On a alors,

$$\mathcal{T}_b^{(12)} = H_N + W_N^*.$$

La suite de la preuve découle des deux lemmes suivants:

**Lemme 3.2.3.**

*Sous les mêmes hypothèses et notations que celles du théorème 3.2.2, on a:*

$$\sigma^{-1}(N)\{H_N - m_1 m_2 S_N(k)\} \xrightarrow{p} 0 \quad \text{où} \quad \sigma^2(N) = 2m_1 m_2 D_N(k).$$

**Lemme 3.2.4.**

*Sous les mêmes hypothèses et notations que celles du théorème 3.2.2, on a:*

$$\sigma^{-1}(N)W_N^* \xrightarrow{d} N(0, 1). \tag{3.2.19}$$

PREUVE DU LEMME 3.2.3.

Dans la démonstration, l'élément  $\Delta$  est une constante de majoration qui prendra, au besoin, plusieurs valeurs distinctes mais qui ne dépendent pas de  $N$ . Remarquons ensuite que  $\mathbb{E}(H_N) = m_1 m_2 S_N(k)$ . En effet, de (3.2.13), (3.2.14) et (3.2.15), on écrit

$$\begin{aligned} \mathbb{E}(H_N) &= \mathbb{E}[H_{1N} + H_{2N}] \\ &= N^{-1} \mathbb{E} \left\{ \sum_{j=0}^{N-1} k^2(j/M) \sum_{t=j+1}^N \|\mathbf{b}_t^{(1)}\|^2 \|\mathbf{b}_{t-j}^{(2)}\|^2 \right. \\ &\quad \left. + \sum_{j=1}^{N-1} k^2(j/M) \sum_{t=j+1}^N \|\mathbf{b}_{t-j}^{(1)}\|^2 \|\mathbf{b}_t^{(2)}\|^2 \right\}. \end{aligned}$$

Par la symétrie du noyau  $k(z)$ , par l'indépendance des deux processus, et puisque par hypothèse  $\mathbb{E}(\|\mathbf{b}_t^{(1)}\|^2) = m_1$ ,  $\mathbb{E}(\|\mathbf{b}_t^{(2)}\|^2) = m_2$ , on trouve

$$\begin{aligned}
\mathbb{E}(H_N) &= m_1 m_2 N^{-1} \left( \sum_{j=0}^{N-1} \sum_{t=j+1}^N k^2(j/M) + \sum_{j=1}^{N-1} \sum_{t=j+1}^N k^2(j/M) \right) \\
&= m_1 m_2 N^{-1} \left( \sum_{j=0}^{N-1} (N-j) k^2(j/M) + \sum_{j=1}^{N-1} (N-j) k^2(j/M) \right) \\
&= m_1 m_2 \sum_{j=1-N}^{N-1} \left(1 - \frac{|j|}{N}\right) k^2(j/M) \\
&= m_1 m_2 S_N(k).
\end{aligned}$$

En considérant le premier terme de  $H_N$ ,  $H_{1N}$  donné par (3.2.13), on a

$$\mathbb{E}(H_{1N} - \mathbb{E}H_{1N})^2 = \mathbb{E} \left[ N^{-1} \sum_{j=0}^{N-1} k^2(j/M) \sum_{t=j+1}^N (\|\mathbf{b}_t^{(1)}\|^2 \|\mathbf{b}_{t-j}^{(2)}\|^2 - m_1 m_2) \right]^2$$

Par l'application de l'inégalité de Minkowski (3.2.11), on obtient

$$\begin{aligned}
\mathbb{E}(H_{1N} - \mathbb{E}H_{1N})^2 &\leq \left\{ \sum_{j=0}^{N-1} \mathbb{E} [k^4(j/M) (N^{-1} \sum_{t=j+1}^N (\|\mathbf{b}_t^{(1)}\|^2 \|\mathbf{b}_{t-j}^{(2)}\|^2 - m_1 m_2)^2)^{1/2} \right\}^2 \\
&\leq \left( \sum_{j=0}^{N-1} k^2(j/M) [\mathbb{E} \{ N^{-1} \sum_{t=j+1}^N (\|\mathbf{b}_t^{(1)}\|^2 \|\mathbf{b}_{t-j}^{(2)}\|^2 - m_1 m_2) \}^2]^{1/2} \right)^2 \\
&\leq \frac{\Delta M^2}{N} \left\{ M^{-1} \sum_{j=0}^{N-1} k^2(j/M) \right\}^2.
\end{aligned}$$

Par l'hypothèse 3.1.1, puisque  $M \rightarrow \infty$  lorsque  $N \rightarrow \infty$ , on peut vérifier que  $M^{-1} \sum_{j=0}^{N-1} k^2(j/M) \rightarrow \int_0^\infty k^2(z) dz < \infty$  et ainsi

$$\mathbb{E}(H_{1N} - \mathbb{E}H_{1N})^2 = O(M^2/N).$$

Aussi, on peut montrer par symétrie que  $\mathbb{E}(H_{2N} - \mathbb{E}H_{2N})^2 = O(M^2/N)$ . Donc, en appliquant à nouveau l'inégalité de Minkowski, on obtient

$$\begin{aligned} \mathbb{E}(H_N - \mathbb{E}H_N)^2 &= \mathbb{E}((H_{1N} - \mathbb{E}H_{1N}) + (H_{2N} - \mathbb{E}H_{2N}))^2 \\ &\leq \{\mathbb{E}^{1/2}(H_{1N} - \mathbb{E}H_{1N})^2 + \mathbb{E}^{1/2}(H_{2N} - \mathbb{E}H_{2N})^2\}^2 \\ &= O(M^2/N). \end{aligned}$$

De plus, puisque  $M^{-1}D_N(k)$  converge vers  $D(k)$  défini par (3.2.5), quand  $N$  tend vers l'infini et que  $M/N$  tend vers 0, on a que

$$\sigma^2(N) = 2m_1m_2MD(k)\{1 + o(1)\} = O(M), \quad (3.2.20)$$

et il s'en suit alors que  $\sigma^{-1}(N)\{H_N - m_1m_2S_N(k)\} \xrightarrow{P} 0$ . Ce qui prouve le premier lemme. □

PREUVE DU LEMME 3.2.4.

Pour démontrer ce lemme, on développe la deuxième partie de notre statistique qui est constituée par le produit des composantes des deux bruits blancs. D'après (3.2.16), (3.2.17) et (3.2.18), et puisque  $\sum_{j=0}^{N-2} \sum_{t=j+2}^N \sum_{s=j+1}^{t-1} = \sum_{t=2}^N \sum_{s=1}^{t-1} \sum_{j=0}^{s-1}$  et  $\sum_{j=1}^{N-2} \sum_{t=j+2}^N \sum_{s=j+1}^{t-1} = \sum_{t=3}^N \sum_{s=1}^{t-1} \sum_{j=1}^{s-1}$ , on écrit

$$\begin{aligned}
W_N^* &= 2N^{-1} \left\{ \sum_{t=2}^N \sum_{s=1}^{t-1} \sum_{j=0}^{s-1} k^2(j/M) \langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle \langle \mathbf{b}_{t-j}^{(2)}, \mathbf{b}_{s-j}^{(2)} \rangle \right. \\
&\quad \left. + \sum_{t=3}^N \sum_{s=2}^{t-1} \sum_{j=1}^{s-1} k^2(j/M) \langle \mathbf{b}_t^{(2)}, \mathbf{b}_s^{(2)} \rangle \langle \mathbf{b}_{t-j}^{(1)}, \mathbf{b}_{s-j}^{(1)} \rangle \right\} \\
&= 2N^{-1} \sum_{t=3}^N \sum_{s=2}^{t-1} \sum_{j=0}^{s-1} k^2(j/M) \langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle \langle \mathbf{b}_{t-j}^{(2)}, \mathbf{b}_{s-j}^{(2)} \rangle \\
&\quad + 2N^{-1} \sum_{t=3}^N \sum_{s=2}^{t-1} \sum_{j=1}^{s-1} k^2(j/M) \langle \mathbf{b}_t^{(2)}, \mathbf{b}_s^{(2)} \rangle \langle \mathbf{b}_{t-j}^{(1)}, \mathbf{b}_{s-j}^{(1)} \rangle \\
&\quad + 2N^{-1} \sum_{t=2}^N \langle \mathbf{b}_t^{(1)}, \mathbf{b}_1^{(1)} \rangle \langle \mathbf{b}_t^{(2)}, \mathbf{b}_1^{(2)} \rangle.
\end{aligned}$$

Ainsi, on peut écrire  $W_N^*$  sous forme de trois sommes qu'on va étudier séparément:

$$\begin{aligned}
W_N^* &= N^{-1} \sum_{t=3}^N (W_{1Nt} + W_{2Nt}) + \omega_N \\
&= N^{-1} \sum_{t=3}^N W_{Nt} + \omega_N,
\end{aligned}$$

où

$$W_{1Nt} = 2 \sum_{s=2}^{t-1} \sum_{j=0}^{s-1} k^2(j/M) \langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle \langle \mathbf{b}_{t-j}^{(2)}, \mathbf{b}_{s-j}^{(2)} \rangle, \quad (3.2.21)$$

$$W_{2Nt} = 2 \sum_{s=2}^{t-1} \sum_{j=1}^{s-1} k^2(j/M) \langle \mathbf{b}_t^{(2)}, \mathbf{b}_s^{(2)} \rangle \langle \mathbf{b}_{t-j}^{(1)}, \mathbf{b}_{s-j}^{(1)} \rangle, \quad (3.2.22)$$

$$\omega_N = 2N^{-1} \sum_{t=2}^N \langle \mathbf{b}_t^{(1)}, \mathbf{b}_1^{(1)} \rangle \langle \mathbf{b}_t^{(2)}, \mathbf{b}_1^{(2)} \rangle. \quad (3.2.23)$$

On remarque que sous l'hypothèse nulle d'indépendance entre les deux sous processus,  $\omega_N = o_p(1)$ , et d'après l'équation (3.2.20),  $\sigma(N)^{-1} = O(M^{-1/2})$ , on obtient

alors  $\sigma^{-1}(N)\omega_N \xrightarrow{P} 0$ . Donc étudier le comportement asymptotique de  $W_N^*$  revient à étudier celui de

$$W_N = N^{-1} \sum_{t=3}^N W_{Nt}. \quad (3.2.24)$$

Avant de continuer la preuve, montrons ce lemme qui nous donne la forme de la variance de  $W_N$ .

**Lemme 3.2.5.**

*Sous les mêmes hypothèses et notations que celles du théorème 3.2.2, on a:*

$$\text{Var}(W_N) = \sigma^2(N) \quad \text{où} \quad \sigma^2(N) = 2m_1m_2D_N(k). \quad (3.2.25)$$

PREUVE DU LEMME 3.2.5:

Étant données l'hypothèse 2.3.1 et l'indépendance des deux bruits blancs, on a

$$\mathbb{E}(W_{1Nt}) = 2 \sum_{s=2}^{t-1} \sum_{j=0}^{s-1} k^2(j/M) \mathbb{E}(\langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle) \mathbb{E}(\langle \mathbf{b}_{t-j}^{(2)}, \mathbf{b}_{s-j}^{(2)} \rangle).$$

Or, pour  $t \neq s$ ,  $\mathbb{E}(\langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle) = \mathbb{E}(\mathbf{b}_t^{(1)T} \mathbf{b}_s^{(1)}) = 0$  puisque  $\mathbf{b}^{(1)}$  est un bruit blanc et ainsi  $\mathbb{E}(W_{1Nt}) = 0$ . Aussi, on peut remarquer que  $\mathbb{E}(W_{1Nt}W_{2Nt}) = 0$  puisque pour  $t > s$ ,  $t > s_1$ , on a

$$\mathbb{E}(\langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle \langle \mathbf{b}_{t-j}^{(2)}, \mathbf{b}_{s-j}^{(2)} \rangle \langle \mathbf{b}_t^{(2)}, \mathbf{b}_{s_1}^{(2)} \rangle \langle \mathbf{b}_{t-j_1}^{(1)}, \mathbf{b}_{s_1-j_1}^{(1)} \rangle) = 0$$

Ainsi,

$$\text{Var}(W_{Nt}) = \mathbb{E}(W_{Nt}^2) = \mathbb{E}(W_{1Nt}^2 + W_{2Nt}^2).$$

En développant  $W_{1Nt}^2$ , on obtient

$$W_{1Nt}^2 = 4 \sum_{s=2}^{t-1} \left( \sum_{j=0}^{s-1} k^2(j/M) \langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle \langle \mathbf{b}_{t-j}^{(2)}, \mathbf{b}_{s-j}^{(2)} \rangle \right)^2 + \beta_1$$

avec

$$\beta_1 = 4 \sum_{s=2}^{t-1} \sum_{\substack{s_1=2 \\ s \neq s_1}}^{t-1} \sum_{j=0}^{s-1} \sum_{\substack{j_1=0 \\ j \neq j_1}}^{s_1-1} k^2(j/M) k^2(j_1/M) \langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle \langle \mathbf{b}_{t-j}^{(2)}, \mathbf{b}_{s-j}^{(2)} \rangle \langle \mathbf{b}_t^{(1)}, \mathbf{b}_{s_1}^{(1)} \rangle \langle \mathbf{b}_{t-j_1}^{(2)}, \mathbf{b}_{s_1-j_1}^{(2)} \rangle$$

Or, pour  $s \neq s_1 \neq t$ , on a  $\mathbb{E}(\beta_1) = 0$ . Aussi,

$$\left( \sum_{j=0}^{s-1} k^2(j/M) \langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle \langle \mathbf{b}_{t-j}^{(2)}, \mathbf{b}_{s-j}^{(2)} \rangle \right)^2 = \sum_{j=0}^{s-1} k^4(j/M) \langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle^2 \langle \mathbf{b}_{t-j}^{(2)}, \mathbf{b}_{s-j}^{(2)} \rangle^2 + \beta_2$$

avec

$$\beta_2 = \sum_{j=0}^{s-1} \sum_{\substack{j_1=0 \\ j \neq j_1}}^{s-1} k^2(j/M) k^2(j_1/M) \langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle \langle \mathbf{b}_{t-j}^{(2)}, \mathbf{b}_{s-j}^{(2)} \rangle \langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle \langle \mathbf{b}_{t-j_1}^{(2)}, \mathbf{b}_{s-j_1}^{(2)} \rangle,$$

et pour  $j \neq j_1, t \neq s$ , on a aussi  $\mathbb{E}(\beta_2) = 0$ . Ainsi, sous l'hypothèse d'indépendance des deux sous processus, et puisque  $\mathbb{E}(\langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle)^2 = m_1$  et  $\mathbb{E}(\langle \mathbf{b}_{t-j}^{(2)}, \mathbf{b}_{s-j}^{(2)} \rangle)^2 = m_2$ , on a donc

$$\begin{aligned} \mathbb{E}(W_{1Nt}^2) &= 4 \sum_{s=2}^{t-1} \sum_{j=0}^{s-1} k^4(j/M) \mathbb{E} \langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle^2 \mathbb{E} \langle \mathbf{b}_{t-j}^{(2)}, \mathbf{b}_{s-j}^{(2)} \rangle^2 \\ &= 4m_1 m_2 \sum_{s=2}^{t-1} \sum_{j=0}^{s-1} k^4(j/M). \end{aligned}$$

De même on montre que  $\mathbb{E}(W_{2Nt}^2) = 4m_1 m_2 \sum_{s=2}^{t-1} \sum_{j=1}^{s-1} k^4(j/M)$ . Alors

$$\text{Var}(W_{Nt}) = 4m_1 m_2 \sum_{s=2}^{t-1} \sum_{|j|=0}^{s-1} k^4(j/M).$$

Puisque  $\mathbb{E}(W_{Nt}) = \mathbb{E}(W_{1Nt}) + \mathbb{E}(W_{2Nt}) = 0$  et par un développement similaire à celui de  $W_{1Nt}$ , on peut montrer aussi que, pour  $t \neq t_1$ ,  $\mathbb{E}(W_{Nt} W_{Nt_1}) = 0$ . Il s'en

suit que:

$$\begin{aligned}
\text{Var}(W_N) &= \frac{1}{N^2} \sum_{t=3}^N \mathbb{E}(W_{Nt}^2) \\
&= \frac{4m_1m_2}{N^2} \sum_{t=3}^N \sum_{s=2}^{t-1} \sum_{|j|=0}^{s-1} k^4(j/M) \\
&= \frac{4m_1m_2}{N^2} \sum_{|j|=0}^{N-2} \sum_{t=|j|+2}^N \sum_{s=2}^t k^4(j/M)
\end{aligned}$$

Ainsi, on obtient

$$\begin{aligned}
\text{Var}(W_N) &= \frac{4m_1m_2}{N^2} \sum_{|j|=0}^{N-2} \frac{(N-|j|-1)(N-|j|)}{2} k^4(j/M) \\
&= 2m_1m_2 \sum_{|j|=0}^{N-2} \left(1 - \frac{|j|}{N}\right) \left(1 - \frac{|j|+1}{N}\right) k^4(j/M) \\
&= 2m_1m_2 D_N(k) \\
&= \sigma^2(N).
\end{aligned}$$

□

Si on note par  $\mathfrak{F}_t$  la  $\sigma$ -algèbre engendrée par  $(\mathbf{b}_s^{(1)}, \mathbf{b}_s^{(2)})^T$ ,  $s \leq t$ , on remarque que  $(W_{Nt}, \mathfrak{F}_{t-1})$  est une suite de différences de martingales puisque  $E(W_{Nt} | \mathfrak{F}_{t-1}) = 0$  étant donné l'indépendance entre  $\{\mathbf{b}_t^{(1)}\}$  et  $\{\mathbf{b}_s^{(2)}\}$ . En effet, on vérifie que  $\mathbb{E}(W_{1Nt} | \mathfrak{F}_{t-1}) = 0$  puisque  $\mathbb{E}(\langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle \langle \mathbf{b}_{t-j}^{(2)}, \mathbf{b}_{s-j}^{(2)} \rangle | \mathfrak{F}_{t-1}) = 0$ . De même pour  $\mathbb{E}(W_{2Nt} | \mathfrak{F}_{t-1}) = 0$ . On va montrer (3.2.19) en utilisant le théorème 3.2.1 de Brown (1971). Ainsi, montrer que  $\{\text{var}(W_N)\}^{-1/2} W_N \xrightarrow{d} N(0, 1)$  repose sur la preuve des deux lemmes suivants.

**Lemme 3.2.6.**

Sous les mêmes hypothèses et notations que celles du théorème 3.2.2, on a:

$$\sigma^{-2}(N) \frac{1}{N^2} \sum_{t=3}^N \mathbb{E}[W_{N,t}^2 \mathbb{I}\{|W_{N,t}| > \epsilon \sigma(N)\}] \rightarrow 0 \quad \forall \epsilon > 0. \quad (3.2.26)$$

**Lemme 3.2.7.**

Sous les mêmes hypothèses et notations que celles du théorème 3.2.2, on a:

$$\sigma^{-2}(N) \frac{1}{N^2} \sum_{t=3}^N \ddot{W}_{Nt}^2 \xrightarrow{p} 1, \quad (3.2.27)$$

où  $\ddot{W}_{Nt}^2 = \mathbb{E}(W_{Nt}^2 | \mathfrak{F}_{t-1})$ .

PREUVE DU LEMME 3.2.6:

Pour montrer (3.2.26), il suffit de vérifier la condition de Lyapounov

$$\sigma^{-4}(N) N^{-4} \sum_{t=3}^N \mathbb{E}(W_{Nt}^4) \rightarrow 0.$$

Pour cela, posons

$$G_{ts}^{(1)} = \sum_{j=1}^{s-1} k^2(j/M) \langle \mathbf{b}_{t-j}^{(1)}, \mathbf{b}_{s-j}^{(1)} \rangle. \quad (3.2.28)$$

$$G_{ts}^{(2)} = \sum_{j=0}^{s-1} k^2(j/M) \langle \mathbf{b}_{t-j}^{(2)}, \mathbf{b}_{s-j}^{(2)} \rangle. \quad (3.2.29)$$

D'après (3.2.21), (3.2.22), (3.2.28) et (3.2.29), on peut écrire

$$W_{1Nt} = 2 \sum_{s=2}^{t-1} \langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle G_{ts}^{(2)} \quad \text{et} \quad W_{2Nt} = 2 \sum_{s=2}^{t-1} \langle \mathbf{b}_t^{(2)}, \mathbf{b}_s^{(2)} \rangle G_{ts}^{(1)}. \quad (3.2.30)$$

Étant données l'hypothèse 2.3.1 et l'indépendance entre les deux sous processus  $\{\mathbf{b}_t^{(1)}\}$  et  $\{\mathbf{b}_s^{(2)}\}$ , on obtient:

$$\begin{aligned}
\mathbb{E}(W_{1Nt}^4) &= 16\mathbb{E}\left[\sum_{s=2}^{t-1} \langle \mathbf{b}_t^{(1)}, \mathbf{b}_s^{(1)} \rangle G_{ts}^{(2)}\right]^4 \\
&\leq 16\mathbb{E}\left[\sum_{s=2}^{t-1} \|\mathbf{b}_t^{(1)}\| \|\mathbf{b}_s^{(1)}\| G_{ts}^{(2)}\right]^4 \\
&= 16\mathbb{E}\|\mathbf{b}_t^{(1)}\|^4 \mathbb{E}\left[\sum_{s=2}^{t-1} \|\mathbf{b}_s^{(1)}\| G_{ts}^{(2)}\right]^4 \\
&\leq 48\mathbb{E}\|\mathbf{b}_t^{(1)}\|^4 \left\{ \sum_{s=2}^{t-1} [\mathbb{E}\|\mathbf{b}_s^{(1)}\|^4 \mathbb{E}(G_{ts}^{(2)})^4]^{1/2} \right\}^2 \\
&\leq \Delta \left\{ \sum_{s=2}^{t-1} [\mathbb{E}(G_{ts}^{(2)})^4]^{1/2} \right\}^2.
\end{aligned}$$

La deuxième inégalité provient du fait que si  $Y_1, \dots, Y_n$  est une suite de variables aléatoires telle que  $\mathbb{E}(Y_i) = 0$  et  $\mathbb{E}(Y_i f(Y_j, Y_k, Y_l)) = 0$  pour  $i \neq j, k, l$  et pour toute fonction  $f$ , alors

$$\mathbb{E}\left(\sum_{i=1}^n Y_i\right)^4 \leq 3 \left\{ \sum_{i=1}^n [\mathbb{E}(Y_i^4)]^{1/2} \right\}^{1/2}. \quad (3.2.31)$$

Aussi, pour  $t > s$ , on peut voir que

$$\begin{aligned}
\mathbb{E}(G_{ts}^{(2)})^4 &= \mathbb{E}\left[\sum_{j=1}^{s-1} k^2(j/M) \langle \mathbf{b}_{t-j}^{(2)}, \mathbf{b}_{s-j}^{(2)} \rangle\right]^4 \\
&\leq \mathbb{E}\left[\sum_{j=1}^{s-1} k^2(j/M) \|\mathbf{b}_{t-j}^{(2)}\| \|\mathbf{b}_{s-j}^{(2)}\|\right]^4,
\end{aligned}$$

et en appliquant l'inégalité (3.2.31), on obtient

$$\begin{aligned}
\mathbb{E}(G_{ts}^{(2)})^4 &\leq 3\left\{\sum_{j=1}^{s-1} k^4(j/M)[\mathbb{E}\|\mathbf{b}_{t-j}^{(2)}\|^4\|\mathbf{b}_{s-j}^{(2)}\|^4]^{1/2}\right\}^2 \\
&\leq 3\left\{\sum_{j=1}^{s-1} k^4(j/M)[\mathbb{E}\|\mathbf{b}_{t-j}^{(2)}\|^4\mathbb{E}\|\mathbf{b}_{s-j}^{(2)}\|^4]^{1/2}\right\}^2 \\
&\leq \Delta M^2\left\{\frac{1}{M}\sum_{j=1}^{s-1} k^4(j/M)\right\}^2 \\
&= \mathcal{O}(M^2).
\end{aligned} \tag{3.2.32}$$

Ainsi, on obtient

$$\begin{aligned}
\mathbb{E}(W_{1Nt}^4) &\leq \Delta t^2 M^2 \\
&= \mathcal{O}(t^2 M^2).
\end{aligned}$$

De la même façon, et par symétrie, on peut montrer que  $\mathbb{E}(W_{2Nt}^4) = \mathcal{O}(t^2 M^2)$ .

Puisque  $(a+b)^4 \leq 8(a^4 + b^4)$ , Il s'en suit donc:

$$\sigma^{-4}(N)N^{-4}\sum_{t=3}^N \mathbb{E}(W_{Nt}^4) \leq \sigma^{-4}(N)\frac{8}{N^4}\sum_{t=3}^N \mathbb{E}(W_{1Nt}^4 + W_{2Nt}^4).$$

Or  $\sigma^{-4}(N) = \mathcal{O}(M^{-2})$  et  $\sum_{t=3}^N \mathbb{E}(W_{1Nt}^4 + W_{2Nt}^4) \leq \Delta \sum_{t=3}^N t^2 M^2 \leq \Delta N^3 M^2$ , donc

$$\sigma^{-4}(N)N^{-4}\sum_{t=3}^N \mathbb{E}(W_{Nt}^4) = \mathcal{O}(N^{-1}).$$

Ainsi la condition (3.2.26) est vérifiée. □

PREUVE DU LEMME 3.2.7:

Nous vérifions maintenant la condition (3.2.27) en montrant que:

$$\sigma^{-4}(N)var\left(N^{-2}\sum_{t=3}^N \ddot{W}_{Nt}^2\right) \rightarrow 0 \tag{3.2.33}$$

Étant donnée l'indépendance des deux sous processus d'innovation,  $\{\mathbf{b}_t^{(1)}\}$  et  $\{\mathbf{b}_s^{(2)}\}$ , on peut écrire par définition de  $\ddot{W}_{Nt}^2$  et en utilisant (3.2.30)

$$\begin{aligned}
\ddot{W}_{Nt}^2 &= \mathbb{E}[W_{Nt}^2 | \mathcal{F}_{t-1}] \\
&= \mathbb{E}[W_{1Nt}^2 | \mathcal{F}_{t-1}] + \mathbb{E}[W_{2Nt}^2 | \mathcal{F}_{t-1}] \\
&= 4\mathbb{E}[(\mathbf{b}_t^{(1)T} \sum_{s=2}^{t-1} \mathbf{b}_s^{(1)} G_{ts}^{(2)})^2 | \mathcal{F}_{t-1}] + 4\mathbb{E}[(\mathbf{b}_t^{(2)T} \sum_{s=2}^{t-1} \mathbf{b}_s^{(2)} G_{ts}^{(1)})^2 | \mathcal{F}_{t-1}] \\
&= 4tr\{\mathbb{E}[\sum_{s=2}^{t-1} G_{ts}^{(2)} \mathbf{b}_s^{(1)T} \mathbf{b}_t^{(1)} \mathbf{b}_t^{(1)T} \sum_{s_1=2}^{t-1} \mathbf{b}_{s_1}^{(1)} G_{ts_1}^{(2)} | \mathcal{F}_{t-1}]\} \\
&\quad + 4tr\{\mathbb{E}[\sum_{s=2}^{t-1} G_{ts}^{(1)} \mathbf{b}_s^{(2)T} \mathbf{b}_t^{(2)} \mathbf{b}_t^{(2)T} \sum_{s_1=2}^{t-1} \mathbf{b}_{s_1}^{(2)} G_{ts_1}^{(1)} | \mathcal{F}_{t-1}]\}
\end{aligned}$$

Si on note par  $\lambda_{1Nt} = \sum_{s=2}^{t-1} G_{ts}^{(1)} \mathbf{b}_s^{(2)T}$  et  $\lambda_{2Nt} = \sum_{s=2}^{t-1} G_{ts}^{(2)} \mathbf{b}_s^{(1)T}$ , puisque  $tr(AB) = tr(BA)$ ,  $tr(AB) \leq tr(A)tr(B)$  et que  $\mathbf{b}_t^{(h)}$  et  $\mathbf{b}_s^{(h)T}$  pour  $h = 1, 2$  sont indépendants, on a

$$\begin{aligned}
\ddot{W}_{Nt}^2 &\leq 4tr\{\mathbb{E}[\lambda_{1Nt} \lambda_{1Nt}^T | \mathcal{F}_{t-1}]\} tr\{\mathbb{E}[\mathbf{b}_t^{(1)} \mathbf{b}_t^{(1)T} | \mathcal{F}_{t-1}]\} \\
&\quad + 4tr\{\mathbb{E}[\lambda_{2Nt} \lambda_{2Nt}^T | \mathcal{F}_{t-1}]\} tr\{\mathbb{E}[\mathbf{b}_t^{(2)} \mathbf{b}_t^{(2)T} | \mathcal{F}_{t-1}]\}
\end{aligned}$$

Notons qu'en conditionnant sur  $\mathcal{F}_{t-1}$ , les termes  $\lambda_{1Nt}$  et  $\lambda_{2Nt}$  deviennent constants, et puisque  $\mathbb{E}[\mathbf{b}_t^{(h)} \mathbf{b}_t^{(h)T}] = m_h$ , pour  $h = 1, 2$ , alors

$$\ddot{W}_{Nt}^2 \leq 4(\ddot{W}_{1Nt}^2 + \ddot{W}_{2Nt}^2),$$

où

$$\ddot{W}_{1Nt}^2 = m_1 \|\lambda_{1Nt}\|^2,$$

$$\ddot{W}_{2Nt}^2 = m_2 \|\lambda_{2Nt}\|^2.$$

Ainsi, il suffit de montrer que  $M^{-2}var(N^{-2} \sum_{l=3}^N \ddot{W}_{lNt}^2) \rightarrow 0$  pour  $l = 1, 2$ .

Premièrement, on considère  $l = 1$  et on écrit:

$$\begin{aligned}
\ddot{W}_{1Nt}^2 &= m_1 \|\lambda_{1Nt}\|^2 \\
&= m_1 tr \left\{ \left( \sum_{s=2}^{t-1} G_{ts}^{(2)} \mathbf{b}_s^{(1)} \right) \left( \sum_{s=2}^{t-1} G_{ts}^{(2)} \mathbf{b}_s^{(1)T} \right) \right\} \\
&= m_1 \sum_{s=2}^{t-1} \|\mathbf{b}_s^{(1)}\|^2 (G_{ts}^{(2)})^2 + 2m_1 tr \left\{ \sum_{s_2=3}^{t-1} \sum_{s_1=2}^{s_2-1} G_{ts_1}^{(2)} \mathbf{b}_{s_1}^{(1)} G_{ts_2}^{(2)} \mathbf{b}_{s_2}^{(1)T} \right\} \\
&= m_1 \sum_{s=2}^{t-1} \|\mathbf{b}_s^{(1)}\|^2 (G_{ts}^{(2)})^2 + 2m_1 \sum_{s_2=3}^{t-1} \sum_{s_1=2}^{s_2-1} G_{ts_1}^{(2)} G_{ts_2}^{(2)} \langle \mathbf{b}_{s_1}^{(1)}, \mathbf{b}_{s_2}^{(1)} \rangle \\
&= \tilde{B}_{1Nt} + \tilde{A}_{1Nt}. \tag{3.2.34}
\end{aligned}$$

Notons que  $\tilde{A}_{1Nt} = \sum_{s_2=3}^{t-1} \tilde{A}_{s_2t}^{(12)}$  avec  $\tilde{A}_{s_2t}^{(12)} = \sum_{s_1=2}^{s_2-1} G_{ts_1}^{(2)} G_{ts_2}^{(2)} \langle \mathbf{b}_{s_1}^{(1)}, \mathbf{b}_{s_2}^{(1)} \rangle$  est une somme de différences de martingale par rapport à  $s_2$  puisque  $\mathbb{E}[\tilde{A}_{s_2t}^{(12)} | \mathcal{F}_{\tau-1}] = 0$  où  $\mathcal{F}_{\tau-1}$  est la  $\sigma$ -algèbre engendrée par  $\mathbf{b}_{s_2}$ ,  $s_2 \leq \tau$ . Pour  $t_2 \geq t_1$ ,

$$\begin{aligned}
\mathbb{E}(\tilde{A}_{1Nt_2} \tilde{A}_{1Nt_1}) &= 4m_1^2 \mathbb{E} \left[ \left( \sum_{s_2=3}^{t_2-1} \sum_{s_1=2}^{s_2-1} G_{t_2s_2}^{(2)} G_{t_2s_1}^{(2)} \langle \mathbf{b}_{s_1}^{(1)}, \mathbf{b}_{s_2}^{(1)} \rangle \right) \left( \sum_{s_2=3}^{t_1-1} \sum_{s_1=2}^{s_2-1} G_{t_1s_2}^{(2)} G_{t_1s_1}^{(2)} \langle \mathbf{b}_{s_1}^{(1)}, \mathbf{b}_{s_2}^{(1)} \rangle \right) \right] \\
&= 4m_1^2 \mathbb{E} \left[ \left( \sum_{s_2=3}^{t_1-1} \sum_{s_1=2}^{s_2-1} G_{t_2s_2}^{(2)} G_{t_2s_1}^{(2)} \langle \mathbf{b}_{s_1}^{(1)}, \mathbf{b}_{s_2}^{(1)} \rangle \right) \left( \sum_{s_2=3}^{t_1-1} \sum_{s_1=2}^{s_2-1} G_{t_1s_2}^{(2)} G_{t_1s_1}^{(2)} \langle \mathbf{b}_{s_1}^{(1)}, \mathbf{b}_{s_2}^{(1)} \rangle \right) \right] \\
&\quad + 4m_1^2 \mathbb{E} \left[ \left( \sum_{s_2=t_1}^{t_2-1} \sum_{s_1=2}^{s_2-1} G_{t_2s_2}^{(2)} G_{t_2s_1}^{(2)} \langle \mathbf{b}_{s_1}^{(1)}, \mathbf{b}_{s_2}^{(1)} \rangle \right) \left( \sum_{s_2=3}^{t_1-1} \sum_{s_1=2}^{s_2-1} G_{t_1s_2}^{(2)} G_{t_1s_1}^{(2)} \langle \mathbf{b}_{s_1}^{(1)}, \mathbf{b}_{s_2}^{(1)} \rangle \right) \right].
\end{aligned}$$

Par l'indépendance de  $\mathbf{b}_{s_1}^{(1)}$  et  $\mathbf{b}_{s_2}^{(2)}$ , et puisque  $\mathbb{E}(\mathbf{b}_t^{(h)}) = 0$ , pour  $h = 1, 2$ , le deuxième terme de la dernière somme est nul. On obtient alors

$$\begin{aligned}
\mathbb{E}(\tilde{A}_{1Nt_2} \tilde{A}_{1Nt_1}) &= 4m_1^2 \mathbb{E} \left[ \left( \sum_{s_2=3}^{t_1-1} \sum_{s_1=2}^{s_2-1} G_{t_2s_2}^{(2)} G_{t_2s_1}^{(2)} \langle \mathbf{b}_{s_1}^{(1)}, \mathbf{b}_{s_2}^{(1)} \rangle \right) \left( \sum_{s_2=3}^{t_1-1} \sum_{s_1=2}^{s_2-1} G_{t_1s_2}^{(2)} G_{t_1s_1}^{(2)} \langle \mathbf{b}_{s_1}^{(1)}, \mathbf{b}_{s_2}^{(1)} \rangle \right) \right] \\
&= 4m_1^3 \sum_{s_2=3}^{t_1-1} \sum_{s_1=2}^{s_2-1} \mathbb{E}(G_{t_2s_2}^{(2)} G_{t_2s_1}^{(2)} G_{t_1s_2}^{(2)} G_{t_1s_1}^{(2)}),
\end{aligned}$$

où la dernière égalité est donnée par l'indépendance entre  $\{\mathbf{b}_t^{(1)}\}$  et  $\{\mathbf{b}_s^{(2)}\}$  et que  $\mathbb{E}(\langle \mathbf{b}_{s_1}^{(1)}, \mathbf{b}_{s_2}^{(1)} \rangle) = m_1$ .

Étant donné la relation (3.2.28) et par des manipulations algébriques simples mais longues, on a pour  $t_2 > s_2$ ,  $t_1 > s_1$ ,

$$\left| \mathbb{E}(G_{t_2 s_2}^{(2)} G_{t_2 s_1}^{(2)} G_{t_1 s_2}^{(2)} G_{t_1 s_1}^{(2)}) \right| \leq \begin{cases} \Delta M^2 \left\{ \frac{1}{M} \sum_{j=0}^{N-1} k^2(j/M) \right\}^2 & \text{si } t_2 = t_1, \\ \Delta M \left\{ \frac{1}{M} \sum_{j=0}^{N-1} k^2(j/M) \right\} & \text{si } t_2 > t_1. \end{cases} \quad (3.2.35)$$

Ainsi, on a

$$\frac{1}{M^2} \mathbb{E} \left( \frac{1}{N^2} \sum_{t=3}^N \tilde{A}_{1Nt} \right)^2 = \frac{1}{M^2 N^4} \sum_{t=3}^N \mathbb{E}(\tilde{A}_{1Nt}^2) + \frac{2}{M^2 N^4} \sum_{t_2=4}^N \sum_{t_1=3}^{t_2-1} \mathbb{E}(\tilde{A}_{1Nt_2} \tilde{A}_{1Nt_1})$$

Puisque  $\mathbb{E}(\tilde{A}_{1Nt}^2) \leq M^2 \sum_{s_2=3}^{t-1} \sum_{s_1=2}^{s_2-1} \Delta \leq \Delta M^2 t^2$ , et d'après (3.2.35),  $\mathbb{E}(\tilde{A}_{1Nt_2} \tilde{A}_{1Nt_1}) \leq \Delta M t_1^2$ , on obtient alors

$$\frac{1}{M^2} \mathbb{E} \left( \frac{1}{N^2} \sum_{t=3}^N \tilde{A}_{1Nt} \right)^2 = \mathcal{O}(N^{-1} + M^{-1}). \quad (3.2.36)$$

Maintenant, on considère le premier terme dans (3.2.34), qui peut être écrit comme suit:

$$\tilde{B}_{1Nt} = m_1^2 \sum_{s=2}^{t-1} (G_{ts}^{(2)})^2 + m_1 \sum_{s=2}^{t-1} (\|\mathbf{b}_s^{(1)}\|^2 - m_1) (G_{ts}^{(2)})^2 = \tilde{B}_{2Nt} + \tilde{A}_{2Nt}. \quad (3.2.37)$$

En conditionnant sur  $(\mathbf{b}_t^{(2)})_{t=1}^N$ ,  $\tilde{A}_{2Nt}$  est une somme de suites de variables pondérées, indépendantes et identiquement distribuées. Il s'en suit, par linéarisation de l'espérance, que:

$$\mathbb{E} \tilde{A}_{2Nt}^2 = \sum_{s=2}^{t-1} \mathbb{E}(\|\mathbf{b}_s^{(1)}\|^2 - m_1)^2 \mathbb{E}(G_{ts}^{(2)})^4 \leq \Delta t M^2 \left\{ \frac{1}{M} \sum_{j=0}^{N-1} k^2(j/M) \right\}^2,$$

où  $\mathbb{E}(G_{ts}^{(2)})^4 \leq \Delta M^2 \left\{ \frac{1}{M} \sum_{j=0}^{N-1} k^2(j/M) \right\}^2$  comme on a montré ci dessus en vérifiant (3.2.32). Ainsi, d'après l'inégalité de Minkowski (3.2.11), on peut écrire:

$$\frac{1}{M^2} \mathbb{E} \left( \frac{1}{N^2} \sum_{t=3}^N \tilde{A}_{2Nt} \right)^2 \leq \frac{1}{M^2 N^4} \left\{ \sum_{t=3}^N (\mathbb{E} \tilde{A}_{2Nt}^2)^{1/2} \right\}^2 = \mathcal{O}(N^{-1}). \quad (3.2.38)$$

Maintenant, on considère le premier terme de (3.2.37). D'après (3.2.28), on peut décomposer  $\tilde{B}_{2Nt}$  en deux termes

$$\begin{aligned}
\tilde{B}_{2Nt} &= m_1^2 \sum_{s=2}^{t-1} (G_{ts}^{(2)})^2 \\
&= m_1^2 \sum_{s=2}^{t-1} \left\{ \sum_{j=0}^{s-1} k^2(j/M) \langle \mathbf{b}_{t-j}^{(2)}, \mathbf{b}_{s-j}^{(2)} \rangle \right\}^2 \\
&= m_1^2 \sum_{s=2}^{t-1} \sum_{j=0}^{s-1} k^4(j/M) \mathbf{b}_{t-j}^{(2)T} \mathbf{b}_{s-j}^{(2)} \mathbf{b}_{s-j}^{(2)T} \mathbf{b}_{t-j}^{(2)} \\
&+ 2m_1^2 \sum_{s=2}^{t-1} \sum_{j_2=1}^{s-1} \sum_{j_1=0}^{j_2-1} k^2(j_2/M) k^2(j_1/M) \mathbf{b}_{t-j_2}^{(2)T} \mathbf{b}_{s-j_2}^{(2)} \mathbf{b}_{s-j_1}^{(2)T} \mathbf{b}_{t-j_1}^{(2)} \\
&= \tilde{B}_{3Nt} + \tilde{A}_{3Nt}.
\end{aligned} \tag{3.2.39}$$

avec

$$\begin{aligned}
\tilde{B}_{3Nt} &= m_1^2 \sum_{s=2}^{t-1} \sum_{j=0}^{s-1} k^4(j/M) \mathbf{b}_{t-j}^{(2)T} \mathbf{b}_{s-j}^{(2)} \mathbf{b}_{s-j}^{(2)T} \mathbf{b}_{t-j}^{(2)}, \\
\tilde{A}_{3Nt} &= 2m_1^2 \sum_{s=2}^{t-1} \sum_{j_2=1}^{s-1} \sum_{j_1=0}^{j_2-1} k^2(j_2/M) k^2(j_1/M) \mathbf{b}_{t-j_2}^{(2)T} \mathbf{b}_{s-j_2}^{(2)} \mathbf{b}_{s-j_1}^{(2)T} \mathbf{b}_{t-j_1}^{(2)} \\
&= \sum_{j_1=0}^{t-3} \sum_{j_2=j_1+1}^{t-2} \sum_{s=0}^{t-1} k^2(j_2/M) k^2(j_1/M) \mathbf{b}_{t-j_2}^{(2)T} \mathbf{b}_{s-j_2}^{(2)} \mathbf{b}_{s-j_1}^{(2)T} \mathbf{b}_{t-j_1}^{(2)}.
\end{aligned}$$

Or, en utilisant le fait que  $\text{tr}(\mathbf{b}_{t-j_2}^{(2)T} \mathbf{b}_{s-j_2}^{(2)} \mathbf{b}_{s-j_1}^{(2)T} \mathbf{b}_{t-j_1}^{(2)}) = \text{tr}(\mathbf{b}_{t-j_1}^{(2)} \mathbf{b}_{t-j_2}^{(2)T} \mathbf{b}_{s-j_2}^{(2)} \mathbf{b}_{s-j_1}^{(2)T})$ , on écrit

$$\tilde{A}_{3Nt} = \text{tr} \left( \sum_{j_1=0}^{t-3} k^2(j_1/M) \mathbf{b}_{t-j_1}^{(2)} \sum_{j_2=j_1+1}^{t-2} k^2(j_2/M) \mathbf{b}_{t-j_2}^{(2)T} \sum_{s=0}^{t-1} \mathbf{b}_{s-j_2}^{(2)} \mathbf{b}_{s-j_1}^{(2)T} \right).$$

Ainsi, notons que  $\tilde{A}_{3Nt}$  est une somme de différences de martingales sur  $j_1$ . Il s'en suit que:

$$\begin{aligned}
\mathbb{E}\tilde{A}_{3Nt}^2 &\leq 4m_1^4 \sum_{j_1=0}^{t-3} k^4(j_1/M) \mathbb{E} \left( \sum_{j_2=j_1+1}^{t-2} k^2(j_2/M) \|\mathbf{b}_{t-j_2}^{(2)}\| \sum_{s=j_2+1}^{t-1} \|\mathbf{b}_{s-j_1}^{(2)}\| \|\mathbf{b}_{s-j_2}^{(2)}\| \right)^2 \\
&\leq 4m_1^4 m_2^2 \sum_{j_1=1}^{t-3} k^4(j_1/M) \left[ \sum_{j_2=j_1+1}^{t-2} k^2(j_2/M) \left\{ \mathbb{E} \left( \sum_{s=j_2+1}^{t-1} \|\mathbf{b}_{s-j_1}^{(2)}\| \|\mathbf{b}_{s-j_2}^{(2)}\| \right)^2 \right\}^{1/2} \right]^2 \\
&\leq \Delta t M^3 \left\{ \frac{1}{M} \sum_{j_1=1}^{N-1} k^4(j_1/M) \right\} \left\{ \frac{1}{M} \sum_{j_2=1}^{N-1} k^2(j_2/M) \right\}^2 \\
&= \mathcal{O}(tM^3),
\end{aligned}$$

où la première inégalité provient du fait que  $tr(AB) \leq tr(A)tr(B)$ , la deuxième inégalité est obtenue en utilisant l'inégalité de Minkowski (3.2.11) et la troisième, par le fait que  $\mathbb{E}(\sum_{s=j_2+1}^{t-1} \|\mathbf{b}_{s-j_1}^{(2)}\| \|\mathbf{b}_{s-j_2}^{(2)}\|)^2 \leq \Delta t$  pour  $t > s > j_2 > j_1$ . Ainsi,

$$\mathbb{E}(\tilde{A}_{3Nt}^2) = \mathcal{O}(tM^3),$$

et donc:

$$\frac{1}{M^2} \mathbb{E} \left( \frac{1}{N^2} \sum_{t=3}^N \tilde{A}_{3Nt} \right)^2 \leq \frac{1}{M^2 N^4} \left\{ \sum_{t=3}^N \mathbb{E}^{1/2}(\tilde{A}_{3Nt}^2) \right\}^2 = \mathcal{O}(M/N). \quad (3.2.40)$$

Finalement, d'après le premier terme de (3.2.39), on a:

$$\begin{aligned}
\tilde{B}_{3Nt} &= m_1^2 m_2^2 \sum_{s=2}^{t-1} \sum_{j=0}^{s-1} k^4(j/M) + m_1^2 \sum_{s=2}^{t-1} \sum_{j=0}^{s-1} k^4(j/M) (\mathbf{b}_{t-j}^{(2)T} \mathbf{b}_{s-j}^{(2)} \mathbf{b}_{s-j}^{(2)T} \mathbf{b}_{t-j}^{(2)} - m_2^2) \\
&= \frac{m_1 m_2}{4} \mathbb{E}(W_{1Nt}^2) + \tilde{A}_{4Nt}.
\end{aligned} \quad (3.2.41)$$

À l'aide de manipulations algébriques similaires à celles utilisées ci dessus, on obtient

$$\mathbb{E}(\tilde{A}_{4Nt})^2 \leq \Delta t^2 M.$$

Il s'en suit que:

$$\frac{1}{M^2 N^4} \mathbb{E} \left( \sum_{t=3}^N \tilde{A}_{4Nt} \right)^2 \leq \frac{1}{M^2 N^4} \left\{ \sum_{t=3}^N \mathbb{E}^{1/2}(\tilde{A}_{4Nt}^2) \right\}^2 = \mathcal{O}(M^{-1}). \quad (3.2.42)$$

Ainsi,  $\ddot{W}_{1Nt}^2$  s'écrit sous la forme suivante

$$\ddot{W}_{1Nt}^2 = \frac{m_1 m_2}{4} \mathbb{E}(W_{1Nt}^2) + \sum_{l=1}^4 \tilde{A}_{lNt},$$

et si on combine les relations (3.2.36), (3.2.38), (3.2.40) et (3.2.42), on obtient

$$\begin{aligned} \frac{1}{M^2} \text{var} \left( \frac{1}{N^2} \sum_{t=3}^N \ddot{W}_{1Nt}^2 \right) &\leq \frac{\Delta}{M^2} \sum_{l=1}^4 \mathbb{E} \left( \frac{1}{N^2} \sum_{t=3}^N \tilde{A}_{lNt} \right)^2 \\ &= \mathcal{O}(N^{-1} + M^{-1}) + \mathcal{O}(N^{-1}) + \mathcal{O}(M/N) + \mathcal{O}(M^{-1}) \\ &= \mathcal{O}(M/N + M^{-1}). \end{aligned}$$

Un résultat similaire peut être démontré par symétrie pour  $\ddot{W}_{2Nt}^2$ . Ainsi la condition (3.2.27) est vérifiée pour  $M/N \rightarrow 0$  et  $M \rightarrow \infty$ .  $\square$

Il s'en suit alors d'après le théorème 3.2.1 de Brown que:

$$\sigma^{-1}(N)W_N \rightarrow N(0, 1).$$

$\square$

Ce qui complète la démonstration du lemme 3.2.4 et ainsi du lemme 3.2.1.  $\square$

PREUVE DU LEMME 3.2.2:

Pour ce lemme, on considère les mêmes notations que celles du lemme 3.2.1. Ainsi, en utilisant une transformation similaire à celle utilisée dans la preuve du lemme 3.2.1, soit  $\hat{\mathbf{b}}_t = \Sigma^{-1/2} \hat{\mathbf{a}}_t$  et en notant que

$$\mathbf{C}_{\hat{\mathbf{b}}}^{(12)}(j) = \Sigma_1^{-1/2} \mathbf{C}_{\hat{\mathbf{a}}}^{(12)}(j) \Sigma_2^{-1/2},$$

on peut écrire

$$\begin{aligned}
\mathcal{T}(\hat{\mathbf{a}}, \boldsymbol{\Sigma}) &= N \sum_{j=1-N}^{N-1} k^2(j/M) \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)^T (\boldsymbol{\Sigma}_2^{-1} \otimes \boldsymbol{\Sigma}_1^{-1}) \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j) \\
&= N \sum_{j=1-N}^{N-1} k^2(j/M) \mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j)^T \mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) \\
&= \mathcal{T}(\hat{\mathbf{b}}, \boldsymbol{\Sigma}).
\end{aligned}$$

Ainsi, on obtient

$$\mathcal{T}(\hat{\mathbf{a}}, \boldsymbol{\Sigma}) - \mathcal{T}(\mathbf{a}, \boldsymbol{\Sigma}) = \mathcal{T}(\hat{\mathbf{b}}, \boldsymbol{\Sigma}) - \mathcal{T}(\mathbf{b}, \boldsymbol{\Sigma}), \quad (3.2.43)$$

et soit à montrer que:

$$\mathcal{T}(\hat{\mathbf{b}}, \boldsymbol{\Sigma}) - \mathcal{T}(\mathbf{b}, \boldsymbol{\Sigma}) = o_p(M^{1/2}). \quad (3.2.44)$$

Notons que dans cette preuve, on ne va pas utiliser le fait que la matrice de covariance du processus  $\mathbf{b}$  est l'identité. Ce résultat reste valide pour tout processus innovation de matrice de covariance quelconque. L'utilisation du processus  $\mathbf{b}$  n'est que pour alléger les calculs.

Pour cela on va d'abord décomposer  $\mathcal{T}(\hat{\mathbf{b}}, \boldsymbol{\Sigma}) - \mathcal{T}(\mathbf{b}, \boldsymbol{\Sigma})$  sous la forme suivante

$$\begin{aligned}
\mathcal{T}(\hat{\mathbf{b}}, \boldsymbol{\Sigma}) - \mathcal{T}(\mathbf{b}, \boldsymbol{\Sigma}) &= N \sum_{j=1-N}^{N-1} k^2(j/M) \{ \mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j)^T \mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) - \mathbf{c}_{\mathbf{b}}^{(12)}(j)^T \mathbf{c}_{\mathbf{b}}^{(12)}(j) \} \\
&= N \sum_{j=1-N}^{N-1} k^2(j/M) (\mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) - \mathbf{c}_{\mathbf{b}}^{(12)}(j))^T (\mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) + \mathbf{c}_{\mathbf{b}}^{(12)}(j)) \\
&\quad + 2N \sum_{j=1-N}^{N-1} k^2(j/M) \mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j)^T \{ \mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) - \mathbf{c}_{\mathbf{b}}^{(12)}(j) \}.
\end{aligned}$$

Alors il suffit de montrer que:

$$N \sum_{j=1-N}^{N-1} k^2(j/M) \|\mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) - \mathbf{c}_{\mathbf{b}}^{(12)}(j)\|^2 = o_p(M^{1/2}), \quad (3.2.45)$$

et

$$N \sum_{j=1-N}^{N-1} k^2(j/M) \langle \mathbf{c}_{\mathbf{b}}^{(12)}(j), \mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) - \mathbf{c}_{\mathbf{b}}^{(12)}(j) \rangle = o_p(M^{1/2}). \quad (3.2.46)$$

Notons que les sommes sur  $j$  peuvent être décomposées en deux sommes, l'une par rapport aux  $j$  positifs et l'autre par rapport aux  $j$  négatifs. Par symétrie, on ne va considérer que les sommes sur les  $j$  positifs, car la preuve sera semblable quand les indices  $j$  sont négatifs.

Montrons d'abord (3.2.45), i.e.

$$T_N^{(1)} = N \sum_{j=0}^{N-1} k^2(j/M) \|\mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) - \mathbf{c}_{\mathbf{b}}^{(12)}(j)\|^2 = o_p(M^{1/2}). \quad (3.2.47)$$

Pour cela, notons par:

$$\hat{\boldsymbol{\delta}}_t = \mathbf{b}_t^{(1)} - \hat{\mathbf{b}}_t^{(1)} \quad ; \quad \hat{\boldsymbol{\eta}}_t = \mathbf{b}_t^{(2)} - \hat{\mathbf{b}}_t^{(2)}, \quad (3.2.48)$$

et donc

$$\begin{aligned} T_N^{(1)} &= N \sum_{j=0}^{N-1} k^2(j/M) \|\mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) - \mathbf{c}_{\mathbf{b}}^{(12)}(j)\|^2 \\ &= N \sum_{j=0}^{N-1} k^2(j/M) \|\mathbf{C}_{\hat{\mathbf{b}}}^{(12)}(j) - \mathbf{C}_{\mathbf{b}}^{(12)}(j)\|_2^2 \\ &= N \sum_{j=0}^{N-1} k^2(j/M) \left\| \frac{1}{N} \sum_{t=j+1}^N (\mathbf{b}_t^{(1)} \mathbf{b}_{t-j}^{(2)T} - \hat{\mathbf{b}}_t^{(1)} \hat{\mathbf{b}}_{t-j}^{(2)T}) \right\|_2^2. \end{aligned}$$

En utilisant l'équation (3.2.48) et l'inégalité de Cauchy-Shwartz, on obtient

$$\begin{aligned} T_N^{(1)} &= N \sum_{j=0}^{N-1} k^2(j/M) \left\| \frac{1}{N} \sum_{t=j+1}^N (\mathbf{b}_t^{(1)} \hat{\boldsymbol{\eta}}_{t-j}^T + \hat{\boldsymbol{\delta}}_t \mathbf{b}_{t-j}^{(2)T} - \hat{\boldsymbol{\delta}}_t \hat{\boldsymbol{\eta}}_{t-j}^T) \right\|_2^2 \\ &\leq 4N (T_{1N} + T_{2N} + T_{3N}), \end{aligned} \quad (3.2.49)$$

avec

$$T_{1N} = \sum_{j=0}^{N-1} k^2(j/M) \left\| \frac{1}{N} \sum_{t=j+1}^N \mathbf{b}_t^{(1)} \hat{\boldsymbol{\eta}}_{t-j}^T \right\|_2^2, \quad (3.2.50)$$

$$T_{2N} = \sum_{j=0}^{N-1} k^2(j/M) \left\| \frac{1}{N} \sum_{t=j+1}^N \hat{\boldsymbol{\delta}}_t \mathbf{b}_{t-j}^{(2)T} \right\|_2^2, \quad (3.2.51)$$

$$T_{3N} = \sum_{j=0}^{N-1} k^2(j/M) \left\| \frac{1}{N} \sum_{t=j+1}^N \hat{\boldsymbol{\delta}}_t \hat{\boldsymbol{\eta}}_{t-j}^T \right\|_2^2. \quad (3.2.52)$$

On va montrer que chacun de ces termes  $T_{jN}$  est  $o_p(M^{1/2}/N)$  pour  $j = 1, 2, 3$ . Rappelons que les techniques qu'on va utiliser dans cette preuve sont similaires à celles utilisées dans la preuve du théorème 2.3.1.

Notons d'abord que:

$$\hat{\boldsymbol{\eta}}_t = \{\boldsymbol{\Phi}(p_2) - \hat{\boldsymbol{\Phi}}(p_2)\} \mathbf{X}_t^{(2)}(p_2) + \boldsymbol{\xi}_t(p_2),$$

avec

$$\boldsymbol{\xi}_t(p_2) = \sum_{l=p_2+1}^{\infty} \boldsymbol{\Phi}_l^{(2)} \mathbf{X}_{t-l}^{(2)},$$

qui représente le biais de l'approximation du processus  $\{\mathbf{X}_t^{(2)}\}$  par un processus autorégressif d'ordre fini  $p_2$ .

Selon (2.3.14), on a

$$\mathbb{E}(\|\boldsymbol{\xi}_t(p_h)\|^2) = O\left(\sum_{l=p_h+1}^{\infty} \|\boldsymbol{\Phi}_l^{(h)}\|_2\right)^2, \quad h = 1, 2.$$

De même, le théorème 1.3.1 dit que

$$\|\hat{\boldsymbol{\Phi}}(p_h) - \boldsymbol{\Phi}(p_h)\|_2^2 = O_p\left\{\frac{p_h}{N}\right\}, \quad h = 1, 2.$$

En utilisant le résultat (2.3.16) du chapitre 2, on peut déduire que

$$\begin{aligned} T_{1N} &= \sum_{j=0}^{N-1} k^2(j/M) \left\| \frac{1}{N} \sum_{t=j+1}^N \hat{\boldsymbol{\eta}}_t \mathbf{b}_{t-j}^{(2)T} \right\|_2^2 \\ &= O_p\left(\frac{p_1^2 M}{N^2}\right) \left\{ \frac{1}{M} \sum_{j=0}^{N-1} k^2(j/M) \right\} \end{aligned}$$

Puisque  $p_1^2 = o\left(\frac{N}{M^{1/2}}\right)$ , on a que

$$T_{1N} = o_p\left(\frac{M^{1/2}}{N}\right).$$

Par symétrie, on peut aussi écrire que:

$$T_{2N} = o_p(M^{1/2}/N).$$

Pour le troisième terme, en utilisant l'inégalité de Cauchy-Schwarz, on obtient:

$$\begin{aligned} T_{3N} &= \sum_{j=0}^{N-1} k^2(j/M) \left\| N^{-1} \sum_{t=j+1}^N \hat{\boldsymbol{\eta}}_t \hat{\boldsymbol{\delta}}_{t-j}^T \right\|_2^2 \\ &\leq \|\boldsymbol{\Phi}(p_1) - \hat{\boldsymbol{\Phi}}(p_1)\|_2^2 \|\boldsymbol{\Phi}(p_2) - \hat{\boldsymbol{\Phi}}(p_2)\|_2^2 \sum_{j=0}^{N-1} k^2(j/M) \left\| N^{-1} \sum_{t=j+1}^N \mathbf{X}_t^{(1)}(p_1) \mathbf{X}_{t-j}^{(2)}(p_2)^T \right\|_2^2 \\ &+ \|\boldsymbol{\Phi}(p_1) - \hat{\boldsymbol{\Phi}}(p_1)\|_2^2 \sum_{j=0}^{N-1} k^2(j/M) \left\| N^{-1} \sum_{t=j+1}^N \mathbf{X}_t^{(1)}(p_1) \boldsymbol{\xi}_{t-j}(p_2)^T \right\|_2^2 \\ &+ \|\boldsymbol{\Phi}(p_2) - \hat{\boldsymbol{\Phi}}(p_2)\|_2^2 \sum_{j=0}^{N-1} k^2(j/M) \left\| N^{-1} \sum_{t=j+1}^N \boldsymbol{\xi}_t(p_1) \mathbf{X}_{t-j}^{(2)}(p_2)^T \right\|_2^2 \\ &+ \sum_{j=0}^{N-1} k^2(j/M) \left\| N^{-1} \sum_{t=j+1}^N \boldsymbol{\xi}_t(p_1) \boldsymbol{\xi}_{t-j}(p_2)^T \right\|_2^2 \end{aligned}$$

Comme on a montré pour  $T_{1N}$ , et en utilisant les résultats du chapitre 2, (2.3.18), (2.3.20), (2.3.21) et (2.3.22), on obtient

$$\sum_{j=0}^{N-1} k^2(j/M) \mathbb{E} \left\| N^{-1} \sum_{t=j+1}^N \mathbf{X}_t^{(1)}(p_1) \mathbf{X}_{t-j}^{(2)}(p_2)^T \right\|_2^2 \leq \frac{\Delta M p_1 p_2}{N},$$

$$\sum_{j=0}^{N-1} k^2(j/M) \mathbb{E} \|N^{-1} \sum_{t=j+1}^N \mathbf{X}_t^{(1)}(p_1) \boldsymbol{\xi}_{t-j}(p_2)^T\|_2^2 \leq \frac{\Delta M p_1}{N} \left( \sum_{l=p_2+1}^{\infty} \|\Phi_l^{(2)}\|_2 \right)^2,$$

$$\sum_{j=0}^{N-1} k^2(j/M) \mathbb{E} \|N^{-1} \sum_{t=j+1}^N \boldsymbol{\xi}_t(p_1) \mathbf{X}_{t-j}^{(2)}(p_2)^T\|_2^2 \leq \frac{\Delta M p_2}{N} \left( \sum_{l=p_1+1}^{\infty} \|\Phi_l^{(1)}\|_2 \right)^2,$$

$$\sum_{j=0}^{N-1} k^2(j/M) \mathbb{E} \|N^{-1} \sum_{t=j+1}^N \boldsymbol{\xi}_t(p_1) \boldsymbol{\xi}_{t-j}(p_2)^T\|_2^2 \leq \frac{\Delta M}{N} \left( \sum_{l=p_1+1}^{\infty} \|\Phi_l^{(1)}\|_2 \right)^2 \left( \sum_{l=p_2+1}^{\infty} \|\Phi_l^{(2)}\|_2 \right)^2.$$

Ainsi, en utilisant les conditions  $p_h = o(N^{1/2}/M^{1/4})$  et  $N \sum_{l=p_h+1}^{\infty} \|\Phi_l^{(h)}\|_2^2 = o(N^{1/2}/M^{1/4})$ , on déduit que

$$T_{3N} = o_p(M^{1/2}/N),$$

Ainsi, si on réunit les trois égalités ci dessus, on obtient:

$$T_N^{(1)} = N \sum_{j=0}^{N-1} k^2(j/M) \|\mathbf{c}_b^{(12)}(j) - \mathbf{c}_b^{(12)}(j)\|^2 = o_p(M^{1/2}). \quad (3.2.53)$$

En utilisant l'équation (3.2.48) et l'inégalité de Cauchy-Schwarz, le membre de gauche de (3.2.46) devient

$$\begin{aligned} T_N^{(2)} &= N \sum_{j=1-N}^{N-1} k^2(j/M) \langle \mathbf{c}_b^{(12)}(j), \mathbf{c}_b^{(12)}(j) - \mathbf{c}_b^{(12)}(j) \rangle \\ &\leq N \sum_{j=1-N}^{N-1} k^2(j/M) \|\mathbf{c}_b^{(12)}(j)\| \|\mathbf{c}_b^{(12)}(j) - \mathbf{c}_b^{(12)}(j)\| \\ &\leq N \left\{ \sum_{j=1-N}^{N-1} k^2(j/M) \|\mathbf{c}_b^{(12)}(j)\|^2 \right\}^{1/2} \left\{ \sum_{j=1-N}^{N-1} k^2(j/M) \|\mathbf{c}_b^{(12)}(j) - \mathbf{c}_b^{(12)}(j)\|^2 \right\}^{1/2}. \end{aligned}$$

Or d'après (3.2.53), on a

$$\sum_{j=0}^{N-1} k^2(j/M) \|\mathbf{c}_b^{(12)}(j) - \mathbf{c}_b^{(12)}(j)\|^2 = O_p(M/N^2) = o_p(M^{1/2}/N).$$

Et aussi, par application de l'inégalité de Markov,  $P(|X| > \epsilon) \leq \frac{\mathbb{E}(|X|)}{\epsilon}$  où  $X$  est une variable aléatoire et  $\epsilon > 0$ , on a

$$\sum_{j=1}^{N-1} k^2(j/M) \|\mathbf{c}_b^{(12)}(j)\|^2 = O_p(M/N), \quad (3.2.54)$$

et puisque  $M/N \rightarrow 0$ , on obtient alors

$$T_N^{(2)} = N \sum_{j=1-N}^{N-1} k^2(j/M) \langle \mathbf{c}_b^{(12)}(j), \mathbf{c}_b^{(12)}(j) - \mathbf{c}_b^{(12)}(j) \rangle = o_p(M^{1/2}).$$

Ainsi la deuxième condition est réalisée, ce qui complète la preuve de le lemme 3.2.2. □

# Chapitre 4

---

## PUISSANCE ASYMPTOTIQUE

### 4.1. INTRODUCTION

À chaque fois que deux tests sont disponibles pour une même hypothèse, il est naturel de comparer leurs fonctions de puissance, et cette comparaison devient intéressante quand les deux tests donnent des résultats distincts. Pour des tests d'hypothèses en séries chronologiques, on ne peut en général utiliser ce genre de comparaisons puisque souvent, seulement la théorie asymptotique sous l'hypothèse nulle est disponible. Ce type de problèmes n'est pas récent, et plusieurs auteurs ont utilisé différentes méthodes pour obtenir une meilleure comparaison du comportement asymptotique de la statistique du test sous les hypothèses alternatives.

Puisque le test d'indépendance, entre deux processus linéaires multivariés que l'on a proposé au chapitre 3 est basé sur une statistique pondérée par un noyau, la mesure de l'efficacité relative d'un test par rapport à un autre paraît nécessaire pour comparer ces tests pour différents noyaux. Dans ce chapitre, nous utiliserons deux méthodes de comparaison, une due à Pitman (1979) et l'autre due à Bahadur (1960). Pour chacune des deux approches, on calculera l'efficacité relative asymptotique d'un test par rapport à un autre. On introduira le concept de la pente approximée de Bahadur. L'approche de Pitman sera utilisée dans

l'étude de la puissance asymptotique locale, et celle de Bahadur, dans le cas de la puissance asymptotique globale. On présentera la distribution asymptotique de la statistique du test sous une classe d'alternatives locales. Par la suite, on trouvera le noyau qui maximise la fonction de puissance asymptotique. Finalement, dans le cas de la puissance asymptotique globale, on montrera la convergence de la statistique  $\mathcal{Q}_N$  sous l'hypothèse de dépendance des deux processus et on calculera la pente asymptotique de Bahadur.

## 4.2. PUISSANCE ASYMPTOTIQUE LOCALE

Pour des tests convergents, la puissance asymptotique converge vers 1 quand  $N \rightarrow \infty$  sous l'hypothèse alternative  $\mathcal{H}_1$  pour tout niveau  $0 < \alpha < 1$ . Pour avoir une puissance inférieure à 1, on peut fixer le niveau et faire varier l'hypothèse alternative tout en s'approchant de  $\mathcal{H}_0$  quand  $N \rightarrow \infty$ . Ainsi, nous allons nous intéresser à la puissance asymptotique locale de la statistique  $\mathcal{Q}_N$  sous une classe déterminée d'hypothèses alternatives locales. Considérons la suite suivante d'alternatives:

$$\mathcal{H}_{\varphi N} : \mathbf{S}_a^{(12)}(w) = \varphi(N)\mathbf{G}(w), \quad w \in [-\pi, \pi],$$

où  $\mathbf{S}_a^{(12)}(w)$  est la fonction de cohérence croisée définie par (1.1.10) et (1.1.11) entre les deux processus  $\{\mathbf{a}_t^{(1)}\}$  et  $\{\mathbf{a}_t^{(2)}\}$ , qu'on peut écrire sous la forme suivante

$$\mathbf{S}_a^{(12)}(w) = \sum_{j=-\infty}^{\infty} \Sigma_1^{-1/2} \Gamma_a^{(12)}(j) \Sigma_2^{-1/2} e^{-iwj}.$$

La fonction  $\mathbf{G}$  est une fonction matricielle complexe définie sur  $[-\pi, \pi]$  telle que  $\varphi(N)\mathbf{G}(w)$  est une fonction de cohérence pour chaque  $N$  et  $\varphi(N) \rightarrow 0$  lorsque  $N \rightarrow \infty$ . Ainsi, l'hypothèse alternative locale  $\mathcal{H}_{\varphi N}$  converge vers  $\mathcal{H}_0$  quand

$N \rightarrow \infty$ . Aussi, notons que l'hypothèse  $\mathcal{H}_{\varphi N}$  peut s'écrire aussi sous une forme vectorielle

$$\mathcal{H}_{\varphi N} : \mathbf{s}_{\mathbf{a}}^{(12)}(w) = \varphi(N)\mathbf{g}(w), \quad w \in [-\pi, \pi],$$

avec  $\mathbf{s}_{\mathbf{a}}^{(12)}(w) = \text{vec}(\mathbf{S}_{\mathbf{a}}^{(12)}(w))$  et  $\mathbf{g}(w) = \text{vec}(\mathbf{G}(w))$ . Puisque  $\text{vec}(ABC) = (C \otimes A)\text{vec}(B)$ , on a

$$\mathbf{s}_{\mathbf{a}}^{(12)}(w) = \sum_{j=-\infty}^{\infty} (\Sigma_2 \otimes \Sigma_1)^{-1/2} \gamma_{\mathbf{a}}^{(12)}(j) e^{-iwj}.$$

Si on définit la norme quadratique suivante  $\|\cdot\|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\cdot|^2 dw$ , où  $|\cdot|^2$  représente le module d'un nombre complexe, on a

$$\|\mathbf{s}_{\mathbf{a}}^{(12)}\|^2 = \sum_{j=-\infty}^{\infty} \gamma_{\mathbf{a}}^{(12)}(j)^T (\Sigma_2 \otimes \Sigma_1)^{-1} \gamma_{\mathbf{a}}^{(12)}(j), \quad (4.2.1)$$

en utilisant la relation  $\frac{1}{2\pi} \int_{-\pi}^{\pi} |e^{-iwj}|^2 dw = 1$ . Aussi, basé sur la série résiduelle  $\{\hat{\mathbf{a}}_t : t = 1, \dots, N\}$ , un estimateur de  $\mathbf{s}_{\mathbf{a}}^{(12)}(w)$  pondéré par un noyau appartenant à la classe  $\mathcal{K}_1$ , est donné par

$$\hat{\mathbf{s}}_{\hat{\mathbf{a}}}^{(12)}(w) = \sum_{j=1-N}^{N-1} k(j/M) \left( \mathbf{C}_{\hat{\mathbf{a}}}^{(22)}(0) \otimes \mathbf{C}_{\hat{\mathbf{a}}}^{(11)}(0) \right)^{-1/2} \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j) e^{-iwj} \quad (4.2.2)$$

Dans tout ce qui suit, on ne va utiliser que la notation vectorielle. Dans ce contexte, la dépendance de  $\mathbf{s}_{\mathbf{a}}^{(12)}(w)$  sur  $N$  est faite de façon implicite pour simplifier la notation. De (4.2.2), il est facile de vérifier que  $\|\hat{\mathbf{s}}_{\hat{\mathbf{a}}}^{(12)}\|^2$ , qui est la distance carrée entre la fonction de cohérence estimée  $\hat{\mathbf{s}}_{\hat{\mathbf{a}}}^{(12)}(w)$  et la fonction de cohérence sous l'hypothèse nulle  $\mathcal{H}_0$  qui est  $\mathbf{s}_{\mathbf{a}}^{(12)}(w) = 0$ , s'écrit

$$\|\hat{\mathbf{s}}_{\hat{\mathbf{a}}}^{(12)}\|^2 = \sum_{j=1-N}^{N-1} k^2(j/M) \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)^T \left( \mathbf{C}_{\hat{\mathbf{a}}}^{(22)}(0) \otimes \mathbf{C}_{\hat{\mathbf{a}}}^{(11)}(0) \right)^{-1} \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j).$$

Ainsi, selon (3.2.2), la statistique  $\mathcal{Q}_N$  est donnée par

$$\mathcal{Q}_N = \frac{N\|\mathbf{s}_{\hat{\mathbf{a}}}^{(12)}\|^2 - m_1 m_2 S_N(k)}{\sqrt{2m_1 m_2 D_N(k)}}.$$

Afin d'étudier la puissance locale, nous allons comme dans Hong (1996), voir aussi Gallant et Jorgenson (1979) et Gallant et White (1988), considérer la distance carrée entre  $\mathbf{s}_{\hat{\mathbf{a}}}^{(12)}(w)$  et  $\mathbf{s}_{\mathbf{a}}^{(12)}(w)$  spécifiée par l'alternative  $\mathcal{H}_{\varphi N}$ , ce qui nous amène à la statistique

$$\mathcal{Q}_N^\varphi = \frac{N\|\mathbf{s}_{\hat{\mathbf{a}}}^{(12)} - \mathbf{s}_{\mathbf{a}}^{(12)}\|^2 - m_1 m_2 S_N(k)}{\{2m_1 m_2 D_N(k)\}^{1/2}},$$

où  $S_N(k)$  et  $D_N(k)$  sont définies respectivement par (3.1.8) et (3.1.9).

Le théorème suivant nous donne la distribution asymptotique de  $\mathcal{Q}_N^\varphi$ .

**Théorème 4.2.1.**

*Supposons que les hypothèses 2.3.1 et 3.1.1 sont satisfaites, et que les deux processus  $\{\mathbf{a}_t^{(1)}\}$  et  $\{\mathbf{a}_s^{(2)}\}$  satisfont la suite d'hypothèses  $\mathcal{H}_{\varphi N}$ . Soient  $M \rightarrow \infty$ ,  $M/N \rightarrow 0$  et  $p_h$ ,  $h = 1, 2$ , deux suites d'entiers positifs qui satisfont les conditions suivantes:*

$$p_h = o\left(\frac{N^{1/2}}{M^{1/4}}\right), \quad N \sum_{l=p_h+1}^{\infty} \|\Phi_l^{(h)}\|_2^2 = o\left(\frac{N^{1/2}}{M^{1/4}}\right).$$

*Si  $\varphi(N) = M^{1/4}/N^{1/2}$ , alors*

$$\mathcal{Q}_N^\varphi \xrightarrow{d} \mathcal{N}(\mu(k), 1),$$

où

$$\mu(k) = \frac{\|\mathbf{g}\|^2}{\{2m_1 m_2 D(k)\}^{1/2}}.$$

PREUVE DU THÉORÈME 4.2.1.

Sous  $\mathcal{H}_{\varphi N}$ ,  $\mathbf{s}_{\hat{\mathbf{a}}}^{(12)}(w) = \varphi(N)\mathbf{g}(w)$ , et par définition de la norme  $\mathcal{L}^2$ , on peut écrire

$$\begin{aligned} \|\mathbf{s}_{\hat{\mathbf{a}}}^{(12)} - \mathbf{s}_{\hat{\mathbf{a}}}^{(12)}\|^2 &= \|\mathbf{s}_{\hat{\mathbf{a}}}^{(12)}\|^2 + \varphi^2(N)\|\mathbf{g}\|^2 \\ &\quad - 2\varphi(N)\operatorname{Re} \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \langle \mathbf{s}_{\hat{\mathbf{a}}}^{(12)*}(w), \mathbf{g}(w) \rangle dw \right) \end{aligned} \quad (4.2.3)$$

où  $\operatorname{Re}(z)$  désigne la partie réelle de  $z$  et  $z^*$  est la valeur conjuguée du complexe  $z$  et  $\langle, \rangle$  désigne le produit scalaire. Considérons d'abord la dernière partie de l'équation ci dessus. On a

$$\begin{aligned} \mathbf{s}_{\hat{\mathbf{a}}}^{(12)}(w) &= \sum_{j=1-N}^{N-1} k(j/M) \left( \mathbf{C}_{\hat{\mathbf{a}}}^{(22)}(0) \otimes \mathbf{C}_{\hat{\mathbf{a}}}^{(11)}(0) \right)^{-1/2} \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j) e^{-iwj} \\ &= \sum_{j=1-N}^{N-1} k(j/M) \left\{ \left( \mathbf{C}_{\hat{\mathbf{a}}}^{(22)}(0) \otimes \mathbf{C}_{\hat{\mathbf{a}}}^{(11)}(0) \right)^{-1/2} - (\boldsymbol{\Sigma}_2 \otimes \boldsymbol{\Sigma}_1)^{-1/2} \right\} \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j) e^{-iwj} \\ &\quad + \sum_{j=1-N}^{N-1} k(j/M) (\boldsymbol{\Sigma}_2 \otimes \boldsymbol{\Sigma}_1)^{-1/2} \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j) e^{-iwj}. \end{aligned}$$

En considérant le bruit blanc décrit dans la preuve du lemme 3.2.1,  $\mathbf{b}_t = \boldsymbol{\Sigma}^{-1/2} \mathbf{a}_t$ , et dont la matrice de covariance est l'identité, et si on note par

$$\mathbf{D}_{\hat{\mathbf{b}}}^{(12)}(w) = \sum_{j=1-N}^{N-1} k(j/M) \mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) e^{-iwj}, \quad \boldsymbol{\beta}(N) = \mathbf{O}_p(N^{-1/2}) (\boldsymbol{\Sigma}_2 \otimes \boldsymbol{\Sigma}_1)^{1/2} + \mathbb{I}$$

où  $\mathbb{I}$  désigne la matrice identité, on peut écrire en utilisant la relation,  $\mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) = (\boldsymbol{\Sigma}_2 \otimes \boldsymbol{\Sigma}_1)^{-1/2} \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j)$ , que

$$\begin{aligned}
\mathbf{s}_{\hat{\mathbf{a}}}^{(12)}(w) &= \sum_{j=1-N}^{N-1} k(j/M) \mathbf{O}_p(N^{-1/2}) \mathbf{c}_{\hat{\mathbf{a}}}^{(12)}(j) e^{-iwj} + \sum_{j=1-N}^{N-1} k(j/M) \mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) e^{-iwj} \\
&= \sum_{j=1-N}^{N-1} k(j/M) \mathbf{O}_p(N^{-1/2}) (\boldsymbol{\Sigma}_2 \otimes \boldsymbol{\Sigma}_1)^{1/2} \mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) e^{-iwj} \\
&\quad + \sum_{j=1-N}^{N-1} k(j/M) \mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) e^{-iwj} \\
&= \boldsymbol{\beta}(N) \mathbf{D}_{\hat{\mathbf{b}}}^{(12)}(w).
\end{aligned}$$

Considérons le dernier terme de l'égalité (4.2.3). Pour montrer que

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \langle \mathbf{s}_{\hat{\mathbf{a}}}^{(12)*}(w), \mathbf{g}(w) \rangle dw = o_p(M^{1/4}/N^{1/2}),$$

il suffit alors de vérifier que  $\frac{1}{2\pi} \int_{-\pi}^{\pi} \langle \mathbf{D}_{\hat{\mathbf{b}}}^{(12)*}(w), \mathbf{g}(w) \rangle dw = o_p(M^{1/4}/N^{1/2})$ . Ainsi, en décomposant  $\mathbf{D}_{\hat{\mathbf{b}}}^{(12)*}(w)$  en deux termes,

$$\mathbf{D}_{\hat{\mathbf{b}}}^{(12)*}(w) = \sum_{j=1-N}^{N-1} k(j/M) (\mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) - \mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j)) e^{iwj} + \sum_{j=1-N}^{N-1} k(j/M) \mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) e^{iwj},$$

et en notant par  $\mathbf{g}_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{g} e^{ijw} dw$ , on obtient que

$$\begin{aligned}
\frac{1}{2\pi} \int_{-\pi}^{\pi} \langle \mathbf{D}_{\hat{\mathbf{b}}}^{(12)*}(w), \mathbf{g}(w) \rangle dw &= \sum_{j=1-N}^{N-1} k(j/M) (\mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) - \mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j))^T \mathbf{g}_j \\
&\quad + \sum_{j=1-N}^{N-1} k(j/M) \mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j)^T \mathbf{g}_j.
\end{aligned}$$

Puisque  $\mathbb{E}\|\mathbf{c}_{\mathbf{b}}^{(12)}(j)\|_2^2 \leq \frac{\Delta}{N}$  et  $\|\mathbf{g}_j\|_2^2 < \infty$ , on a alors

$$\begin{aligned} \mathbb{E} \left| \sum_{j=1-N}^{N-1} k(j/M) \mathbf{c}_{\mathbf{b}}^{(12)}(j)^T \mathbf{g}_j \right|^2 &\leq \frac{\Delta}{N} \sum_{j=1-N}^{N-1} \|\mathbf{g}_j\|_2^2 k^2(j/M) \\ &= O(N^{-1}). \end{aligned}$$

Aussi, en utilisant l'inégalité de Cauchy-Schwarz, on obtient

$$\left| \sum_{j=1-N}^{N-1} k(j/M) (\mathbf{c}_{\mathbf{b}}^{(12)}(j) - \mathbf{c}_{\mathbf{b}}^{(12)}(j))^T \mathbf{g}_j \right| \leq \|\mathbf{g}\|^2 \left\{ \sum_{j=1-N}^{N-1} k^2(j/M) \|\mathbf{c}_{\mathbf{b}}^{(12)}(j) - \mathbf{c}_{\mathbf{b}}^{(12)}(j)\|_2^2 \right\}^{1/2}.$$

Or, d'après l'équation (3.2.45) du chapitre 3, on a

$$\sum_{j=1-N}^{N-1} k^2(j/M) \|\mathbf{c}_{\mathbf{b}}^{(12)}(j) - \mathbf{c}_{\mathbf{b}}^{(12)}(j)\|_2^2 = o_p(M^{1/2}/N).$$

On obtient alors

$$\left| \sum_{j=1-N}^{N-1} k(j/M) (\mathbf{c}_{\mathbf{b}}^{(12)}(j) - \mathbf{c}_{\mathbf{b}}^{(12)}(j))^T \mathbf{g}_j \right| = o_p(M^{1/4}/N^{1/2})$$

et

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} \langle \mathbf{D}_{\mathbf{b}}^{(12)*}(w), \mathbf{g}(w) \rangle dw &= O_p(N^{-1/2}) + o_p(M^{1/4}/N^{1/2}) \\ &= o_p(M^{1/4}/N^{1/2}). \end{aligned}$$

Finalement, quand  $\varphi(N) = M^{1/4}/N^{1/2}$ , on obtient

$$\|\mathbf{s}_{\hat{\mathbf{a}}}^{(12)} - \mathbf{s}_{\mathbf{a}}^{(12)}\|^2 = \|\mathbf{s}_{\hat{\mathbf{a}}}^{(12)}\|^2 + \|\mathbf{g}\|^2 M^{1/2}/N + o_p(M^{1/2}/N).$$

Il s'en suit que

$$\begin{aligned} Q_N^\varphi &= \frac{N \|\mathbf{s}_{\hat{\mathbf{a}}}^{(12)} - \mathbf{s}_{\mathbf{a}}^{(12)}\|^2 - m_1 m_2 S_N(k)}{\{2m_1 m_2 D_N(k)\}^{1/2}} \\ &= \frac{N \|\mathbf{s}_{\hat{\mathbf{a}}}^{(12)}\|^2 - m_1 m_2 S_N(k)}{\{2m_1 m_2 D_N(k)\}^{1/2}} + \frac{\|\mathbf{g}\|^2 M^{1/2}}{\{2m_1 m_2 D_N(k)\}^{1/2}} + \frac{o_p(M^{1/2})}{\{2m_1 m_2 D_N(k)\}^{1/2}} \end{aligned}$$

D'après le théorème 3.2.2 et la relation (3.2.4), on peut conclure que

$$Q_N^\varphi \xrightarrow{d} \mathcal{N}(\mu(k), 1),$$

$$\text{où } \mu(k) = \frac{\|g\|^2}{\{2m_1 m_2 D(k)\}^{1/2}}. \quad \square$$

Le test  $Q_N^\varphi$  permet de discriminer entre  $\mathcal{H}_0$  et une classe d'hypothèses alternatives locales qui convergent vers  $\mathcal{H}_0$  à la vitesse  $\varphi(N) = \frac{M^{1/4}}{N^{1/2}}$ . Plus  $M$  croît lentement, plus le test est puissant. Ceci est en contraste avec le fait que l'approximation la loi normale asymptotique s'améliore lorsque l'ordre autorégressif  $p_h$  croît rapidement. Puisque la vitesse  $\frac{M^{1/4}}{N^{1/2}}$  croît moins rapidement que la vitesse paramétrique  $N^{-1/2}$ , notre test est moins efficace que celui d'El Himdi et Roy (1997) sous l'hypothèse  $\mathcal{H}_{\varphi N}$ . En effet, El Himdi et Roy (1997) supposent un  $M$  arbitraire mais fixe. C'est le prix à payer pour obtenir un test convergent quand on considère une large classe d'alternatives.

D'après le théorème 3.2.1, la puissance asymptotique du test  $Q_N^\varphi$  au niveau  $\alpha \in (0, 1)$  est donnée par

$$\lim_{N \rightarrow \infty} \mathbb{P}(Q_N^\varphi > z_\alpha) = 1 - \Phi\{z_\alpha - \mu(k)\}, \quad (4.2.4)$$

où  $\Phi$  est la fonction de distribution de la  $\mathcal{N}(0, 1)$ , et  $z_\alpha$  est le quantile d'ordre  $\alpha$  de la loi normale  $\mathcal{N}(0, 1)$ . Cette puissance est une fonction du noyau  $k$ .

Supposons que  $M = N^\nu$  où  $0 < \nu < 1$ . D'après Pitman (1979, chap. 7) et Ghosh et Huang (1991, p. 1007), pour deux tests utilisant les deux noyaux  $k_1$  et  $k_2$ , l'efficacité relative asymptotique de Pitman de  $k_2$  par rapport à  $k_1$  est donnée par

$$ERA_P(k_2, k_1) = \left\{ \frac{D(k_1)}{D(k_2)} \right\}^{\frac{1}{2-\nu}}. \quad (4.2.5)$$

Par exemple, d'après Hong (1996), l'efficacité relative du noyau de Bartlett  $k_B(z) = (1 - |z|)\mathbb{I}\{|z| \leq 1\}$  par rapport au noyau tronqué  $k_T(z) = \mathbb{I}\{|z| \leq 1\}$  est

$$ERA_P(k_B, k_T) = 5^{1/(2-\nu)} > 5^{1/2} \simeq 2.23$$

pour tout  $0 < \nu < 1$ , et  $\mathbb{I}$  est la fonction indicatrice telle que définie au chapitre 3. Ainsi,  $k_B$  est plus efficace que  $k_T$ , ce dernier noyau correspond au test d'El Himdi et Roy (1997). D'autres noyaux sont aussi plus puissants que le noyau tronqué. Nous considérons maintenant le noyau optimal qui maximise la puissance de  $\mathcal{Q}_N^{\mathcal{Q}}$  dans une classe de noyaux. Notons d'abord que le biais asymptotique de l'estimateur par noyau dépend du comportement ou du lissage du noyau au point zéro et aussi du lissage de la matrice de densité spectrale au même point. D'après Parzen (1957), on définit

$$k_r := \lim_{z \rightarrow 0} \left\{ \frac{1 - k(z)}{|z|^r} \right\} \text{ pour } r \in [0, \infty).$$

Plus le noyau est lisse au point zéro, plus la valeur de  $r$  pour laquelle  $k_r$  est fini est grande. Si  $r$  est un entier pair, on a

$$k_r = -\frac{1}{r!} \frac{d^r k(z)}{dz^r} \Big|_{z=0},$$

et  $k_r < \infty$  si et seulement si  $k(z)$  est  $r$  fois différentiable au point zéro. *L'exposant caractéristique* de la fonction  $k(z)$  est le plus grand entier  $r \geq 0$  tel que  $k_r$  est fini et non nul. Par exemple, pour le noyau tronqué,  $k_r = 0$  pour tout  $r < \infty$ , avec le noyau de Bartlett,  $k_1 = 1$ ,  $k_r = 0$  pour  $r < 1$  et  $k_r = \infty$  pour  $r > 1$ . Pour les noyaux de Parzen, Tukey-Hamming et Bartlett-Priestley,  $k_2 = 6$ ,  $\pi^2/4$ , et 1.42, respectivement,  $k_r = 0$  pour  $r < 2$ , et  $k_r = \infty$  pour  $r > 2$  et ceci pour les trois noyaux. Ces définitions s'appliquent aussi à la fonction de densité spectrale.

Nous considérons maintenant la classe suivante de noyaux, avec  $r = 2$

$$\kappa(\tau) = \left\{ k \in \mathcal{K}_1, k_2 = \frac{1}{2}\tau^2, K(\lambda) \geq 0, \lambda \in \mathbb{R} \right\},$$

où  $K(\lambda)$  est la transformée de Fourier du noyau  $k(z)$ , appelée aussi *fenêtre génératrice spectrale* définie par

$$K(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} k(z) e^{-iz\lambda} dz.$$

Cette famille de noyaux contient tous les noyaux qui appartiennent à  $\mathcal{K}_1$  qui génèrent, pour des échantillons finis, un estimateur semi défini positif. Entre autres, les noyaux de Daniell, Parzen et Bartlett-Priestley appartiennent à  $\kappa(\tau)$ , mais cette dernière ne contient pas le noyau tronqué, le noyau de Bartlett ni le noyau de Tukey-Hamming. La condition  $k_2 = \frac{1}{2}\tau^2$  joue un rôle de normalisation, puisque le biais de l'estimateur par noyau dépend de  $k_2$ . Ainsi, pour le même paramètre de lissage, noté ici  $M$ , tous les noyaux appartenant à  $\kappa(\tau)$  vont donner le même biais asymptotique et donc la comparaison sera équitable. Le théorème qui suit, dont la preuve se trouve dans Hong (1996), nous donne le noyau qui maximise la puissance asymptotique dans la classe de noyaux  $\kappa(\tau)$ .

#### **Théorème 4.2.2.**

*Supposons que les conditions de théorème 4.2.1 sont réalisées. Alors le noyau de Daniell*

$$k_D(z) = \frac{\sin\left(3^{\frac{1}{2}}\tau z\right)}{\left(3^{\frac{1}{2}}\tau z\right)}, \quad z \in (-\infty, \infty)$$

*maximise la puissance asymptotique de  $\mathcal{Q}_N^o$  sur  $\kappa(\tau)$ .*

Dans un contexte d'estimation de densité spectrale, c'est le noyau de Bartlett-Priestley qui maximise la puissance asymptotique selon différents critères d'erreur quadratique moyenne.

### 4.3. PUISSANCE ASYMPTOTIQUE GLOBALE

Tandis que l'analyse locale de la puissance nous fournit des informations utiles sur l'efficacité des tests au voisinage de  $\mathcal{H}_0$ , il reste qu'elle ne nous donne pas un compte complet sur l'ensemble des propriétés asymptotiques de la puissance. C'est ce qui nous a poussé à examiner la puissance asymptotique globale de la statistique  $\mathcal{Q}_N$ . Pour établir le théorème de convergence sous l'hypothèse alternative de dépendance  $\mathcal{H}_1$ , on va imposer une condition classique de dépendance entre les deux processus linéaires multivariés  $\{\mathbf{a}_t^{(1)}\}$  et  $\{\mathbf{a}_t^{(2)}\}$ . Considérons ainsi l'hypothèse suivante:

#### Hypothèse 4.3.1.

*Les deux processus innovations  $\{\mathbf{a}_t^{(1)}\}$  et  $\{\mathbf{a}_t^{(2)}\}$  sont stationnaires d'ordre quatre et vérifient les conditions suivantes:*

$$\sum_{j=-\infty}^{+\infty} \|\Gamma_{\mathbf{a}}^{(12)}(j)\|_2^2 < \infty, \quad \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} |\kappa_{uvuv}(0, i, j, l)| < \infty,$$

où  $\kappa_{uvuv}(0, i, j, l)$  est le cumulants d'ordre quatre de la distribution de  $(a_{u,t}^{(1)}, a_{v,t+i}^{(2)}, a_{u,t+j}^{(1)}, a_{v,t+l}^{(2)})$ .

Sous cette hypothèse, la suite  $(\Gamma_{\mathbf{a}}^{(12)}(j))_{j=-\infty}^{\infty}$  n'a pas besoin d'être absolument sommable, comme dans le cas des processus à mémoire longue pour lesquels la densité spectrale n'existe pas à la fréquence  $w = 0$ . La condition sur les cumulants est standard dans la littérature des séries chronologiques multivariées, voir Hannan (1970, p. 211) et Anderson (1971, p. 465). Elle caractérise la dépendance temporelle du couple  $(\mathbf{a}_t^{(1)}, \mathbf{a}_t^{(2)})$ . Brillinger (1981, p. 26) fait même une hypothèse plus forte en supposant que les cumulants de tous les ordres existent et sont absolument sommables. À prime abord, quand  $(\mathbf{a}_t^{(1)}, \mathbf{a}_t^{(2)})$  a une

loi jointe normale bivariée, la condition sur les cumulants devient triviale puisque  $\kappa_{uvuv}(0, i, j, l) = 0, \forall i, j, l$ . De même, selon Hannan (1970, p. 211), tous les processus linéaires stationnaires à l'ordre quatre, dont les coefficients sont absolument sommables et dont les moments d'ordre quatre des innovations existent, satisfont la condition des cumulants de l'hypothèse 4.3.1.

**Théorème 4.3.1.**

*Supposons que les hypothèses 2.3.1 et 4.3.1 sont vérifiées. Soient  $M = M(N) \rightarrow \infty$ ,  $M/N \rightarrow 0$  et  $p_h$  avec  $h = 1, 2$ , deux suites d'entiers positifs qui satisfont les conditions suivantes:*

$$p_h^2 = o\left(\frac{N}{M}\right) \quad , \quad \sum_{j=p_h+1}^{\infty} \|\Phi_j^{(h)}\|_2^2 = o(M^{-1}).$$

*Alors, nous avons,*

$$\left(\frac{M^{1/2}}{N}\right) \mathcal{Q}_N \xrightarrow{p} \frac{\|\mathbf{s}_x^{(12)}\|^2}{\{2m_1m_2D(k)\}^{1/2}}. \quad (4.3.1)$$

PREUVE DU THÉORÈME 4.3.1.

D'après (3.2.3), on peut écrire

$$\begin{aligned} \left(\frac{M^{1/2}}{N}\right) \mathcal{Q}_N &= \frac{M^{1/2}\|\mathbf{s}_a^{(12)}\|^2 - \left(\frac{M^{1/2}}{N}\right) m_1m_2S_N(k)}{\{2m_1m_2D(k)\}^{1/2}} \\ &= \frac{\|\mathbf{s}_a^{(12)}\|^2}{\{2m_1m_2M^{-1}D_N(k)\}^{1/2}} - \frac{N^{-1}S_N(k)}{\{2M^{-1}D_N(k)\}^{1/2}} (m_1m_2)^{1/2}. \end{aligned} \quad (4.3.2)$$

En appliquant les résultats de convergence (3.2.4), le deuxième terme de l'égalité ci dessus converge vers zéro. Par l'invariance de la fonction de cohérence sous les transformations linéaires, voir Priestley (1981, p. 662), on a  $\|\mathbf{s}_x^{(12)}\| = \|\mathbf{s}_a^{(12)}\|$  où  $\|\mathbf{s}_a^{(12)}\|$  est définie par (4.2.1). En utilisant la transformation utilisée au chapitre 3,  $\mathbf{b}_t = \Sigma^{-1/2}\mathbf{a}_t$ , on a donc  $\|\mathbf{s}_a^{(12)}\| = \|\mathbf{s}_b^{(12)}\|$ . Ainsi, montrer (4.3.1) revient à

montrer que

$$\|\mathbf{s}_{\hat{\mathbf{b}}}^{(12)}\|^2 - \|\mathbf{s}_{\mathbf{b}}^{(12)}\|^2 \xrightarrow{p} 0.$$

qui découle des deux lemmes suivants.  $\square$

Avant d'énoncer les deux lemmes, notons que  $\tilde{\mathbf{s}}_{\mathbf{b}}^{(12)}(w)$  est défini de la même façon que  $\mathbf{s}_{\hat{\mathbf{b}}}^{(12)}(w)$  en remplaçant les séries résiduelles  $(\hat{\mathbf{b}}_t^{(1)}, \hat{\mathbf{b}}_t^{(2)})_{t=1}^N$  par les séries innovations  $(\mathbf{b}_t^{(1)}, \mathbf{b}_t^{(2)})_{t=1}^N$ .

**Lemme 4.3.1.**

*Sous les hypothèses du théorème 4.3.1, on a :*

$$\|\mathbf{s}_{\hat{\mathbf{b}}}^{(12)}\|^2 - \|\tilde{\mathbf{s}}_{\mathbf{b}}^{(12)}\|^2 \xrightarrow{p} 0$$

**Lemme 4.3.2.**

*Sous les hypothèses du théorème 4.3.1, on a :*

$$\|\tilde{\mathbf{s}}_{\mathbf{b}}^{(12)}\|^2 - \|\mathbf{s}_{\mathbf{b}}^{(12)}\|^2 \xrightarrow{p} 0$$

**PREUVE DU LEMME 4.3.1.**

Par définition de  $\mathbf{s}_{\hat{\mathbf{b}}}^{(12)}$  et  $\tilde{\mathbf{s}}_{\mathbf{b}}^{(12)}$  et en utilisant les mêmes développements qu'au chapitre 3, on a

$$\begin{aligned} \|\mathbf{s}_{\hat{\mathbf{b}}}^{(12)}\|^2 - \|\tilde{\mathbf{s}}_{\mathbf{b}}^{(12)}\|^2 &= \sum_{j=1-N}^{N-1} k^2(j/M) (\|\mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j)\|_2^2 - \|\mathbf{c}_{\mathbf{b}}^{(12)}(j)\|_2^2) \\ &= \sum_{j=1-N}^{N-1} k^2(j/M) \|\mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) - \mathbf{c}_{\mathbf{b}}^{(12)}(j)\|_2^2 \\ &\quad + 2 \sum_{j=1-N}^{N-1} k^2(j/M) \langle \mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j), \mathbf{c}_{\hat{\mathbf{b}}}^{(12)}(j) - \mathbf{c}_{\mathbf{b}}^{(12)}(j) \rangle. \end{aligned}$$

Il suffit alors de montrer que le premier terme converge en probabilité vers 0, car en utilisant l'inégalité de Cauchy-Schwarz, le deuxième terme s'écrit comme

produit du premier terme et d'une quantité finie.

En utilisant les notations du chapitre 3, et l'équation (3.2.49), on peut écrire

$$\sum_{j=1-N}^{N-1} k^2(j/M) \|\mathbf{c}_{\mathbf{b}}^{(12)}(j) - \mathbf{c}_{\mathbf{b}}^{(12)}(j)\|_2^2 \leq 4 \sum_{l=1}^3 T_{lN},$$

où  $T_{lN}$ ,  $l = 1, 2, 3$ , sont définies par (3.2.50), (3.2.51) et (3.2.52) respectivement.

Commençons par montrer que  $T_{1N} \rightarrow 0$  en probabilité. En appliquant l'inégalité de Cauchy-Schwarz, on obtient

$$\begin{aligned} T_{1N} &= \sum_{j=0}^{N-1} k^2(j/M) \left\| \sum_{t=j+1}^N \mathbf{b}_t^{(1)} \hat{\boldsymbol{\eta}}_{t-j}^T \right\|_2^2 \\ &\leq M \left\{ \frac{1}{M} \sum_{j=0}^{N-1} k^2(j/M) \right\} \left\{ \frac{1}{N} \sum_{t=1}^N \|\mathbf{b}_t^{(1)}\|_2^2 \right\} \left\{ \frac{1}{N} \sum_{t=1}^N \|\hat{\boldsymbol{\eta}}_{t-j}\|_2^2 \right\}. \end{aligned}$$

Or, par (3.2.48), on a  $\hat{\boldsymbol{\eta}}_t = \mathbf{b}_t^{(2)} - \hat{\mathbf{b}}_t^{(2)}$  et il s'en suit que

$$\begin{aligned} \frac{1}{N} \sum_{t=j+1}^N \|\hat{\boldsymbol{\eta}}_{t-j}\|_2^2 &= \frac{1}{N} \sum_{t=j+1}^N \|\mathbf{b}_{t-j}^{(2)} - \hat{\mathbf{b}}_{t-j}^{(2)}\|_2^2 \\ &\leq \frac{1}{N} \sum_{t=j+1}^N \{ \|\boldsymbol{\Phi}(p_2) - \hat{\boldsymbol{\Phi}}(p_2) \mathbf{X}_{t-j}^{(2)}(p_2)\|_2^2 + \|\sum_{l=p_2+1}^{\infty} \boldsymbol{\Phi}_l^{(2)} \mathbf{X}_{t-l}^{(2)}\|_2^2 \} \\ &= O_p\left(\frac{p_2^2}{N}\right) + O_p(1) \sum_{l=p_2+1}^{\infty} \|\boldsymbol{\Phi}_l^{(2)}\|_2^2. \end{aligned}$$

Ainsi, en invoquant les hypothèses sur le processus  $\mathbf{b}$ , sur  $p_2$  et sur les paramètres  $(\boldsymbol{\Phi}_l^{(2)})$ , on conclut que

$$\begin{aligned} T_{1N} &= O_p\left(\frac{Mp_2^2}{N}\right) + O_p\left(M \sum_{l=p_2+1}^{\infty} \|\boldsymbol{\Phi}_l^{(2)}\|_2^2\right) \\ &= o_p(1). \end{aligned}$$

Par symétrie, on montre que  $T_{2N} = o_p(1)$ . Pour  $T_{3N}$ , on écrit

$$\begin{aligned} T_{3N} &= \sum_{j=0}^{N-1} k^2(j/M) \left\| \frac{1}{N} \sum_{t=j+1}^N \hat{\delta}_t \hat{\eta}_{t-j}^T \right\|_2^2 \\ &\leq M \left\{ \frac{1}{M} \sum_{j=0}^{N-1} k^2(j/M) \right\} \left\{ \frac{1}{N} \sum_{t=1}^N \|\hat{\delta}_t\|_2^2 \right\} \left\{ \frac{1}{N} \sum_{t=1}^N \|\hat{\eta}_{t-j}\|_2^2 \right\}. \end{aligned}$$

Par symétrie, on montre que  $\frac{1}{N} \sum_{t=1}^N \|\hat{\delta}_t\|_2^2 = O_p\left(\frac{p_1^2}{N}\right) + O_p(1) \sum_{l=p_1+1}^{\infty} \|\Phi_l^{(1)}\|_2^2$ , et ainsi, en utilisant les mêmes hypothèses que celles utilisées pour  $T_{1N}$ , on déduit que  $T_{3N} = o_p(1)$ . Finalement, on conclut que

$$\|\mathbf{s}_{\hat{\mathbf{b}}}^{(12)}\|^2 - \|\tilde{\mathbf{s}}_{\mathbf{b}}^{(12)}\|^2 = o_p(1).$$

Ce qui termine la preuve du lemme 4.3.1. □

#### PREUVE DU LEMME 4.3.2.

Pour montrer ce lemme, on effectuera des développements similaires à ceux de la preuve précédente. Par définition de  $\|\tilde{\mathbf{s}}_{\mathbf{b}}^{(12)}\|^2$  et  $\|\mathbf{s}_{\mathbf{b}}^{(12)}\|^2$ , on écrit

$$\begin{aligned} \|\tilde{\mathbf{s}}_{\mathbf{b}}^{(12)}\|^2 - \|\mathbf{s}_{\mathbf{b}}^{(12)}\|^2 &= \sum_{j=1-N}^{N-1} k^2(j/M) \|\mathbf{c}_{\mathbf{b}}^{(12)}(j)\|_2^2 - \sum_{-\infty}^{\infty} \|\gamma_{\mathbf{b}}^{(12)}(j)\|_2^2 \\ &= \sum_{j=1-N}^{N-1} k^2(j/M) [\|\mathbf{c}_{\mathbf{b}}^{(12)}(j)\|_2^2 - \|\gamma_{\mathbf{b}}^{(12)}(j)\|_2^2] \\ &\quad + \sum_{j=1-N}^{N-1} (k^2(j/M) - 1) \|\gamma_{\mathbf{b}}^{(12)}(j)\|_2^2 + \sum_{|j|>N} \|\gamma_{\mathbf{b}}^{(12)}(j)\|_2^2. \end{aligned}$$

Le deuxième terme de la dernière égalité tend vers 0 quand  $M \rightarrow \infty$  d'après le théorème de convergence dominé dont les conditions sont vérifiées par les hypothèses 3.1.1 et 4.3.1. Le dernier terme converge aussi vers 0 quand  $N \rightarrow \infty$  car  $\sum_{j=-\infty}^{+\infty} \|\Gamma_{\mathbf{X}}^{(12)}(j)\|_2^2 < \infty$ . Donc il reste à montrer que le premier terme converge

en probabilité vers 0. Puisque

$$\begin{aligned} \sum_{j=1-N}^{N-1} k^2(j/M) [\|\mathbf{c}_b^{(12)}(j)\|_2^2 - \|\boldsymbol{\gamma}_b^{(12)}(j)\|_2^2] &= \sum_{j=1-N}^{N-1} k^2(j/M) \|\mathbf{c}_b^{(12)}(j) - \boldsymbol{\gamma}_b^{(12)}(j)\|_2^2 \\ &+ 2 \sum_{j=1-N}^{N-1} k^2(j/M) \lambda(j). \end{aligned}$$

où  $\lambda(j) = \langle \boldsymbol{\gamma}_b^{(12)}(j), \mathbf{c}_b^{(12)}(j) - \boldsymbol{\gamma}_b^{(12)}(j) \rangle$ . Comme dans la preuve du lemme 4.3.1, il suffit alors de vérifier que  $\sum_{j=1-N}^{N-1} k^2(j/M) \|\mathbf{c}_b^{(12)}(j) - \boldsymbol{\gamma}_b^{(12)}(j)\|_2^2 = \sum_{u,v=1}^{m_1, m_2} \sum_{j=-N+1}^{N-1} k^2(j/M) \{c_{uv}^{(12)}(j) - \gamma_{uv}^{(12)}(j)\}^2 = o_p(1)$  où  $c_{uv}^{(12)}(j)$  et  $\gamma_{uv}^{(12)}(j)$  désignent l'élément  $(u, v)$  des matrices  $\mathbf{C}_b^{(12)}(j)$  et  $\boldsymbol{\Gamma}_b^{(12)}(j)$  respectivement.

Étant données les hypothèses 2.3.1, 4.3.1 et d'après Priestley (1981, p. 325-26), on a

$$Var\{c_{uv}^{(12)}(j)\} = \frac{1}{N} \sum_{i=-N+1}^{N-1} (1 - \frac{|i|}{N}) \{\gamma_{uv}^{(12)}(i+j)\gamma_{uv}^{(12)}(i-j) + \kappa_{uvuv}(0, j, i, i+j)\}.$$

Il s'en suit, d'après les hypothèses 3.1.1 et 4.3.1, que

$$\begin{aligned} \sum_{j=-N+1}^{N-1} k^2(j/M) Var\{c_{uv}^{(12)}(j)\} &= \frac{1}{N} \sum_{j=-N+1}^{N-1} k^2(j/M) \\ &\times \sum_{i=-N+1}^{N-1} (1 - \frac{|i|}{N}) \gamma_{uv}^{(12)}(i+j)\gamma_{uv}^{(12)}(i-j) \\ &+ \frac{1}{N} \sum_{j=-N+1}^{N-1} k^2(j/M) \sum_{i=-N+1}^{N-1} (1 - \frac{|i|}{N}) \kappa_{uvuv}(0, j, i, i+j) \\ &= O(\frac{M}{N} + \frac{1}{N}). \end{aligned}$$

Ainsi, on conclut que  $\sum_{j=1-N}^{N-1} k^2(j/M) \|\mathbf{c}_b^{(12)}(j) - \boldsymbol{\gamma}_b^{(12)}(j)\|_2^2 = o_p(1)$ . Ce qui prouve le lemme 3.2.  $\square$

Le théorème 4.3.1 implique que  $\mathcal{Q}_N$  tend vers l'infini à la vitesse  $N/M^{1/2}$  sous des alternatives fixes. Asymptotiquement, plus  $M$  croît lentement, plus  $\mathcal{Q}_N$  diverge rapidement vers l'infini, et plus le test  $\mathcal{Q}_N$  est puissant. Cette conclusion est la même dans le cas d'alternatives locales.

Pour comparer l'efficacité de deux tests sous des alternatives fixes, le critère de Pitman (1979) est inapproprié car la puissance de  $Q_N$  s'approche de 1 quand  $N \rightarrow \infty$  pour tout niveau  $\alpha \in (0, 1)$ .

Ce qui nous mène à utiliser le critère de la pente asymptotique de Bahadur, qui est pertinent pour des tests basés sur des échantillons de grandes tailles et sous des alternatives fixes. La pente asymptotique de Bahadur est définie comme étant la vitesse de convergence du niveau de signification expérimental ou valeur- $p$ , asymptotique vers 0 quand  $N \rightarrow \infty$ . Et puisque  $Q_N$  est asymptotiquement  $\mathcal{N}(0, 1)$  sous l'hypothèse nulle, la probabilité de rejet asymptotique est donnée par  $1 - \Phi(Q_N)$ .

Si on définit maintenant

$$\aleph_N(k) = -2 \ln \{1 - \Phi(Q_N)\}.$$

D'après Bahadur (1960, p. 283),  $\ln \{1 - \Phi(\xi)\} = -\frac{1}{2}\xi^2 \{1 + o(1)\}$  quand  $\xi \rightarrow \infty$ , et par (4.3.2), puisque le deuxième terme de cette équation tend vers zéro, on obtient

$$\left(\frac{M^{1/2}}{N}\right) \aleph_N(k) \rightarrow \frac{\|\mathbf{s}_x^{(12)}\|^2}{\{2m_1m_2D(k)\}^{1/2}} \quad (4.3.3)$$

en probabilité sous des alternatives fixes quand  $M \rightarrow \infty$  et  $M/N \rightarrow 0$ . La quantité  $\frac{\|\mathbf{s}_x^{(12)}\|^2}{\{2m_1m_2D(k)\}^{1/2}}$  est la pente asymptotique de  $Q_N$  selon Bahadur (1960).

Si on suppose que  $M = N^\nu$ , l'efficacité relative asymptotique de Bahadur entre deux noyaux sera ainsi donnée par

$$ERA_B(k_2, k_1) = \left\{ \frac{D(k_1)}{D(k_2)} \right\}^{\frac{1}{2-\nu}}. \quad (4.3.4)$$

Pour le calcul de  $ERA_B(k_2, k_1)$ , on fait un raisonnement similaire à celui pour l'efficacité relative asymptotique de Pitman  $ERA_P(k_2, k_1)$ . En fait

$$ERA_B(k_2, k_1) = \lim_{N_1 \rightarrow \infty, N_2 \rightarrow \infty} \left( \frac{N_2}{N_1} \right) \quad (4.3.5)$$

où  $N_2$  est la taille échantionnale nécessaire afin que le deuxième test basé sur le noyau  $k_2$  ait la même puissance asymptotique que le premier test basé sur le noyau  $k_1$  avec une taille échantionnale  $N_1$ . D'après (4.3.3) et puisque  $M = N^\nu$ , on déduit alors facilement (4.3.4).

Une grande pente asymptotique implique une vitesse rapide de convergence vers 0 de la valeur-p asymptotique quand  $N \rightarrow \infty$ . En plus, la vitesse de divergence de  $\aleph_N(k)$  vers l'infini est d'ordre  $N^2/M$  et Bahadur (1960) a souligné que cette vitesse est plus rapide que celle des tests paramétriques incluant ceux basés sur la normalité asymptotique et aussi les tests  $\chi^2$ , qui sont d'ordre  $N$ .

## Chapitre 5

---

### ÉTUDE PAR SIMULATION

Nous avons proposé dans les chapitres 2 et 3 deux types de procédures pour tester l'indépendance de deux séries multivariées autorégressives d'ordre infini VAR( $\infty$ ). Pour les deux procédures, on suppose que les deux séries multivariées sont générées par des processus VAR( $\infty$ ) stationnaires et inversibles. Après une procédure de filtrage par une autorégression finie qui permet d'estimer les paramètres autorégressifs pour chaque série, les séries résiduelles résultantes sont utilisées pour le calcul des statistiques de tests  $Q_{\hat{\mathbf{a}}}^{(12)}(j)$ ,  $Q_{\hat{\mathbf{a}},M}^{(12)}$  et  $\mathcal{Q}_N$  pour différents noyaux, tailles échantillonales et différentes valeurs du paramètre de troncature  $M$ . Afin d'alléger la notation, les statistiques  $Q_{\hat{\mathbf{a}}}^{(12)}(j)$  et  $Q_{\hat{\mathbf{a}},M}^{(12)}$  seront notées par  $Q_{\hat{\mathbf{a}}}(j)$  et  $Q_{\hat{\mathbf{a}},M}$  respectivement. Les distributions des statistiques  $Q_{\hat{\mathbf{a}}}(j)$  et  $Q_{\hat{\mathbf{a}},M}$  sont approximées par des distributions du khi-deux et celle de  $\mathcal{Q}_N$  par une distribution  $\mathcal{N}(0, 1)$ . Dans ce chapitre, nous nous proposons d'étudier la précision de l'approximation des distributions exactes des statistiques  $\mathcal{Q}_N$  et  $Q_{\hat{\mathbf{a}},M}$  par les distributions asymptotiques correspondantes. Pour la statistique  $Q_{\hat{\mathbf{a}},M}$  les distributions exactes des statistiques des tests  $\mathcal{H}_0^M : \boldsymbol{\rho}_{\mathbf{X}}^{(12)}(j) = \mathbf{0}, |j| \leq M$ , est estimée pour les trois percentiles  $\alpha = 0.01, 0.05$  et  $0.10$ . Pour  $\mathcal{Q}_N$ , l'hypothèse  $\mathcal{H}_0 : \boldsymbol{\rho}_{\mathbf{X}}^{(12)}(j) = \mathbf{0}, |j| \leq N - 1$ , est aussi estimée pour les trois percentiles. Pour étudier l'influence de la taille échantillonnale et du paramètre de troncature  $M$

sur la distribution de ces statistiques, différentes tailles seront considérées ainsi que différentes valeurs de  $M$ .

## 5.1. MÉTHODOLOGIE DE LA SIMULATION.

### 5.1.1. Choix des modèles de processus.

Pour étudier la performance des statistiques  $Q_{\hat{a},M}$ ,  $Q_N$  et  $Q_N^*$  introduites dans les chapitres précédents, nous avons simulé des réalisations suivant les trois modèles de processus stationnaires et inversibles  $\mathbf{AR}(1)$ ,  $\mathbf{MA}(1)$ ,  $\mathbf{AR}_\delta(1)$  décrits dans le tableau 5.2. Les deux premiers modèles nous permettront de déterminer le comportement des différentes statistiques de tests sous l'hypothèse nulle d'indépendance des deux séries. Le modèle  $\mathbf{AR}_\delta(1)$  représente un modèle autorégressif  $\mathbf{AR}(1)$  qui décrit une alternative directe de dépendance entre les deux processus multivariés stationnaires  $\mathbf{X}^{(1)}$  et  $\mathbf{X}^{(2)}$  et qui permettra d'évaluer la puissance des tests. La matrice de covariance croisée dépend d'un paramètre  $\delta$  qui prend les valeurs  $\delta = 1.0, 1.5$  et  $2$ , et qui représente la dépendance instantanée entre les deux processus innovations. Les valeurs des paramètres des différents modèles sont données dans le tableau 5.2.

### 5.1.2. Description de la simulation

On va se limiter à des séries bivariées  $\{\mathbf{X}_t^{(1)}\}$  et  $\{\mathbf{X}_t^{(2)}\}$  stationnaires et inversibles. On va générer pour chaque modèle des séries de longueurs  $N = 100$  et  $200$ . Nous avons utilisé le logiciel Splus, voir Venables et Ripley (2000) pour programmer une fonction qui génère des innovations gaussiennes vectorielles  $\mathbf{a}_1, \dots, \mathbf{a}_N$  de matrice de covariance  $\Sigma_{\mathbf{a}}$ . Ainsi, une réalisation de longueur  $N$  du processus conjoint  $(\mathbf{X}_t^{(1)}, \mathbf{X}_t^{(2)})$  se déduit directement à partir de récurrences définissant le

modèle de processus  $\text{MA}(1)$ . Pour les deux autres modèles  $\text{AR}(1)$ ,  $\text{AR}_\delta(1)$ , afin d'avoir une même structure de corrélation pour les valeurs initiales de la série  $\{\mathbf{X}_t\}$  que pour le reste des observations,  $\mathbf{X}_1$  est générée suivant sa distribution exacte qui est gaussienne centrée de covariance  $\mathbf{\Gamma}_\mathbf{X}(0)$  à l'aide de l'algorithme d'Ansley (1980) qui est aussi programmé dans le langage S. Les autres observations sont obtenues en résolvant l'équation aux différences définissant ces trois modèles. Afin de réduire le temps de calcul, quelques fonctions de récurrence ont été programmées dans le langage Fortran77. Par la suite, pour chaque sous-série bivariée  $\{\mathbf{X}_t^{(h)}\}$ ,  $h = 1, 2$ , nous avons estimé par la méthode des moindres carrés un modèle autorégressif. L'ordre  $p$  a été déterminé en minimisant le critère AIC d'Akaike pour  $1 \leq p \leq P$ , où  $P$  est fixé. À partir des séries résiduelles  $\hat{\mathbf{a}}_t^{(h)}$ ,  $h = 1, 2$ , nous avons calculé les matrices de corrélations croisées résiduelles,  $\mathbf{R}_{\hat{\mathbf{a}}}^{(12)}(j)$  définies à la section 1.1.

### 5.1.3. Procédure des tests

Pour les tests basés sur les statistiques  $\mathcal{Q}_N$  et  $\mathcal{Q}_N^*$  définies par

$$\mathcal{Q}_N = \frac{N \sum_{j=1-N}^{N-1} k^2(j/M) Q_{\hat{\mathbf{a}}}(j) - m_1 m_2 S_N(k)}{\sqrt{2m_1 m_2 D_N(k)}}$$

$$\mathcal{Q}_N^* = \frac{N \sum_{j=1-N}^{N-1} k^2(j/M) Q_{\hat{\mathbf{a}}}(j) - m_1 m_2 MS(k)}{\sqrt{2M m_1 m_2 D(k)}},$$

on calcule pour chaque noyau les quantités  $S_N(k)$ ,  $D_N(k)$ ,  $S(k)$  et  $D(k)$  définies par (3.1.8), (3.1.9), (3.2.5), à l'aide d'une fonction programmée dans le langage S. Les différents noyaux qu'on va utiliser pour pondérer les statistiques du test et qui appartiennent à la classe  $\mathcal{K}_1$  sont le noyau Uniforme Tronqué (TR), le noyau de Bartlett (BAR), le noyau de Daniell (DAN), le noyau de Parzen (PAR), le noyau de Bartlett-Priestley (BP) et le noyau de Tukey-Hamming (TH) qui sont

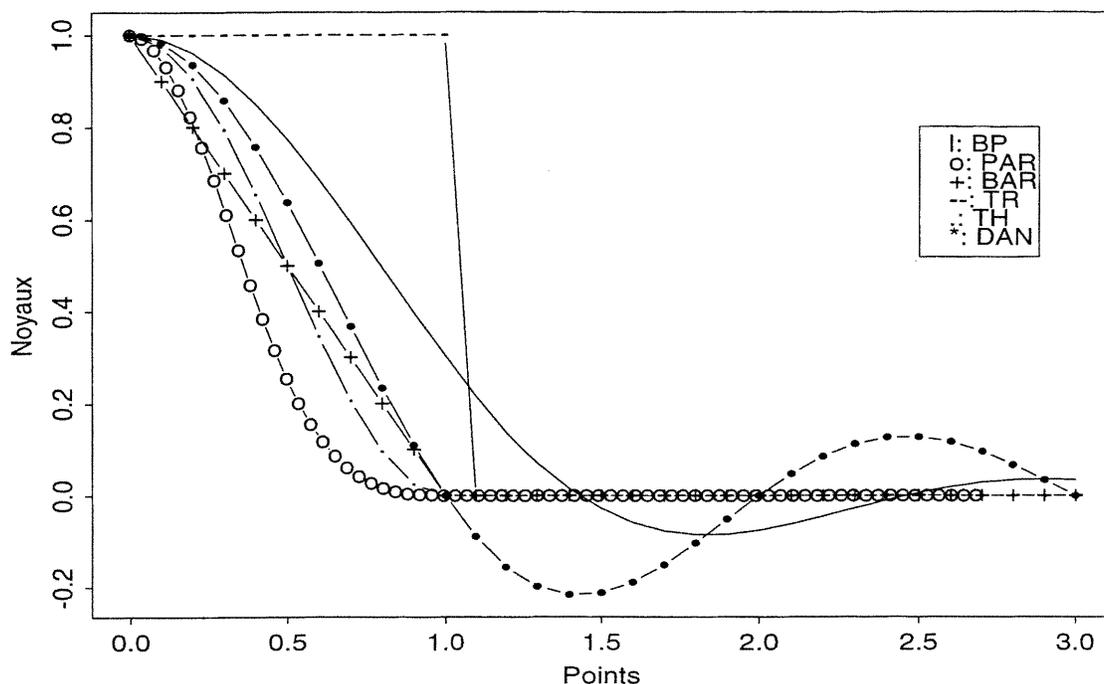


FIG. 5.1. *Comparaison des différents noyaux utilisés*

définis par (3.1.2) à (3.1.7). Les noyaux TR, PAR et TH ont un support compact puisque  $k(z) = 0$  pour  $|z| > 1$ . Pour ces fonctions,  $M$  est le nombre qui représente le délai de troncature, puisque les délais d'ordre  $j > M$  ont un poids nul. Par contre, les noyaux de Daniell et de Bartlett-Priestley ont un support non borné. Pour ces noyaux, toutes les  $(N - 1)$  autocorrélations échantionnales sont utilisées, et  $M$  peut être considéré comme un paramètre de lissage.

La figure 5.1 représente le graphique des six noyaux utilisés pour comparer les tests. Ces noyaux ont été normalisés afin que les variances asymptotiques des statistiques de tests soient identiques. Seuls les biais asymptotiques diffèrent.

Cette renormalisation paraît nécessaire afin que la comparaison soit adéquate pour chaque noyau et pour chaque valeur du paramètre  $M(N)$ , voir Andrews (1991). Et puisque la variance asymptotique dépend de  $D(k)$ , alors les noyaux renormalisés  $\tilde{k}(x)$  vérifient  $\int_{-\infty}^{\infty} \tilde{k}(x)^4 dx = 1$ . Par la suite, les noyaux utilisés dans les statistiques de tests sont renormalisés par  $\tilde{k}(x) = k(cx)$  avec  $c = \int k(x)^4 dx$ , où  $k$  représente un des noyaux cités ci dessus. Ainsi,  $c(TR) = 2$ ,  $c(BAR) = 2/5$ ,  $c(PAR) = 0.382615$ ,  $c(DAN) = 2/3$ ,  $c(TH) = 35/64$ ,  $c(BP) = 0.867532$ . Pour le paramètre de lissage  $M$  trois choix vont être utilisés,  $M = [\ln(N)]$ ,  $[3N^{0.2}]$  et  $[3N^{0.3}]$ , où  $[x]$  représente la partie entière de  $x$ . Ces choix sont basés sur l'optimisation du choix du paramètre de lissage, voir Andrews (1991), Robinson (1991).

Pour chaque modèle, la simulation sera faite selon les étapes suivantes:

- (1) On génère indépendamment  $N$  vecteurs  $\mathbf{a}_t$  de dimension 4 d'une loi  $N(\mathbf{0}, \Sigma)$  en utilisant la fonction *rnorm* de Splus.
- (2) Pour chacun des modèles AR(1) et MA(1) dont les valeurs des paramètres sont données dans le tableau 5.2, les valeurs de  $\mathbf{X}_1, \dots, \mathbf{X}_N$  sont obtenues en résolvant les équations aux différences (1.2.1) et (1.2.3). 5000 et 2000 réalisations indépendantes ont été générées pour les longueurs de série  $N = 100$  et 200 respectivement.
- (3) Les deux séries  $\{\mathbf{X}_t^{(h)}\}_{t=1}^N$ ,  $h = 1, 2$ , sont modélisées de façon indépendante en estimant des modèles autorégressifs par la méthode des moindres carrés conditionnels. L'ordre autorégressif a été déterminé en minimisant le critère AIC d'Akaike pour  $p \leq P$  où  $P$  est fixé à 12. Les séries résiduelles  $\{\hat{\mathbf{a}}_t^{(h)}\}$ ,  $h = 1, 2$ , sont utilisées dans le calcul des corrélations croisées définies par (1.1.14).
- (4) Les valeurs des statistiques  $Q_{\hat{\mathbf{a}}, M}^*$  seront calculées pour chaque  $M = 1, \dots, 12$ . Pour chaque test, la valeur de la statistique sera comparée à la valeur critique

$N = 5000$	NR5	NR1	$N = 2000$	NR5	NR1
$\alpha = 1$	[0.7, 1.3]	[0.6, 1.4]	$\alpha = 1$	[0.6, 1.4]	[0.4, 1.6]
$\alpha = 5$	[4.4, 5.6]	[4.2, 5.8]	$\alpha = 5$	[4.0, 6.0]	[3.7, 6.3]
$\alpha = 10$	[9.2, 10.8]	[8.9, 11.1]	$\alpha = 10$	[8.7, 11.3]	[8.3, 11.7]

TABLE 5.1. Zones de non signification à 5% et 1% du pourcentage de rejets basé sur 5000 et 2000 réalisations respectivement pour chacun des trois niveaux nominaux.

obtenue à partir de la distribution de khi-deux correspondante. Les statistiques  $\mathcal{Q}_N$  et  $\mathcal{Q}_N^*$  sont calculées pour chaque noyau mentionné ci-dessus et sont comparées à la valeur critique de la distribution normale  $\mathcal{N}(0, 1)$  pour les deux premiers modèles. Pour le troisième modèle  $AR(1)_\delta$ , nous allons évaluer la puissance pour différentes valeurs de  $\delta$ .

(5) Finalement, pour chaque niveau de signification nominal,  $\alpha = 0.01, 0.05$  et  $0.10$ , pour les deux premiers modèles et pour la longueur de série  $N = 100$ , nous obtenons à partir des 5000 réalisations les niveaux empiriques (fréquences de rejets de l'hypothèse nulle d'indépendance). Les erreurs types des niveaux empiriques basés sur 5000 réalisations sont 0.14% pour 1%, 0.31% pour 5% et 0.42% pour 10%. Pour la longueur  $N = 200$ , nous avons généré 2000 réalisations de chaque modèle. Cette restriction à 2000 est due à la lenteur du programme en Splus. Les erreurs types sont 0.22% pour 1%, 0.49% pour 5% et 0.67% pour 10%. Ces valeurs permettent d'identifier, pour chaque test, les niveaux empiriques qui diffèrent significativement des niveaux nominaux correspondants.

(6) Finalement, afin d'étudier la puissance, nous avons généré pour les deux longueurs  $N = 100$  et  $200$ , 2000 réalisations du modèle  $AR(1)_\delta$ , avec  $\delta = 1.0, 1.5$  et  $2$ , dont les valeurs des paramètres sont données dans le tableau 5.2.

## 5.2. ANALYSE DES RÉSULTATS DE SIMULATION.

### 5.2.1. Discussion de l'étude de niveau

Les résultats de l'étude de niveau sont rapportés dans les tableaux 5.3 à 5.6. Ces résultats sont analysés selon les zones de non signification à 5% et à 1% du pourcentage de rejets pour chacun de trois niveaux nominaux, qui ont été notées respectivement NR5 et NR1 dans le tableau 5.1.3.

Pour le premier modèle, on note que pour des séries de taille  $N = 100$ , nous obtenons de bons résultats pour la deux tests basés sur les statistiques  $Q_N$  et  $Q_N^*$ , et aussi pour tous les noyaux, sauf le noyau tronqué. Ainsi, on remarque que les niveaux observés sont pour la plupart à l'intérieur des zones de non-rejet NR5 pour les niveaux 5% et 10% et se situent, pour les mêmes niveaux, très près des zones de non rejet NR1. Ces niveaux deviennent de plus en plus précis quand  $M$  croît. On note aussi une sous estimation du niveau de la part du noyau PAR et surtout pour une petite valeur de  $M$ . Aussi, un sur rejet est observé avec les noyaux DAN et BP pour le niveau 10%. Pour le noyau BAR, les niveaux sont généralement bien approximés. Pour la statistique  $Q_{\hat{a},M}^*$ , on note que le niveau est assez bien contrôlé et les résultats sont presque similaires à ceux du noyau TR. Pour  $N = 200$ , et comme on s'y attendait, l'effet de la taille est remarquable. Pour tous les noyaux, il y a eu une amélioration très nette de l'approximation du niveau. Le niveau 10% est surtout très bien contrôlé. Ces approximations deviennent meilleures quand  $M$  prend des valeurs plus grandes. On remarque aussi que la statistique  $Q_{\hat{a},M}^*$  s'améliore quand  $M$  croît.

Finalement, on peut dire qu'avec les deux modèles considérés, le niveau est en général bien contrôlé, spécialement à 5% et cela même avec des séries de longueur  $N = 100$ . L'approximation du niveau s'améliore un peu quand la longueur des

séries croît. Aussi, il en est de même lorsque le paramètre de lissage  $M$  croît.

L'utilisation d'un facteur de correction similaire à celui proposé par Li et McLeod (1981) dans le cas multivarié (voir aussi Kheoh et McLeod (1992) pour le cas univarié) conduirait possiblement à un meilleur contrôle du niveau mais ce point n'est pas discuté ici. Aussi, on peut conclure qu'il n'y a pratiquement pas de différence d'un point de vue pratique entre les deux statistiques  $Q_N$  et  $Q_N^*$ .

### 5.2.2. Discussion de l'étude de puissance

Pour les modèles  $AR(1)_\delta$ , la matrice de covariance croisée n'est pas nulle et est fonction d'un paramètre  $\delta$  qui fait varier la dépendance entre les deux séries. Pour le modèle  $AR(1)_\delta$ , les résultats sont donnés dans les tableaux 5.7 à 5.12 respectivement pour  $\delta = 1.0, 1.5$  et  $2$ . Pour  $\delta = 1$  et  $N = 100$ , on observe un gain de puissance des tests basés sur les noyaux considérés, autre que le noyau tronqué  $TR$ , par rapport au test  $Q_{\hat{a},M}^{(12)*}$ . Au niveau 5%, Le noyau  $BP$  performe bien pour toutes les valeurs de  $M$  considérées. Pour le même niveau, le noyau  $DAN$ , performe aussi bien pour les petites valeurs de  $M$  que pour les grandes. Pour 10%, le comportement des noyaux est similaire, et on remarque une légère performance du noyau  $PAR$  par rapport aux autres. Pour la même valeur de  $\delta$  et  $N = 200$ , on observe une certaine amélioration de la puissance pour toutes les valeurs de  $M$  et pour tous les niveaux. On remarque aussi une meilleure performance du noyau  $DAN$  pour toutes les valeurs de  $M$ . Les résultats pour les autres noyaux autre que  $TR$  sont plutôt similaires. Aussi, quand la valeur de  $\delta$  augmente, la dépendance est plus forte et ainsi le test a plus de facilité à la détecter. On remarque pour les deux tailles échantionnales considérées que lorsque  $M$  prend une grande valeur, la puissance diminue et ceci est normal, puisque les processus considérés

sont à mémoire courte et quand  $M$  croît, on prend plus de délais. Avec le noyau tronqué uniforme  $TR$ , les résultats sont presque identiques à ceux du test  $Q_{\hat{a},M}^{(12)*}$ . Finalement, on peut conclure sur la puissance qu'il n'y a pas de différence importante entre les noyaux excepté le noyau TR qui donne des puissances rapprochées à celles de la statistique  $Q_{\hat{a},M}^{(12)*}$ . De même que l'approximation du niveau, quand la longueur des séries passe de 100 à 200, la puissance augmente considérablement. Par contre, quand  $M$  augmente, la puissance décroît.

TAB. 5.2. Modèles et paramètres des processus utilisés dans l'étude par simulation

MODÈLES	ÉQUATIONS	$\Sigma_a$
AR(1)	$\begin{bmatrix} \mathbf{X}_t^{(1)} \\ \mathbf{X}_t^{(2)} \end{bmatrix} = \begin{bmatrix} \Phi^{(1)} & 0 \\ 0 & \Phi^{(2)} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{t-1}^{(1)} \\ \mathbf{X}_{t-1}^{(2)} \end{bmatrix} + \begin{bmatrix} \mathbf{a}_t^{(1)} \\ \mathbf{a}_t^{(2)} \end{bmatrix}$	$\begin{bmatrix} \Sigma_a^{(1)} & 0 \\ 0 & \Sigma_a^{(2)} \end{bmatrix}$
MA(1)	$\begin{bmatrix} \mathbf{X}_t^{(1)} \\ \mathbf{X}_t^{(2)} \end{bmatrix} = \begin{bmatrix} \Psi^{(1)} & 0 \\ 0 & \Psi^{(2)} \end{bmatrix} \begin{bmatrix} \mathbf{a}_{t-1}^{(1)} \\ \mathbf{a}_{t-1}^{(2)} \end{bmatrix} + \begin{bmatrix} \mathbf{a}_t^{(1)} \\ \mathbf{a}_t^{(2)} \end{bmatrix}$	$\begin{bmatrix} \Sigma_a^{(1)} & 0 \\ 0 & \Sigma_a^{(2)} \end{bmatrix}$
AR(1) $_{\delta}$	$\begin{bmatrix} \mathbf{X}_t^{(1)} \\ \mathbf{X}_t^{(2)} \end{bmatrix} = \begin{bmatrix} \Phi^{(1)} & 0 \\ 0 & \Phi^{(2)} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{t-1}^{(1)} \\ \mathbf{X}_{t-1}^{(2)} \end{bmatrix} + \begin{bmatrix} \mathbf{a}_t^{(1)} \\ \mathbf{a}_t^{(2)} \end{bmatrix}$	$\begin{bmatrix} \Sigma_a^{(1)} & \Sigma_{a,\delta}^{(12)} \\ \Sigma_{a,\delta}^{(12)} & \Sigma_a^{(2)} \end{bmatrix}$
PARAMÈTRES		
$\Phi^{(1)} = \begin{bmatrix} 1.2 & -0.5 \\ 0.6 & 0.3 \end{bmatrix}$	$\Phi^{(2)} = \begin{bmatrix} -0.6 & 0.3 \\ 0.3 & 0.6 \end{bmatrix}$	$\Psi^{(1)} = \begin{bmatrix} -0.2 & 0.3 \\ -0.6 & 1.1 \end{bmatrix}$
$\Psi^{(2)} = \begin{bmatrix} 0.8 & 0.3 \\ 0.1 & 0.6 \end{bmatrix}$	$\Sigma_a^{(1)} = \begin{bmatrix} 1.0 & 0.5 \\ 0.5 & 1.0 \end{bmatrix}$	$\Sigma_a^{(2)} = \begin{bmatrix} 1.0 & 0.75 \\ 0.75 & 1.0 \end{bmatrix}$
$\Sigma_{a,\delta}^{(12)} = \begin{bmatrix} 0.1\delta & 0 \\ 0 & 0.05\delta \end{bmatrix}$		

TAB. 5.3. Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle  $AR(1)$  avec  $Q_N$  et  $Q_N^*$  basés sur 5000 réalisations pour  $N = 100$  et 2000 réalisations pour  $N = 200$  en utilisant les noyaux  $DAN$ ,  $PAR$ ,  $BP$  et  $TR$ .

			$Q_N$				$Q_N^*$			
			$DAN$	$PAR$	$BP$	$TR$	$DAN$	$PAR$	$BP$	$TR$
$N = 100$	$M = 5$	1%	0.8	0.5	0.7	0.6	0.9	0.6	0.8	0.7
		5%	6.1	3.4	4.8	4.6	4.8	4.2	5.1	4.7
		10%	9.4	7.8	10.2	8.4	9.7	8.2	10.3	9.3
	$M = 8$	1%	0.9	0.6	1.3	0.6	1.5	0.8	1.2	0.9
		5%	5.4	3.9	5.2	3.6	5.2	4.2	4.8	4.1
		10%	10.5	8.9	10.7	7.1	9.4	9.3	10.2	8.9
	$M = 12$	1%	0.7	0.6	0.9	0.5	0.8	0.7	1.1	0.6
		5%	5.1	4.6	4.8	4.1	4.2	4.3	4.7	4.4
		10%	10.6	8.6	11.1	7.5	10.3	9.2	10.4	7.9
$N = 200$	$M = 5$	1%	0.8	0.6	0.8	0.7	0.8	0.8	1.1	0.8
		5%	5.9	4.8	5.7	4.2	6.1	4.5	5.2	4.8
		10%	9.1	8.3	9.4	7.9	8.5	8.4	9.6	8.6
	$M = 9$	1%	0.8	0.6	0.8	0.7	0.8	0.7	1.2	0.8
		5%	6.3	4.2	5.7	4.2	5.8	4.4	5.6	4.7
		10%	9.7	8.9	10.2	7.3	8.9	9.0	9.7	8.9
	$M = 15$	1%	1.2	0.7	0.8	0.7	0.9	0.8	0.8	0.7
		5%	6.4	4.1	5.8	3.9	5.1	4.3	5.2	4.3
		10%	10.2	8.9	10.4	6.9	9.7	9.1	10.6	8.5

TAB. 5.4. Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle  $AR(1)$  avec  $Q_N, Q_N^*$  et  $Q_{\hat{a},M}^*$  basés sur 5000 réalisations pour  $N = 100$  et 2000 réalisations pour  $N = 200$  en utilisant les noyaux TH et BAR.

			$Q_N$		$Q_N^*$		$Q_{\hat{a},M}^*$
			TH	BAR	TH	BAR	
$N = 100$	$M = 5$	1%	0.7	0.7	0.6	0.8	0.7
		5%	5.8	5.5	5.1	4.8	4.4
		10%	10.2	9.7	9.9	8.5	9.1
	$M = 8$	1%	0.8	0.7	1.1	0.7	0.7
		5%	5.4	5.7	4.9	4.8	4.7
		10%	9.4	9.2	9.6	9.8	8.9
	$M = 12$	1%	0.8	0.8	1.1	0.9	0.7
		5%	5.2	4.8	4.7	4.7	4.2
		10%	9.2	8.7	9.5	9.2	8.2
$N = 200$	$M = 5$	1%	0.6	0.7	0.7	0.9	0.8
		5%	5.4	5.6	4.6	4.8	4.1
		10%	9.4	9.6	9.6	9.6	8.7
	$M = 9$	1%	0.8	0.7	0.9	0.7	0.7
		5%	5.8	5.7	5.6	5.7	4.4
		10%	10.9	9.7	10.0	9.6	9.0
	$M = 15$	1%	1.0	0.8	0.7	0.6	0.8
		5%	4.8	4.5	5.4	4.6	4.6
		10%	10.6	9.7	10.2	9.8	8.9

TAB. 5.5. Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle  $MA(1)$  avec  $Q_N$  et  $Q_N^*$  basés sur 5000 réalisations pour  $N = 100$  et 2000 réalisations pour  $N = 200$  en utilisant les noyaux  $DAN$ ,  $PAR$ ,  $BP$  et  $TR$ .

			$Q_N$				$Q_N^*$			
			$DAN$	$PAR$	$BP$	$TR$	$DAN$	$PAR$	$BP$	$TR$
$N = 100$	$M = 5$	1%	0.8	0.6	0.9	0.6	0.9	0.7	1.1	0.7
		5%	5.9	4.2	5.2	4.6	4.8	4.3	5.1	4.7
		10%	10.3	7.8	10.2	8.2	9.6	9.5	9.8	8.7
	$M = 8$	1%	0.8	0.5	1.3	0.6	1.1	0.6	1.1	0.7
		5%	5.4	4.2	5.6	3.6	5.2	4.3	4.8	4.1
		10%	10.3	8.2	10.1	7.1	9.4	8.6	9.2	7.9
	$M = 12$	1%	0.8	0.6	0.9	0.7	0.8	0.7	1.1	0.8
		5%	5.4	4.7	5.8	4.6	5.2	4.9	5.2	4.8
		10%	9.3	8.6	9.2	7.5	9.5	8.4	9.4	7.9
$N = 200$	$M = 5$	1%	0.7	0.6	0.8	0.7	0.8	0.7	1.1	0.8
		5%	5.9	4.5	5.7	4.2	6.1	4.4	5.2	4.8
		10%	9.2	8.3	9.0	7.9	9.5	8.4	9.6	8.6
	$M = 9$	1%	0.8	0.9	0.8	0.7	0.8	0.9	1.0	0.8
		5%	6.3	4.1	5.7	4.2	5.8	4.5	5.6	4.7
		10%	9.7	8.9	10.5	7.3	9.2	9.1	9.7	8.9
	$M = 15$	1%	1.1	0.8	0.9	0.7	0.8	0.8	1.1	0.9
		5%	6.4	4.4	5.8	4.5	5.1	4.5	5.4	4.6
		10%	10.2	9.1	9.2	6.9	9.7	9.5	10.1	9.5

TAB. 5.6. Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle  $MA(1)$  avec  $Q_N, Q_N^*$  et  $Q_{a,M}^*$  basés sur 5000 réalisations pour  $N = 100$  et 2000 réalisations pour  $N = 200$  en utilisant les noyaux  $TH$  et  $BAR$ .

			$Q_N$		$Q_N^*$		$Q_{a,M}^*$
			$TH$	$BAR$	$TH$	$BAR$	
$N = 100$	$M = 5$	1%	0.8	0.7	1.1	0.6	0.8
		5%	6.1	5.4	5.2	4.6	4.8
		10%	10.3	9.8	10.2	8.2	8.7
	$M = 8$	1%	0.9	1.1	1.0	0.8	1.1
		5%	5.4	5.9	5.2	5.2	5.2
		10%	10.5	10.6	10.1	7.1	9.4
	$M = 12$	1%	0.8	0.9	0.8	0.6	0.8
		5%	5.1	4.9	4.8	4.6	4.3
		10%	9.3	8.6	9.1	8.9	7.9
$N = 200$	$M = 5$	1%	0.7	0.8	0.9	0.7	0.7
		5%	5.9	5.7	5.6	5.2	6.1
		10%	9.4	9.6	9.2	8.9	8.5
	$M = 9$	1%	0.9	0.8	0.8	0.7	0.8
		5%	6.3	6.1	5.8	5.2	5.8
		10%	9.7	9.5	10.2	9.3	8.9
	$M = 15$	1%	1.1	0.9	0.7	0.8	0.8
		5%	5.7	6.0	5.6	5.7	5.1
		10%	10.2	9.8	9.3	8.9	8.7

TAB. 5.7. Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle  $AR(1)_\delta$ ,  $\delta = 1$  avec  $Q_N$  et  $Q_N^*$  basés sur 2000 réalisations en utilisant les noyaux  $DAN$ ,  $PAR$ ,  $BP$  et  $TR$ .

			$Q_N$				$Q_N^*$			
			$DAN$	$PAR$	$BP$	$TR$	$DAN$	$PAR$	$BP$	$TR$
$N = 100$	$M = 5$	1%	27.2	26.4	24.6	16.2	28.2	26.0	25.6	16.8
		5%	32.6	34.6	33.0	22.0	32.8	32.4	34.2	20.6
		10%	42.4	44.6	44.2	26.2	43.6	40.6	42.8	24.8
	$M = 8$	1%	22.4	20.8	23.0	12.4	22.6	24.2	23.8	11.8
		5%	30.2	27.6	28.4	14.1	28.8	31.0	28.6	13.8
		10%	34.2	36.0	32.8	16.0	32.6	34.8	32.0	15.4
	$M = 12$	1%	18.0	16.4	16.8	7.6	20.2	18.8	18.4	8.4
		5%	25.2	24.4	26.2	11.8	24.2	26.2	26.4	12.4
		10%	32.2	32.6	30.8	16.0	34.2	30.2	32.8	16.8
$N = 200$	$M = 5$	1%	46.2	44.8	42.4	26.2	44.6	46.4	44.2	27.0
		5%	52.4	54.4	56.0	28.6	58.2	54.2	56.2	30.8
		10%	64.2	60.2	61.8	34.2	62.6	64.4	66.2	36.8
	$M = 9$	1%	34.0	32.6	30.8	22.6	32.6	34.8	36.0	24.0
		5%	38.2	36.6	36.2	26.2	36.2	36.4	34.6	28.2
		10%	41.2	42.0	38.0	28.2	40.0	38.6	40.2	30.4
	$M = 15$	1%	27.4	26.0	28.2	17.2	27.2	24.6	26.2	18.0
		5%	32.4	32.0	34.6	20.2	32.0	34.6	34.2	22.6
		10%	34.2	32.8	30.2	24.2	36.2	36.4	34.4	26.0

TABLE 5.8. Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle  $AR(1)_\delta$ ,  $\delta = 1$  avec  $Q_N$ ,  $Q_N^*$  et  $Q_{\hat{\alpha},M}^*$  basés sur 2000 réalisations en utilisant les noyaux TH et BAR.

			$Q_N$		$Q_N^*$		$Q_{\hat{\alpha},M}^*$
			TH	BAR	TH	BAR	
$N = 100$	$M = 5$	1%	26.8	27.4	24.5	26.6	18.8
		5%	32.4	36.2	34.2	34.6	22.8
		10%	44.3	42.1	42.2	40.2	24.7
	$M = 8$	1%	20.8	18.6	19.2	20.4	11.4
		5%	28.4	26.9	27.2	24.6	13.2
		10%	33.4	34.6	36.1	36.4	15.4
	$M = 12$	1%	18.2	15.9	16.8	17.4	8.2
		5%	25.1	23.7	24.8	25.1	12.6
		10%	31.3	30.6	28.1	30.5	15.8
$N = 200$	$M = 5$	1%	44.6	46.2	44.8	48.4	28.6
		5%	52.8	54.6	52.7	56.2	31.2
		10%	62.2	60.3	58.2	57.6	32.2
	$M = 9$	1%	32.4	30.8	30.6	31.2	22.4
		5%	36.2	35.1	38.2	36.2	25.8
		10%	40.7	38.8	40.0	36.6	28.4
	$M = 15$	1%	26.2	28.3	24.6	25.6	18.2
		5%	32.2	30.2	28.8	30.8	22.2
		10%	32.6	34.2	32.2	36.2	25.2

TAB. 5.9. Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle  $AR(1)_\delta$ ,  $\delta = 1.5$  avec  $Q_N$  et  $Q_N^*$  basés sur 2000 réalisations en utilisant les noyaux  $DAN$ ,  $PAR$ ,  $BP$  et  $TR$ .

			$Q_N$				$Q_N^*$			
			$DAN$	$PAR$	$BP$	$TR$	$DAN$	$PAR$	$BP$	$TR$
$N = 100$	$M = 5$	1%	32.8	27.4	29.1	22.4	32.8	27.7	30.1	21.2
		5%	42.7	40.3	39.8	25.8	40.1	38.9	41.7	28.7
		10%	50.7	47.9	48.1	30.2	47.4	48.5	50.3	31.7
	$M = 8$	1%	32.9	31.7	33.4	30.5	22.7	31.9	32.0	17.8
		5%	28.4	30.9	31.2	22.6	32.2	30.0	32.8	20.1
		10%	28.3	24.6	25.1	19.1	27.9	24.6	25.5	19.9
	$M = 12$	1%	27.4	26.9	27.8	16.4	29.8	26.8	27.9	18.6
		5%	29.1	28.7	28.8	19.4	30.2	32.3	29.0	17.8
		10%	33.4	32.5	29.6	21.5	32.9	30.6	29.5	21.6
$N = 200$	$M = 5$	1%	54.4	54.6	52.6	32.2	52.8	50.6	52.0	33.4
		5%	57.9	55.7	59.7	31.2	60.1	58.9	59.8	36.8
		10%	61.2	58.2	64.2	38.9	58.5	60.4	60.6	40.6
	$M = 9$	1%	45.8	44.9	47.6	28.7	44.4	42.9	46.9	29.8
		5%	52.3	50.2	52.7	33.2	54.9	53.4	51.6	32.5
		10%	57.8	52.9	55.0	35.3	58.9	53.4	56.8	33.4
	$M = 15$	1%	41.0	38.4	43.7	22.4	42.8	45.4	45.1	25.9
		5%	44.4	43.5	45.7	21.7	45.1	46.4	45.1	24.3
		10%	56.8	54.8	57.2	30.9	58.6	56.6	59.0	29.5

TABLE 5.10. Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle  $AR(1)_\delta$ ,  $\delta = 1.5$  avec  $Q_N$ ,  $Q_N^*$  et  $Q_{a,M}^*$  basés sur 2000 réalisations en utilisant les noyaux TH et BAR.

			$Q_N$		$Q_N^*$		$Q_{a,M}^*$
			TH	BAR	TH	BAR	
N = 100	M = 5	1%	32.4	30.7	33.4	31.8	23.9
		5%	42.4	40.6	43.1	42.6	24.8
		10%	46.3	47.9	45.1	46.9	29.6
	M = 8	1%	25.4	23.6	24.0	24.6	18.4
		5%	30.4	31.9	32.4	32.6	20.5
		10%	36.5	35.4	37.9	34.7	23.4
	M = 12	1%	25.3	22.9	26.8	23.5	15.8
		5%	27.8	24.9	27.8	25.1	17.5
		10%	31.1	28.5	29.9	27.8	22.9
N = 200	M = 5	1%	50.5	48.9	48.8	49.7	34.8
		5%	55.9	53.7	56.6	54.2	38.1
		10%	61.8	60.2	63.0	61.6	40.9
	M = 9	1%	47.2	48.9	48.8	50.7	28.7
		5%	52.3	50.1	53.7	51.4	30.8
		10%	57.7	55.9	57.9	53.5	32.9
	M = 15	1%	40.0	39.4	42.7	41.7	21.8
		5%	45.6	47.0	46.7	49.6	24.5
		10%	56.3	54.7	57.5	54.8	28.6

TAB. 5.11. Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle  $AR(1)_\delta$ ,  $\delta = 2$  avec  $Q_N$  et  $Q_N^*$  basés sur 2000 réalisations en utilisant les noyaux  $DAN$ ,  $PAR$ ,  $BP$  et  $TR$ .

			$Q_N$				$Q_N^*$			
			$DAN$	$PAR$	$BP$	$TR$	$DAN$	$PAR$	$BP$	$TR$
$N = 100$	$M = 5$	1%	53.8	51.4	56.3	32.4	54.2	52.7	53.1	31.2
		5%	58.3	60.2	59.4	33.8	59.1	61.9	58.7	34.2
		10%	65.1	62.4	68.1	37.2	64.8	66.2	67.3	36.7
	$M = 8$	1%	49.4	51.2	50.8	26.5	51.2	48.6.9	49.0	25.7
		5%	54.4	52.6	56.2	29.2	56.2	54.0	55.0	30.1
		10%	60.3	57.8	62.1	32.1	62.3	60.6	63.5	31.9
	$M = 12$	1%	45.4	44.2	46.8	24.4	46.2	44.7	45.3	23.4
		5%	52.1	49.7	50.8	27.4	51.2	48.3	52.0	26.4
		10%	53.4	52.6	54.2	27.5	54.9	52.6	53.5	26.8
$N = 200$	$M = 5$	1%	72.4	74.5	74.6	48.2	76.8	72.6	74.0	50.8
		5%	82.4	80.4	80.6	54.2	84.2	82.6	84.8	52.8
		10%	92.2	89.6	90.8	58.4	92.5	90.4	90.6	60.6
	$M = 9$	1%	64.2	62.8	60.6	42.4	66.4	62.9	63.8	40.2
		5%	70.2	68.4	72.8	45.4	71.2	72.4	74.6	42.5
		10%	76.8	78.9	74.8	48.2	77.8	75.4	78.8	47.4
	$M = 15$	1%	52.0	52.4	54.7	30.2	54.8	53.4	55.1	29.9
		5%	61.4	57.5	62.7	30.7	62.1	56.4	61.1	31.3
		10%	60.8	51.8	58.2	30.9	62.6	54.6	60.6	29.5

TAB. 5.12. Pourcentages de rejet de l'hypothèse nulle d'indépendance sous le modèle  $AR(1)_\delta$ ,  $\delta = 2$  avec  $Q_N$ ,  $Q_N^*$  et  $Q_{\hat{a},M}^*$  basés sur 2000 réalisations en utilisant les noyaux TH et BAR.

			$Q_N$		$Q_N^*$		$Q_{\hat{a},M}^*$
			TH	BAR	TH	BAR	
$N = 100$	$M = 5$	1%	54.2	52.8	55.1	51.8	27.8
		5%	57.6	58.4	55.2	52.5	28.4
		10%	66.5	65.2	69.1	61.4	29.3
	$M = 8$	1%	46.2	42.7	49.4	41.5	19.4
		5%	44.4	47.6	48.4	44.2	20.2
		10%	47.2	46.3	51.9	47.3	22.6
	$M = 12$	1%	25.3	22.9	26.8	23.5	19.8
		5%	27.8	24.9	27.8	25.1	20.5
		10%	31.1	28.5	29.9	27.8	22.9
$N = 200$	$M = 5$	1%	77.5	79.4	81.8	78.7	54.8
		5%	81.9	84.4	80.6	83.5	57.1
		10%	88.4	89.2	90.2	86.7	58.9
	$M = 9$	1%	67.4	70.9	72.3	67.8	45.7
		5%	73.6	72.8	73.2	69.7	47.9
		10%	74.8	71.9	72.5	73.5	48.6
	$M = 15$	1%	57.0	55.9	56.4	54.8	39.8
		5%	59.6	60.1	59.7	58.6	42.5
		10%	62.3	63.8	65.5	62.8	44.6

## Chapitre 6

---

### APPLICATION

En vue d'illustrer les résultats obtenus dans les chapitres précédents et pour comparer nos résultats à ceux d'El Himdi et Roy (1997), nous considérons le même jeu de données. Ainsi, nous considérons un ensemble de sept séries de données trimestrielles d'indicateurs économiques canadiens et américains utilisés initialement lors d'une étude de la politique monétaire canadienne, voir Racette et Raynauld (1992). Les indicateurs économiques canadiens sont la production intérieure brute en dollars constants de 1982 (PIB), l'indice implicite de prix de la production intérieure brute (PIBD), le taux d'intérêt nominale à court terme (TX.CA) et la base monétaire (M1). Les trois autres variables représentent le produit national brut américain en dollars constants de 1982 (GNP), l'indice implicite des prix du produit national brut américain (GNPD) et le taux d'intérêt américain de court terme (TX.US). Dans cette étude, la période d'observation s'étend du premier trimestre de 1970 au dernier trimestre de 1989, soit  $N = 80$  observations pour chaque série.

Nous nous proposons dans cette étude d'appliquer différents tests étudiés précédemment en vue d'explorer la liaison entre les indicateurs économiques canadiens et américains. Dans une analyse préliminaire de chacune des sept séries économiques considérées, nous avons effectué différentes transformations et/ou

différenciation qui visent à rendre ces séries stationnaires et à stabiliser leurs variances. Les séries ainsi obtenues définissent deux séries chronologiques vectorielles canadienne et américaine, de dimension 4 et 3 respectivement. Dans une deuxième étape, nous procédons à la construction d'un modèle autorégressif vectoriel pour chacune de ces deux séries. Les séries résiduelles issues de l'estimation de ces deux modèles permettent de calculer la statistique  $Q_N$ , pour différents noyaux ainsi que les statistiques  $Q_{\hat{a}}^*(j)$  et  $Q_{\hat{a},M}^*$ .

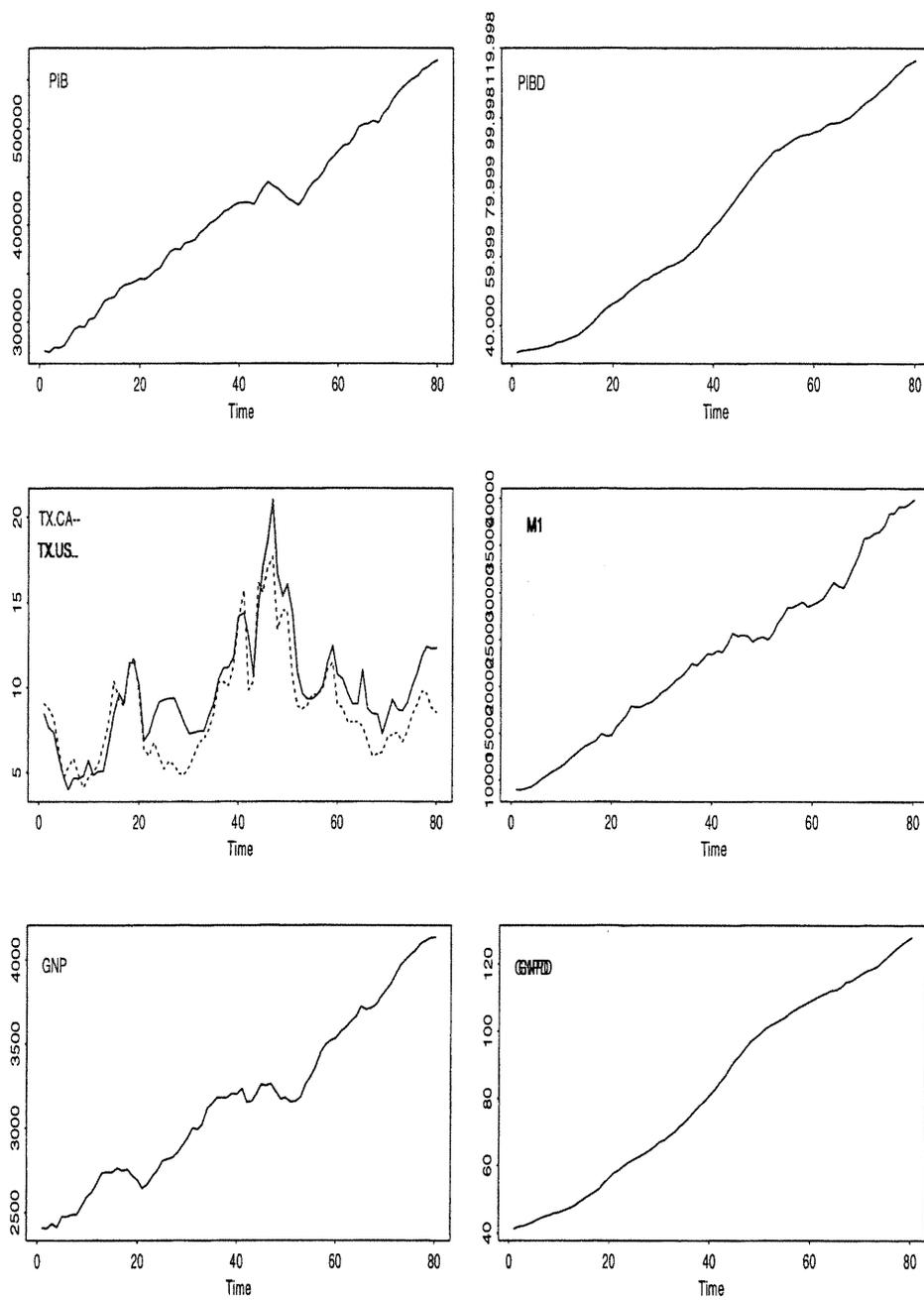


FIG. 6.1. *Séries d'indicateurs économiques canadiens et américains.*

## 6.1. ANALYSE ET MODÉLISATION DES SÉRIES

Les graphiques des sept séries sont contenues dans la figure 6.1. Un examen rapide de ces graphiques montre d'une part que les deux séries de taux d'intérêt se comportent de manière irrégulière mais assez similaire. Les cinq autres variables ont par contre une allure plus lisse et proche de celle d'une promenade aléatoire. Afin de stabiliser la variance et de rendre ces séries stationnaires, nous effectuons des transformations et des différenciations. Nous utilisons les mêmes transformations qu'El Himdi et Roy (1997). Les séries stationnaires canadiennes et américaine sont représentées par  $\{\mathbf{W}_t^{(1)}\}$  et  $\{\mathbf{W}_t^{(2)}\}$  respectivement, où

$$\mathbf{W}_t^{(1)} = \begin{bmatrix} W_{1t}^{(1)} \\ W_{2t}^{(1)} \\ W_{3t}^{(1)} \\ W_{4t}^{(1)} \end{bmatrix} \quad \text{avec} \quad \begin{cases} W_{1t}^{(1)} = \frac{1}{1000}(1 - B)PIB_t, \\ W_{2t}^{(1)} = 10(1 - B)PIBD_t, \\ W_{3t}^{(1)} = TX.CA_t, \\ W_{4t}^{(1)} = 100(1 - B)\ln(M1_t), \end{cases}$$

et

$$\mathbf{W}_t^{(2)} = \begin{bmatrix} W_{1t}^{(2)} \\ W_{2t}^{(2)} \\ W_{3t}^{(2)} \end{bmatrix} \quad \text{avec} \quad \begin{cases} W_{1t}^{(2)} = \frac{1}{10}(1 - B)GNP_t, \\ W_{2t}^{(2)} = 10(1 - B)GNPD_t, \\ W_{3t}^{(2)} = TX.US_t. \end{cases}$$

Les facteurs multiplicatifs ont été choisis pour que les variances des composantes de chaque série multivariée soient de même ordre de grandeur.

L'identification des modèles pour les séries  $\{\mathbf{W}_t^{(1)}\}$  et  $\{\mathbf{W}_t^{(2)}\}$  a été faite à l'aide de la procédure STEPAR du progiciel SCA. Pour chaque série, nous avons trouvé dans un premier temps un modèle  $AR(p)$ . L'ordre  $p$  étant obtenu en minimisant le critère AIC d'Akaike et en utilisant le test khi-deux pour vérifier l'égalité

à zéro du dernier coefficient autorégressif. Pour la série canadienne  $\{\mathbf{W}_t^{(1)}\}$ , nous avons d'abord trouvé  $p = 11$  et par la suite, nous avons retranché les coefficients non significatifs au niveau 5% pour aboutir à un modèle comportant uniquement les valeurs retardées à  $t - 1$ ,  $t - 2$ ,  $t - 11$ , tel que défini par l'équation (6.1.1) suivante :

$$\mathbf{W}_t^{(1)} = \Theta_0^{(1)} + \Phi_1^{(1)}\mathbf{W}_{t-1}^{(1)} + \Phi_2^{(1)}\mathbf{W}_{t-2}^{(1)} + \Phi_{11}^{(1)}\mathbf{W}_{t-11}^{(1)} + \mathbf{a}_t^{(1)}, \quad t \in \mathbb{Z}. \quad (6.1.1)$$

Avec la série  $\{\mathbf{W}_t^{(2)}\}$ , nous avons trouvé  $p = 3$  et toutes les valeurs retardées ont été retenues; le modèle est décrit par l'équation (6.1.2) suivante :

$$\mathbf{W}_t^{(2)} = \Theta_0^{(2)} + \Phi_1^{(2)}\mathbf{W}_{t-1}^{(2)} + \Phi_2^{(2)}\mathbf{W}_{t-2}^{(2)} + \Phi_3^{(2)}\mathbf{W}_{t-3}^{(2)} + \mathbf{a}_t^{(2)}, \quad t \in \mathbb{Z}. \quad (6.1.2)$$

Par la suite, nous avons estimé les paramètres, par la méthode du maximum de vraisemblance, à l'aide de la procédure MESTIM. Nous n'avons retenu que les paramètres estimés dont la valeur absolue est supérieure à une fois son écart-type. Les autres valeurs sont fixés à zéro tout au long de la phase d'estimation. Nous nous sommes assurés que les modèles obtenus sont stationnaires en calculant les racines des polynômes autorégressifs définissant ces deux modèles. Les paramètres de la série canadienne et sa matrice de covariance résiduelles sont données par :

$$\hat{\Theta}_0^{(1)} = \begin{bmatrix} 0.256 \\ 1.275 \\ -0.357 \\ 3.339 \end{bmatrix} \quad \hat{\Phi}_1^{(1)} = \begin{bmatrix} 0.000 & -0.107 & 0.000 & 0.000 \\ 0.000 & 0.576 & 0.000 & 0.000 \\ 0.179 & 0.177 & 0.888 & 0.106 \\ 0.000 & 0.445 & -1.718 & -0.253 \end{bmatrix},$$

$$\hat{\Phi}_2^{(1)} = \begin{bmatrix} 0.000 & 0.000 & 0.000 & 0.516 \\ 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.039 & 0.000 & 0.000 \\ -1.129 & 0.000 & 0.778 & 0.000 \end{bmatrix},$$

$$\hat{\Phi}_{11}^{(1)} = \begin{bmatrix} -0.109 & 0.000 & 0.177 & 0.049 \\ 0.000 & -0.189 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 \end{bmatrix},$$

$$\hat{\Sigma}_a^{(1)} = \begin{bmatrix} 0.746 & -0.327 & 0.033 & -0.007 \\ -0.327 & 0.507 & 0.021 & 0.097 \\ 0.033 & 0.021 & 0.108 & -0.006 \\ -0.007 & 0.097 & -0.006 & 0.447 \end{bmatrix}.$$

Les graphiques des composantes de la série résiduelle canadienne sont présentées dans les figures 6.2. Chaque composante semble se comporter comme un bruit blanc de moyenne zéro et de variance constante. Les corrélations résiduelles sont fournies à la figure 6.3 et nous remarquons qu'il y a seulement sept corrélations légèrement à l'extérieur des limites de signification à 5%. Ceci concorde parfaitement avec les résultats obtenus en utilisant l'analyse des résidus effectuée dans la dernière phase de la modélisation par le progiciel SCA qui nous a permis de conclure que les deux modèles multivariés estimés sont adéquats.

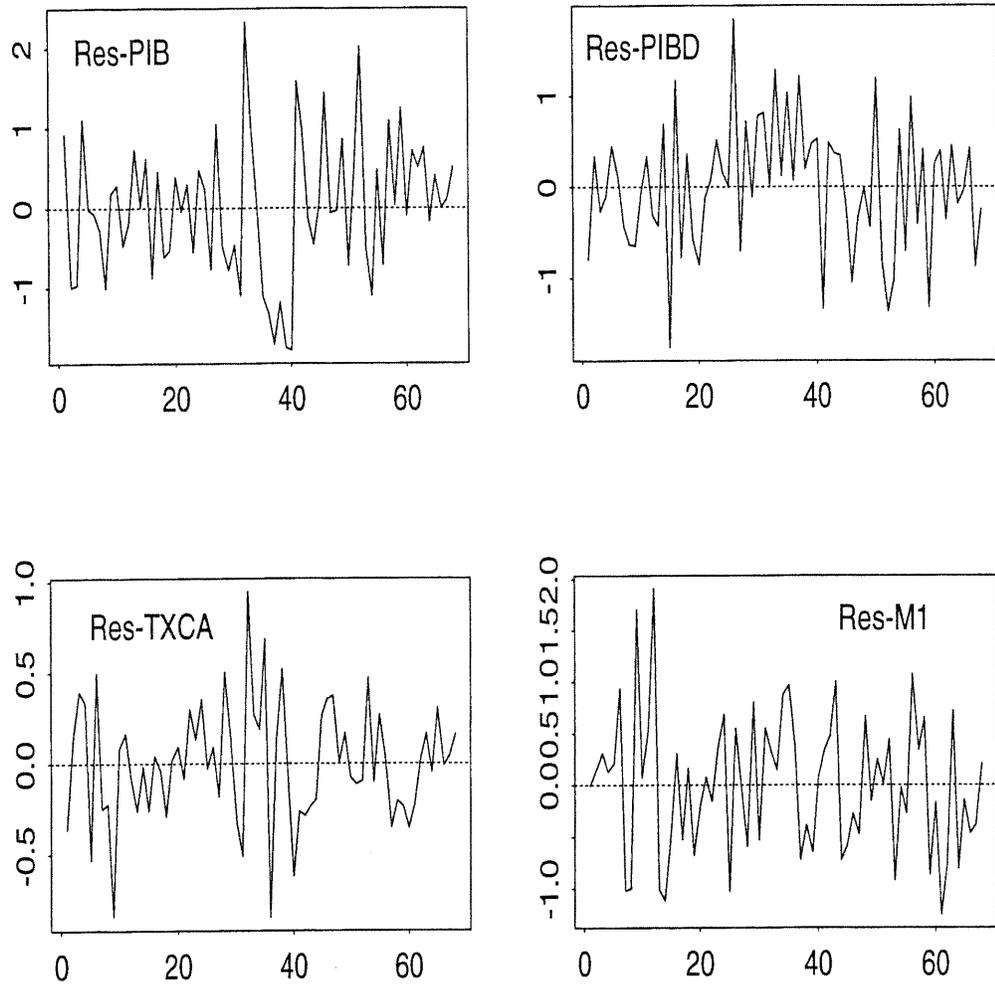


FIG. 6.2. Composantes de la série résiduelle correspondant à la série canadienne différenciée.

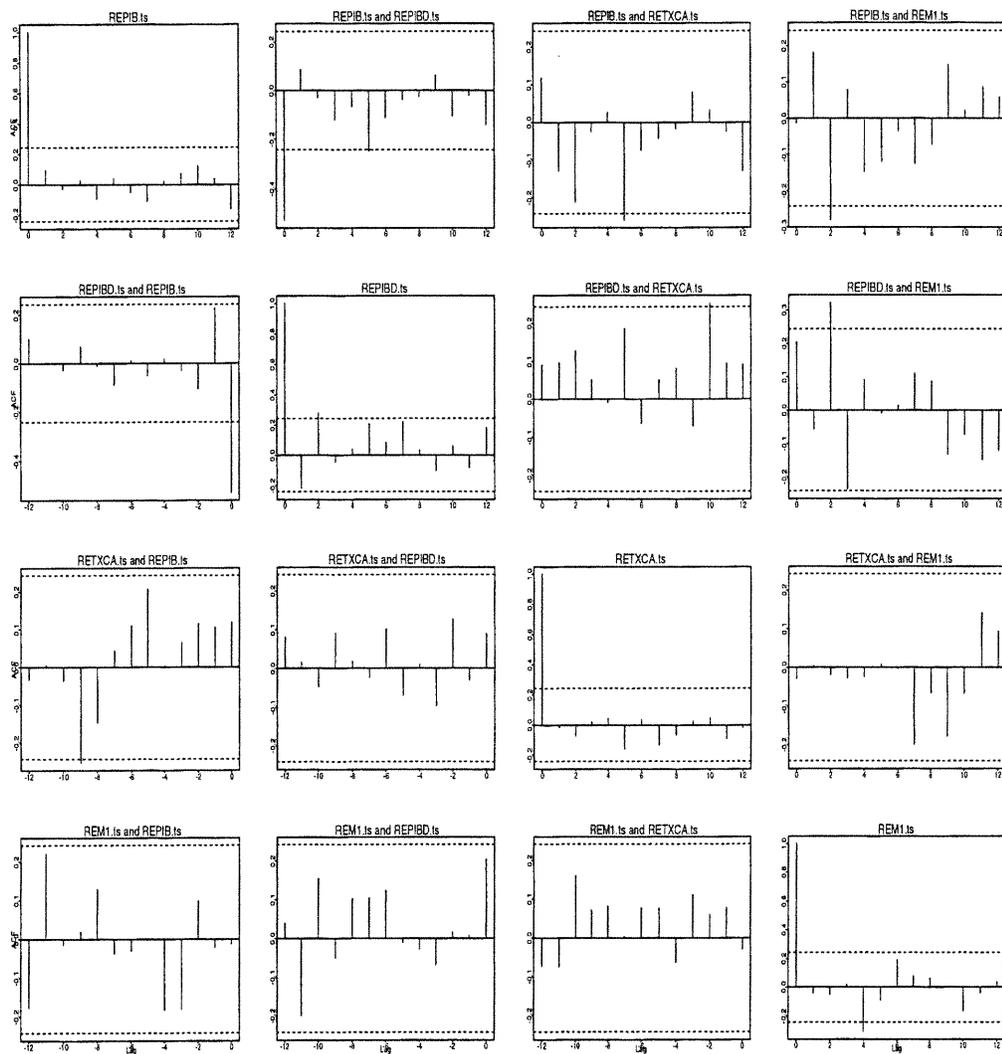


FIG. 6.3. Autocorrélations et corrélations croisées résiduelles de la série canadienne différenciée.

Pour la série américaine, les estimateurs sont donnés par :

$$\hat{\Theta}_0^{(2)} = \begin{bmatrix} 1.618 \\ 0.486 \\ -0.141 \end{bmatrix} \quad \hat{\Phi}_1^{(2)} = \begin{bmatrix} 0.000 & -0.000 & 0.000 \\ 0.000 & 0.576 & 0.510 \\ 0.000 & 0.122 & 0.880 \end{bmatrix},$$

$$\hat{\Phi}_2^{(2)} = \begin{bmatrix} 0.219 & 0.000 & -0.938 \\ 0.000 & 0.000 & 0.314 \\ 0.184 & 0.197 & 0.566 \end{bmatrix},$$

$$\hat{\Phi}_3^{(2)} = \begin{bmatrix} 0.000 & 0.000 & 0.563 \\ 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.422 \end{bmatrix},$$

$$\hat{\Sigma}_a^{(2)} = \begin{bmatrix} 0.686 & -0.078 & 0.058 \\ -0.078 & 0.386 & 0.036 \\ 0.058 & 0.036 & 0.153 \end{bmatrix}.$$

Les graphiques de la série résiduelle américaine et de ses corrélations sont données aux 6.4 et 6.5. À nouveau, les composantes de la série résiduelle semblent se comporter comme des bruits blancs et il n'y a que 4 corrélations se situant à l'extérieur des limites de significations à 5%.

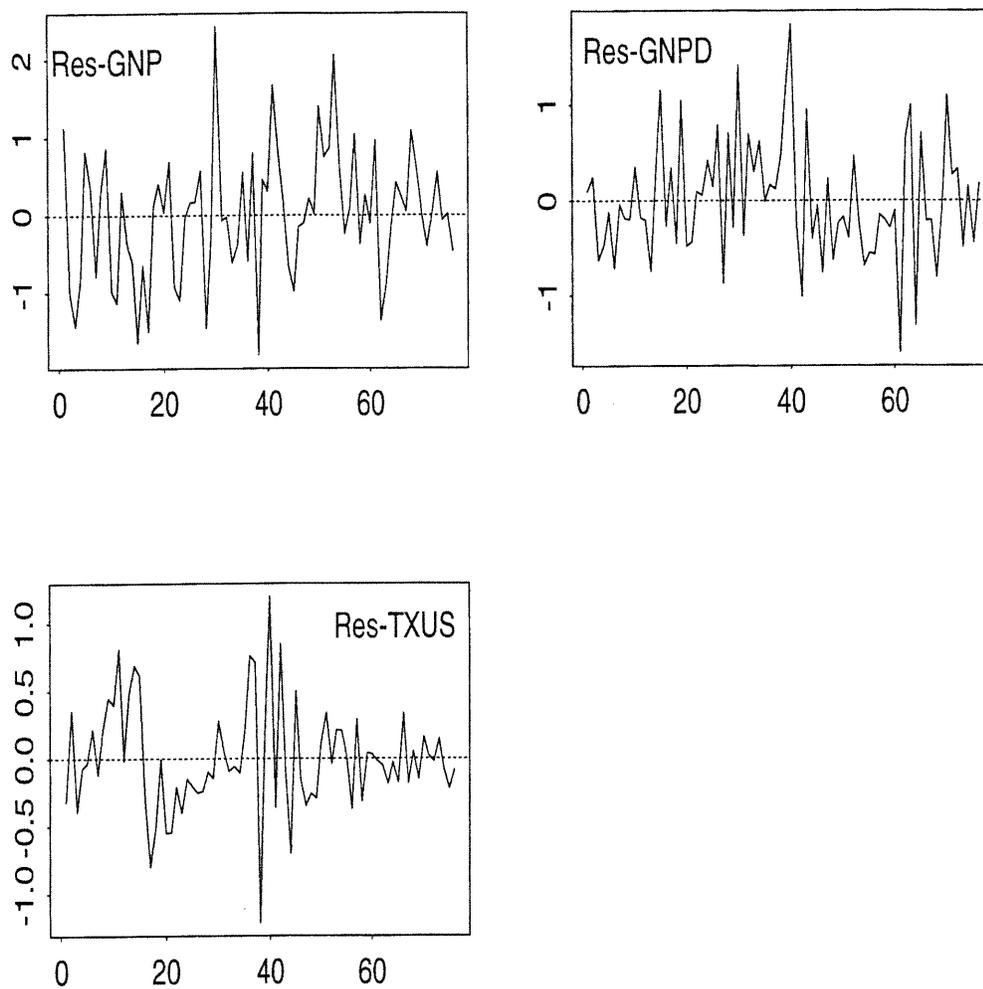


FIG. 6.4. Composantes de la série résiduelle correspondant à la série américaine différenciée.

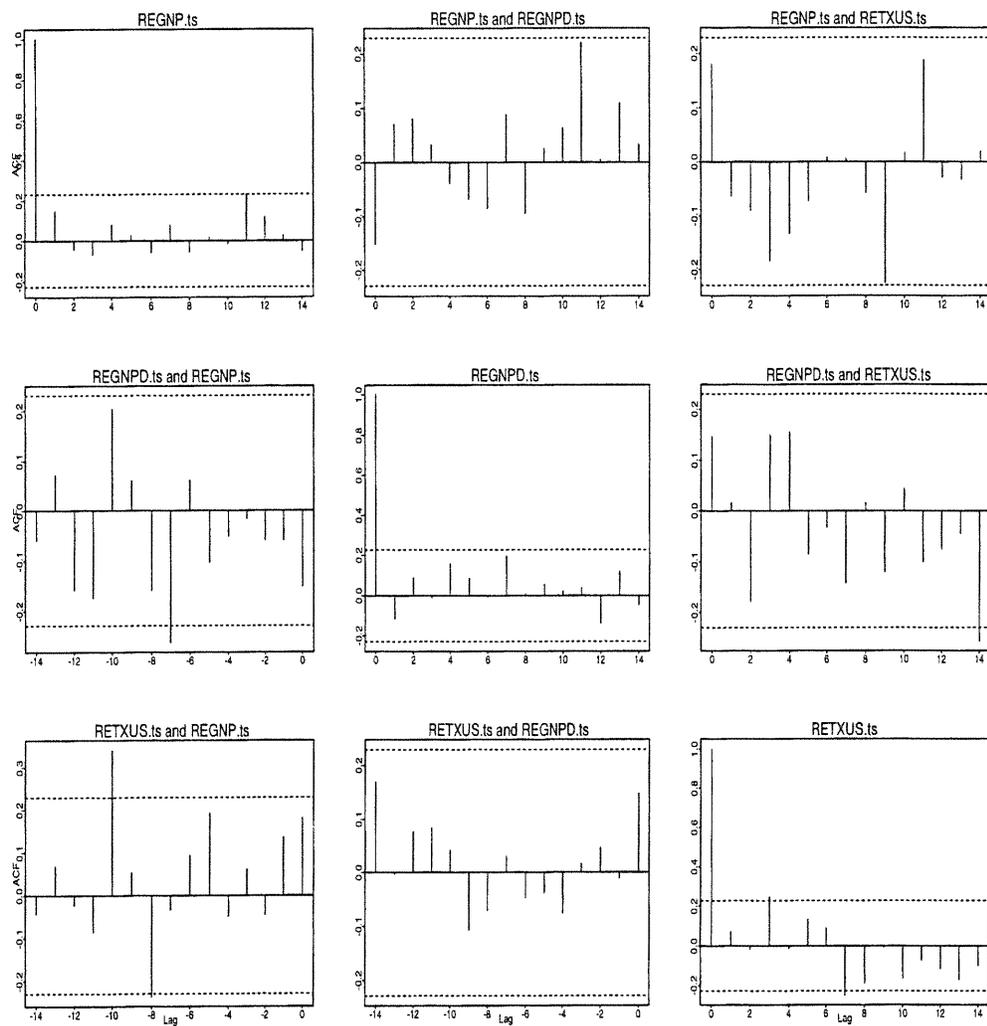


FIG. 6.5. Autocorrélations et corrélations croisées résiduelles de la série américaine différenciée.

## 6.2. TESTS DE NON CORRÉLATION

Les valeurs de la statistique  $Q_{\hat{\mathbf{a}}}^{(12)}(j)^*$  pour  $|j| = 0, 1, \dots, 12$  calculées à partir des séries résiduelles des deux modèles estimés sont représentées à la figure 6.6. La ligne pointillée horizontale représente la valeur critique au niveau de signification  $\alpha = 0.05$  pour tester l'hypothèse nulle  $\mathcal{H}_0$  de non corrélation entre les deux séries  $\{\mathbf{W}_t^{(1)}\}$  et  $\{\mathbf{W}_t^{(2)}\}$  contre l'alternative  $\mathcal{H}_{1j} : \boldsymbol{\rho}_{\mathbf{a}}^{(12)}(j) \neq \mathbf{0}$ . Cette valeur critique, indépendante du délai  $j$ , est égale à  $\chi_{12,0.95}^2 = 21.026$  et d'après la figure 6.6 nous constatons que  $\boldsymbol{\rho}_{\mathbf{a}}^{(12)}(j)$  est significativement différent de zéro pour les délais  $j=0$  et 2, et l'hypothèse de non corrélation sera probablement rejetée avec un test global. Nous pouvons aussi faire des tests simultanés sur plusieurs délais, par exemple les délais  $j = -M, \dots, M$  où  $M$  est un entier prédéterminé. L'hypothèse alternative est alors

$$\mathcal{H}_1^{(M)} : \text{il existe au moins un } j \text{ tel que } |j| \leq M \text{ et } \boldsymbol{\rho}_{\mathbf{a}}^{(12)}(j) \neq \mathbf{0}.$$

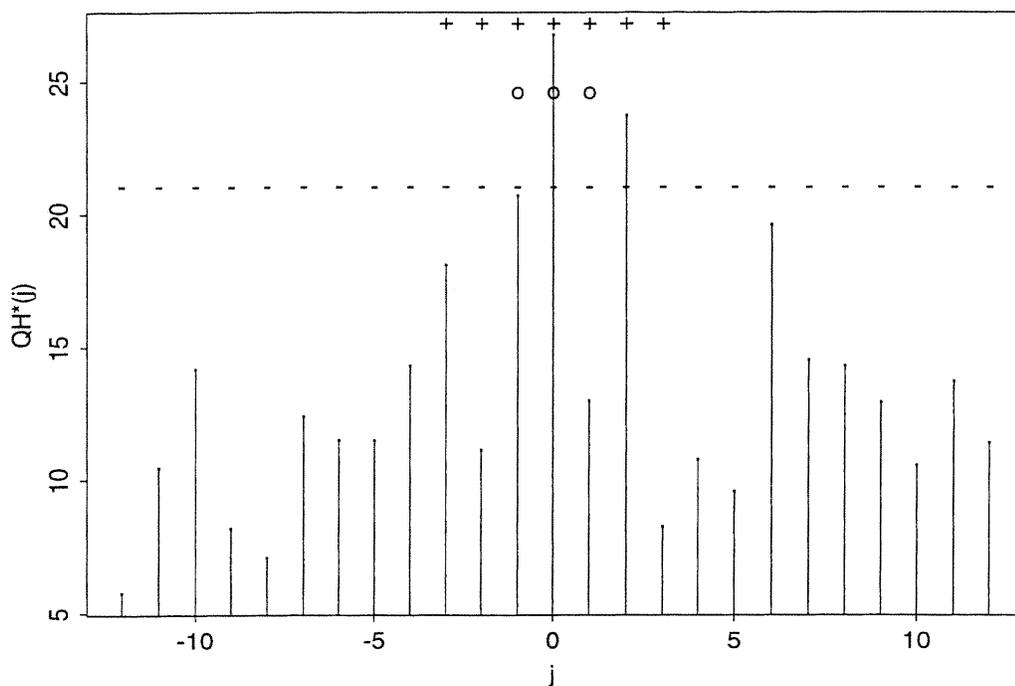
Comme les statistiques  $Q_{\hat{\mathbf{a}}}^{(12)}(j)^*$  sont asymptotiquement indépendantes et identiquement distribuées sous l'hypothèse nulle, il faut utiliser le niveau marginal  $\alpha_0 = 1 - (1 - \alpha)^{\frac{1}{2M+1}}$  afin d'avoir un niveau global égale à  $\alpha$ . Dans la figure 6.6, nous présentons les valeurs critiques au niveau global  $\alpha = 0.05$  pour  $M = 1$  et 3 qui sont respectivement 24.578 et 27.171. Comme  $Q_{\hat{\mathbf{a}}}^{(12)}(0)^* = 26.765$ , nous rejetons  $\mathcal{H}_0$  à  $M = 1$ , mais nous ne rejetons pas à  $M = 3$ .

Les résultats du test global basé sur la statistique portemanteau  $Q_{\hat{\mathbf{a}},M}^{(12)*}$  ainsi que les niveaux de signification empiriques  $\alpha_M$  correspondants sont donnés au tableau 6.2 pour  $M = 1, \dots, 12$ . Au niveau de signification  $\alpha = 0.05$ , l'hypothèse de non corrélation est rejetée pour toutes les valeurs de  $M$  inférieures ou égales

à 11.

Pour le test global basé sur les statistiques  $\mathcal{Q}_N$  et  $\mathcal{Q}_N^*$ , les valeurs des statistiques ainsi que les niveaux de significations empiriques  $\alpha_M$  sont donnés dans le tableau 6.2 pour  $M = 4, 7$  et 11. Ces choix de valeurs du paramètre de lissage  $M$  correspondent respectivement à  $[\ln(N)]$ ,  $[3N^{0.2}]$  et  $[3N^{0.3}]$ , où  $[x]$  représente la partie entière de  $x$ . Ainsi, pour  $M = 4$ , l'hypothèse de non corrélation est rejetée quand on considère la statistique  $\mathcal{Q}_N$  sauf pour les noyaux *BAR*. Pour la même valeur de  $M$ , on rejette aussi l'hypothèse nulle pour les noyaux *BAR* et *TH* quand on considère la statistique  $\mathcal{Q}_N^*$ . Quand  $M = 7$ , on rejette l'hypothèse de non corrélation pour tous les noyaux et pour les deux statistiques. Finalement, pour  $M = 11$ , on rejette l'hypothèse nulle pour tous les noyaux avec la statistique  $\mathcal{Q}_N$ . Avec la statistique modifiée  $\mathcal{Q}_N^*$ , il y a seulement le noyau *TR* avec lequel on ne rejette pas l'hypothèse nulle. Pour cette valeur de  $M$ , la différence entre les valeurs des statistiques  $\mathcal{Q}_N$  et  $\mathcal{Q}_N^*$  est probablement due au fait que les séries sont courtes ( $N = 68$ ) et que  $MS(k)$  et  $MD(k)$  ne fournissent pas de bonnes approximations de  $S_N(k)$  et  $D_N(k)$ . Il semble donc préférable pour des séries de cette longueur d'utiliser la statistique non modifiée.

FIG. 6.6. Valeurs de la statistique  $Q_{\hat{a}}^{(12)}(j)^*$  à différents délais pour les séries canadienne et américaine différenciées. La ligne horizontale pointillée représente la valeur critique marginale au niveau  $\alpha = 0.05$ . Les autres lignes représentent les valeurs critiques des tests simultanés aux délais  $|j| \leq M$  au niveau global  $\alpha = 0.05$  pour  $M = 1(-o-)$  et  $M = 3(-+-)$ .



M	$Q_{\hat{a},M}^*$	$\alpha_M$	M	$Q_{\hat{a},M}^*$	$\alpha_M$
1	60.422	0.007	7	225.707	0.012
2	95.266	0.003	8	247.067	0.021
3	121.615	0.005	9	268.165	0.035
4	146.668	0.008	10	292.873	0.039
5	167.708	0.020	11	317.050	0.045
6	198.807	0.012	12	334.216	0.085

TAB. 6.1. Valeurs de la statistique  $Q_{\hat{a},M}^*$  et niveaux empiriques en fonction de  $M$  pour la série canadienne et américaine différenciées.

Noyaux	TR	BAR	DAN	TH	BP
$Q_N(M = 4)$	2.922	1.241	2.191	1.750	2.672
$\alpha$	0.002	0.107	0.014	0.040	0.004
$Q_N^*(M = 4)$	3.586	1.374	1.931	1.658	2.474
$\alpha$	0.0001	0.085	0.027	0.05	0.007

Noyaux	TR	BAR	TH	DAN	BP
$Q_N(M = 7)$	2.640	2.343	2.916	2.779	2.964
$\alpha$	0.004	0.010	0.002	0.003	0.002
$Q_N^*(M = 7)$	2.680	2.225	2.367	2.560	2.537
$\alpha$	0.004	0.013	0.009	0.005	0.006

Noyaux	TR	BAR	TH	DAN	BP
$Q_N(M = 11)$	2.189	2.7043	2.918	2.969	2.838
$\alpha$	0.014	0.003	0.002	0.002	0.002
$Q_N^*(M = 11)$	1.544	2.324	1.950	2.564	2.054
$\alpha$	0.061	0.01	0.026	0.005	0.020

TAB. 6.2. Valeur des statistiques  $Q_N$  et  $Q_N^*$  avec  $M = 4, 7, 11$  pour la série canadienne et américaine différenciées.

## CONCLUSION

---

Dans cette thèse, nous avons proposé deux approches pour tester l'indépendance de deux séries multivariées stationnaires autorégressives d'ordre infini VAR( $\infty$ ).

La première approche est une généralisation de celle proposée par El Himdi et Roy (1997) qui teste la non corrélation entre deux séries ARMA stationnaires multivariées. Dans le cas de deux séries VAR( $\infty$ ), nous avons vu qu'un ensemble fini de corrélations croisées entre les séries résiduelles issues de l'approximation des deux séries par des autorégressions d'ordres finis possède une distribution asymptotique multinormale. Ceci nous a permis de déduire une procédure asymptotique pour tester l'hypothèse de non corrélation (ou d'indépendance dans le cas gaussien) par rapport à une alternative de corrélation à un délai particulier  $j$  entre les deux séries d'innovations correspondantes. En utilisant la statistique portemanteau  $Q_{\hat{\mathbf{a}}, M}^{(12)*}$ , nous pouvons confronter l'hypothèse nulle d'indépendance à l'alternative de corrélation sérielle croisée à un nombre fixé de délais, disons pour  $j = -M, \dots, M$ . L'approche utilisée ici ne nécessite pas l'estimation des vrais modèles ARMA décrivant les deux séries et nous protège ainsi d'erreurs éventuelles de spécification lors de la modélisation. Ceci est un avantage par rapport à la méthode d'El Himdi et Roy (1997) qui repose sur l'estimation des vrais modèles pour chacune des séries.

Les tests basés sur  $Q_{\hat{a}}^{(12)}(j)^*$ ,  $|j| \leq M$ , comme le test portemanteau  $Q_{\hat{a},M}^{(12)*}$  ne sont pas des tests convergents s'il y a corrélation sérielle à des délais  $j$  tels que  $|j| > M$  entre les deux séries d'innovations. C'est la raison pour laquelle dans la deuxième partie, nous avons proposé la statistique  $\mathcal{Q}_N$  qui prend en compte tous les délais possibles (les délais  $j$  tels que  $|j| \leq N - 1$  pour des séries de longueur  $N$ ). Les délais sont pondérés à l'aide d'une fonction de noyau et d'un point de troncature comme dans l'estimation de la densité spectrale. Il s'agit d'une version multivariée de la statistique de Hong (1996) et nous avons établi au chapitre 3 la normalité asymptotique sous l'hypothèse nulle.

Puisque ces tests sont basés sur des statistiques pondérées, nous avons étudié au chapitre 4 la puissance locale, selon l'approche de Pitman (1979), et la puissance globale selon l'approche de Bahadur (1960). Nous avons établi la normalité asymptotique de la statistique  $\mathcal{Q}_N$  sous des hypothèses alternatives locales. Nous avons aussi identifié le noyau optimal dans une certaine classe de noyaux. Pour une classe générale d'alternatives globales (fixe par rapport à  $N$ ), nous avons montré que le test  $\mathcal{Q}_N$  est convergent. Finalement, nous avons effectué une expérience de simulation afin d'étudier la précision de l'approximation des distributions exactes des statistiques  $Q_{\hat{a},M}^{(12)*}$ ,  $\mathcal{Q}_N$  et  $\mathcal{Q}_N^*$  par les distributions asymptotiques correspondantes. On a observé qu'avec les modèles choisis, le niveau est bien contrôlé avec tous les noyaux considérés et que les différents noyaux conduisent à des puissances similaires excepté le noyau uniforme tronqué pour lequel la puissance est moindre. En pratique, il semble donc que si l'on veut faire un test global, on a avantage à utiliser la statistique  $\mathcal{Q}_N$  avec un noyau autre que le noyau uniforme tronqué qui produit un test moins puissant. Étant donné sa forme simple, le noyau de Bartlett semble un choix judicieux. De l'expérience de simulation, il se dégage

aussi que la statistique  $Q_N$  et sa version modifiée  $Q_N^*$  conduisent à des résultats très similaires, tout au moins pour les longueurs de séries considérées.

Dans ce travail, nous avons fait l'hypothèse de stationnarité des deux séries d'intérêt. Cependant, une majorité de séries réelles sont non stationnaires. Par exemple, en économie, plusieurs séries multivariées sont cointégrées, c'est à dire que chaque composante peut être non stationnaire mais il existe au moins une combinaison qui forme une série univariée stationnaire. En s'appuyant sur Pham, Roy et Cédras (2001) et sur Saikkonen (1992), les résultats théoriques de cette thèse devraient pouvoir s'étendre au cas de deux séries partiellement non stationnaires ou cointégrées comme c'est souvent le cas avec des données économiques.

## BIBLIOGRAPHIE

---

- Anderson, T. W. (1971), *The Statistical Analysis of Time Series*, Wiley and Sons: New York.
- Andrews, D. W. K. (1991), Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica* **58**, 817–858.
- Ansley, C. F. (1980), Computation of the theoretical autocovariance function for a vector ARMA process, *Journal of Statistical Computation and Simulation* **12**, 15–24.
- Bahadur, R. R. (1960), Stochastic comparison of tests, *Annals of Mathematical Statistics* **31**, 276–295.
- Berk, K. N. (1974), Consistent autoregressive spectral estimates, *Annals of Statistics* **2**, 489–502.
- Brillinger, D. R. (1981), *Time Series: Data Analysis and Theory.*, New York: Holden-Day.
- Brockwell, P. J. et Davis, R. A. (1991), *Times Series: Theory and Methods*, second edn, Springer-Verlag: New York.
- Brown, B. (1971), Martingale central limit theorems, *Annals of Mathematical Statistics* **42**, 59–66.
- Chitturi, R. (1976), Distribution of multivariate white noise autocorrelations, *Journal of the American Statistical Association.* **71**, 928–934.
- Cléroux, R. et Roy, R. (1988), On vector autocorrelation in multivariate time series, *1987 Proceedings of the Business and Economic Statistics Section, ASA, Washington, DC.* pp. 654–658.
- El Himdi, K. et Roy, R. (1997), Tests for noncorrelation of two multivariate arma time series, *Canadian Journal of Statistics* **25**, 233–256.
- Escoufier, Y. (1973), Le traitement des variables vectorielles, *Biometrics* **29**, 751–760.
- Fuller, W. A. (1996), *Introduction to Statistical Time Series*, second edn, Wiley and Sons: New York.
- Gallant, A. R. et White, H. (1988), *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models.*, Basil Blackwell: Oxford.

- Gallant, N. H. et Jorgenson, D. (1979), Statistical inference for a system of simultaneous non-linear implicit equations in the context of instrumental variables estimation, *Journal of Econometrics* **11**, 275–302.
- Ghosh, B. K. et Huang, W. M. (1991), The power and optimal kernel of the Bickel-Rosenblatt test for goodness of fit, *Annals of Statistics* **17**, 999–1009.
- Hannan, E. (1970), *Multiple Time Series*, Wiley and Sons: New York.
- Haugh, L. D. (1976), Checking the independence of two covariance-stationary time series: a univariate residual cross correlation approach, *Journal of the American Statistical Association*. **71**, 376–385.
- Haugh, L. D. et Box, G. (1977), Identification of dynamic regression (distributed lag) models connecting two time series, *Journal of the American Statistical Association*. **72**, 121–130.
- Hong, Y. (1996), Testing for independence between two covariance stationary time series, *Biometrika* **83**, 615–625.
- Hosking, J. (1980), The multivariate portemanteau statistic, *Journal of the American Statistical Association*. **75**, 602–608.
- Kheoh, T.S. et McLeod, A.I. (1992), Comparison of two modified portmanteau tests for model adequacy, *Computational Statistics and Data Analysis* **14**, 99–106.
- Lewis, R. et Reinsel, G. (1985), Prediction of multivariate time series by autoregressive model fitting, *Journal of Multivariate Analysis* **16**, 393–411.
- Li, W. et McLeod, A. (1981), Distribution of the residual autocorrelation in multivariate ARMA time series models, *Journal of the Royal Statistical Society B* **43**, 231–239.
- Lütkepohl, H. (1989), A note on the asymptotic distribution of impulse response functions of estimated VAR models with orthogonal residuals, *Journal of Econometrics* **42**, 371–376.
- Lütkepohl, H. (1991), *Introduction to Multiple Time Series Analysis*, Berlin: Springer-Verlag.
- McLeod, I.A. (1979), Distribution of the residual cross correlation in univariate ARMA time series models, *Journal of the American Statistical Association* **74**, 849–855.
- Paparoditis, E. (1996), Bootstrapping autoregressive and moving average parameter estimates of infinite order vector autoregressive processes, *Journal of Multivariate Analysis* **57**, 277–296.
- Parzen, E. (1957), On consistent estimates of the spectrum of a stationary time series, *Annals of Mathematical Statistics* **28**, 329–348.

- Pham, D., Roy, R. et Cédras, L. (2001), Tests for non-correlation of two cointegrated ARMA time series, *Journal of Time Series Analysis*, à paraître.
- Pitman, E. J. G. (1979), *Some Basic Theory for Statistical Inference*, Chapman Hall: London.
- Priestley, M. B. (1981), *Spectral Analysis and Time Series*, Vol. 1 et 2, London: Academic Press.
- Racette, D. et Raynauld, J. (1992), Un modèle bivar de prévision de la dépense nominale et d'analyse de la politique monétaire canadienne, *Séminaire sur les questions monétaires, Banque du Canada, mai 1992*, 347–425.
- Robinson, P. M. (1991), Consistent nonparametric entropy-based testing, *Review of Economic Studies* **58**, 437–425.
- Roy, R. (1989), Asymptotic covariance structure of serial correlation in multivariate time series, *Biometrika* **76**, 824–827.
- Roy, R. et Cléroux, R. (1993), On vector cross-correlation in time series and applications, *International Statistical Review* **61**, 447–464.
- Saikkonen, P. (1992), Estimation and testing of cointegrated systems by an autoregressive approximation, *Econometric Theory* **8**, 1–27.
- Taniguchi, M. et Kakizawa, Y. (2000), *Asymptotic Theory of Statistical Inference for Time Series*, Springer-Verlag: New York.
- Venables, R. et Ripley, M. (2000), *The S Programming*, Wadsworth and Brooks/Cole: California.
- White, H. (1984), *Asymptotic Theory for Econometricians*, Academic Press: London.