

Université de Montréal

Tests d'ajustement reposant sur les méthodes
d'ondelettes dans les modèles ARMA avec un terme
d'erreur qui est une différence de martingales
conditionnellement hétéroscédastique

par

Chu Pheuil Liou

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

Mémoire présenté à la Faculté des études supérieures et postdoctorales
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en Statistique

septembre 2019

© Chu Pheuil Liou, 2019

Sommaire

L'étude porte sur le développement d'une procédure à base d'ondelettes afin de tester la qualité d'ajustement d'un modèle autorégressif moyenne mobile (ARMA), où l'innovation est une différence de martingales avec hétéroscédasticité conditionnelle. Dans le cadre du modèle ARMA semi-fort, les innovations ne sont pas indépendantes, mais sont présumées des différences de martingales, voir par exemple Francq et al. (2005). Conséquemment, les distributions asymptotiques de type chi-carré pour les statistiques de test de Box-Pierce-Ljung, établies sous l'hypothèse d'indépendance du terme d'erreur, ne sont pas forcément adéquates. De plus, les irrégularités saisonnières dans la densité spectrale de l'innovation pourraient affecter la puissance des tests classiques. Ceci motive l'étude des méthodes d'ondelettes dans le présent contexte, où le terme d'erreur est présumé plus général que composé de variables aléatoires indépendantes.

Afin de trouver la distribution asymptotique de nouveaux tests reposant sur les ondelettes, nous commençons par établir la distribution asymptotique des autocovariances résiduelles basées sur les résidus du modèle ajusté. La densité spectrale exprimée en base d'ondelettes repose sur des coefficients, dits coefficients d'ondelettes. Nous trouvons les distributions asymptotiques des versions empiriques de ces coefficients d'ondelettes lorsque le modèle ARMA est correctement spécifié. Nous montrons que les variances et covariances asymptotiques des coefficients d'ondelettes empiriques dépendent des coefficients du modèle ARMA ainsi que de la structure de covariance du moment d'ordre deux des innovations.

Ces résultats sont utilisés pour construire de nouveaux tests. Nous démontrons que ces statistiques de tests suivent des distributions de type chi-carré asymptotiquement. La performance des statistiques de test proposées pour plusieurs tailles d'échantillon sont

étudiées par des simulations de Monte Carlo. Les résultats de nos simulations suggèrent que les nouveaux tests convergent raisonnablement rapidement vers des distributions de type chi-carré, et qu'ils offrent des propriétés de puissance avantageuses comparativement aux statistiques de test classiques, au moins pour certains modèles. Un exemple avec des données réelles illustre la méthodologie.

Titre court: Tests d'ajustement à base d'ondelettes dans les modèles ARMA.

2010 Mathematics Subject Classification: Primaire: 62M10; Secondaire: 62M15

Mots clés: Qualité d'ajustement; Modèle ARMA; Différence de martingales conditionnellement hétéroscédastique; Méthode d'ondelettes; Autocorrélation résiduelle; Densité spectrale.

Summary

Wavelet-based test procedures for lack-of-fit in autoregressive moving average (ARMA) models with conditionally heteroskedastic martingale difference innovations are investigated. In the framework of so-called semistrong ARMA models, the innovations are not independent, but are presumed a martingale difference, see for example Francq et al. (2005). Consequently, the Chi-square distributions of the popular Box-Pierce-Ljung test statistics, established under the hypothesis of independence in the error term, are not necessarily adequate. Moreover, the seasonal irregularities in the spectral density of the innovation can affect the power of the classical tests. This provides motivations for studying wavelet methods in the current context, where the error term is presumed more general than independent random variables.

To find the asymptotic distribution of the new wavelet-based tests, we begin by establishing the asymptotic distribution of the residual autocovariances based on the residuals of the fitted model. The wavelet-based spectral density relies on certain coefficients, the so-called wavelet coefficients. We derive the asymptotic distribution of the empirical versions of those wavelet coefficients when the ARMA model is correctly specified. We show that the asymptotic variances and covariances of the empirical wavelet coefficients depend on both the coefficients of the ARMA model, and the covariance structure of the second moment of the innovations.

Those results are used to construct new tests. We demonstrate that those test statistics have Chi-square distributions asymptotically. The performances of the proposed test statistics are investigated via Monte Carlo simulation studies for several sample sizes. The results of our simulations suggest that the new tests converge reasonably quickly to

the Chi-square distributions and offer favorable power properties over the classical statistical tests, at least in some models. An example with real data illustrates the methodology.

Short title: Wavelet-based testing for lack of fit in ARMA models.

2010 Mathematics Subject Classification: Primary: 62M10; Secondary: 62M15

Keywords: Lack of fit tests; ARMA model; Conditionally heteroscedastic martingale difference; Wavelet method; Residual autocorrelation; Spectral density.

Table des matières

Sommaire	iii
Summary	v
Liste des tableaux	ix
Liste des figures	xiii
Remerciements	xv
Chapitre 1. Introduction	1
Chapitre 2. Préliminaires	7
2.1. Processus stationnaires	7
2.2. Processus ARMA et ARIMA	8
2.3. Processus ARMA semi-fort	17
2.4. Distribution limite dans les modèles ARMA	21
2.5. Introduction aux ondelettes	25
2.6. Ondelettes et estimateurs de densité spectrale à base d'ondelettes	30
2.7. Distributions asymptotiques des coefficients d'ondelettes empiriques	37
2.8. Modèle GARCH	39
Chapitre 3. Statistiques de test	43
3.1. Tests de Box-Pierce-Ljung	43

3.2. Tests de Francq-Roy-Zakoïan	45
3.3. Nouveaux tests reposant sur les ondelettes	51
Chapitre 4. Expériences empiriques	57
4.1. Étude des niveaux empiriques	58
4.2. Discussion de l'étude sur le niveau exact	62
4.3. Discussion de l'étude sur la puissance	69
Chapitre 5. Exemple avec données réelles	77
Conclusion	83
Annexe	85
Bibliographie	89

Liste des tableaux

- 4.1 Niveaux empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(1) avec des innovations GARCH(1,1), $\phi \in \{0,1; 0,2; 0,3\}$, $\alpha \in \{0,3; 0,6\}$, $\beta = 0,1$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$ 63
- 4.2 Niveaux empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(1) avec des innovations GARCH(1,1), $\phi \in \{0,1; 0,2; 0,3\}$, $\alpha = 0,1$, $\beta \in \{0,3; 0,6\}$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$ 64
- 4.3 Niveaux empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(2) avec des innovations GARCH(1,1), $\phi \in \{0,1; 0,2; 0,3\}$, $\alpha \in \{0,3; 0,6\}$, $\beta = 0,1$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$ 66
- 4.4 Niveaux empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(2) avec des innovations GARCH(1,1), $\phi \in \{0,1; 0,2; 0,3\}$, $\alpha = 0,1$, $\beta \in \{0,3; 0,6\}$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$ 67
- 4.5 Niveaux empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(1) avec des innovations iid, $\phi \in \{0,1; 0,2; 0,3\}$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$ 68
- 4.6 Niveaux empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(2) avec des innovations iid, $\phi \in \{0,1; 0,2; 0,3\}$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$ 69

4.7	Puissances empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(1)×AR(4) avec des innovations GARCH(1,1), $(\phi_1; \phi_4) \in \{(0,2; 0,1); (0,3; 0,1); (0,3; 0,2)\}$, $\alpha \in \{0,3; 0,6\}$, $\beta = 0,1$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$	71
4.8	Puissances empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(1)×AR(4) avec des innovations GARCH(1,1), $(\phi_1; \phi_4) \in \{(0,2; 0,1); (0,3; 0,1); (0,3; 0,2)\}$, $\alpha = 0,1$, $\beta \in \{0,3; 0,6\}$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$	72
4.9	Puissances empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(2)×AR(4) avec des innovations GARCH(1,1), $(\phi_2; \phi_4) \in \{(0,2; 0,1); (0,3; 0,1); (0,3; 0,2)\}$, $\alpha \in \{0,3; 0,6\}$, $\beta = 0,1$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$	73
4.10	Puissances empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(2)×AR(4) avec des innovations GARCH(1,1), $(\phi_2; \phi_4) \in \{(0,2; 0,1); (0,3; 0,1); (0,3; 0,2)\}$, $\alpha = 0,1$, $\beta \in \{0,3; 0,6\}$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$	74
4.11	Puissances empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(1)×AR(4) avec des innovations iid, $(\phi_1, \phi_4) \in \{(0,2; 0,1); (0,3; 0,1); (0,3; 0,2)\}$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$	75
4.12	Puissances empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(2)×AR(4) avec des innovations iid, $(\phi_1, \phi_4) \in \{(0,2; 0,1); (0,3; 0,1); (0,3; 0,2)\}$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$	76
5.1	Valeurs-p des différents tests pour les résidus des rendements du S&P500 avec l'innovation suivant un modèle GARCH(0,0) et les rendements suivant des modèles (a) AR(0), (b) AR(1), (c) AR(2) et (d) AR(4).	80

5.2	Valeurs d'autocorrélation des résidus des rendements du S&P500 suivant des modèles AR avec l'innovation suivant des modèles GARCH.	81
-----	---	----

Liste des figures

2.1	Données quotidiennes à la fermeture d'Amazon du 17 septembre 2008 jusqu'au 17 septembre 2018.....	13
2.2	Données quotidiennes à la fermeture d'Amazon du 17 septembre 2008 jusqu'au 17 septembre 2018 après une première différentiation.....	14
2.3	Données quotidiennes à la fermeture d'Amazon du 17 septembre 2008 jusqu'au 17 septembre 2018 après une transformation logarithmique et une première différentiation.....	15
5.1	Graphique d'autocorrélation des résidus pour les rendements du S&P500 avec l'innovation suivant un modèle GARCH(0,0) et les rendements suivant des modèles (a) AR(0), (b) AR(1), (c) AR(2) et (d) AR(4).....	81
5.2	Graphique d'autocorrélation des résidus non-standardisés pour les rendements du S&P500 avec l'innovation suivant un modèle GARCH(1,1) et les rendements suivant des modèles (a) AR(0), (b) AR(1), (c) AR(2) et (d) AR(4).....	82
5.3	Graphique d'autocorrélation des résidus standardisés pour les rendements du S&P500 avec l'innovation suivant un modèle GARCH(1,1) et les rendements suivant des modèles (a) AR(0), (b) AR(1), (c) AR(2) et (d) AR(4).....	82

Remerciements

Je tiens à remercier mon directeur de recherche, Monsieur Pierre Duchesne, pour sa supervision, sa patience et son soutien tout au long du mémoire. Grâce à lui, j'en ressors plus confiant, discipliné et plus mature, en plus de toutes les connaissances acquises.

Je tiens à remercier ma femme pour son support et mes beaux-parents pour leur patience afin de me laisser compléter ce mémoire.

Je tiens aussi à remercier mes parents et les fonds de recherche de mon directeur pour le soutien financier.

Chapitre 1

Introduction

Une étape fondamentale dans l'étude de la qualité d'ajustement d'un modèle de séries chronologiques est d'étudier le comportement des résidus du modèle ajusté. En particulier, les statistiques de test de Box-Pierce-Ljung développées lors des travaux fondateurs de Box et Pierce (1970) et Ljung et Box (1978) ont été largement utilisées lors de la modélisation des modèles de séries chronologiques autorégressifs et moyennes mobiles (ARMA). Des références utiles incluent Davies, Triggs et Newbold (1977), Li et McLeod (1981), Robinson (1991), Anderson (1993), Hong (1996), Andrews et Ploberger (1996), Lobato et Robinson (1998), Hong (1999), Paparoditis (2000), Lee et Hong (2001), entre autres. Le livre de Li (2004) décrit plusieurs de ces statistiques de test et présente aussi d'autres procédures populaires dans les modèles univariés et multivariés.

Dans un modèle ARMA, l'innovation (ou le bruit) est typiquement présumée seulement non-corrélée ou bruit blanc (ARMA faible tel que défini dans Francq et al. (2005)). Néanmoins, lors des études théoriques de la plupart des statistiques de test courantes (par exemple, les tests portemanteaux cités précédemment), les distributions asymptotiques sont obtenues sous des hypothèses plus restrictives sur le terme d'innovation. Typiquement, le modèle ARMA est présumé fort, ce qui veut dire que le terme d'erreur est supposé être composé de variables aléatoires indépendantes et identiquement distribuées (iid).

Les propriétés générales du modèle ARMA sont valides sous des conditions beaucoup plus larges sur le terme d'erreur. Par exemple, la structure de corrélation du modèle ARMA nécessite seulement que le bruit blanc soit faible. Cependant, dans l'étude statistique des estimateurs et lors de l'inférence statistique, l'étude mathématique dans un

cadre asymptotique est grandement facilitée en supposant davantage, comme par exemple l'hypothèse d'indépendance sur le bruit.

A priori, rien n'assure que les distributions asymptotiques des tests classiques obtenues sous des hypothèses restrictives, comme l'hypothèse d'indépendance, continuent d'être valides, même approximativement, lorsque le bruit est non-corrélé mais dépendant. En fait, il a été trouvé que les distributions asymptotiques peuvent être assez différentes dans le cas général, comparativement aux lois trouvées sous l'hypothèse d'indépendance du bruit. Francq et al. (2005) fournissent plusieurs exemples de processus ARMA avec innovations non-corrélées mais non indépendantes.

Comme mentionné précédemment, la classe des modèles ARMA dans sa forme la plus générale suppose que le terme d'erreur est simplement non-corrélé. De ce point de vue, supposer un bruit blanc iid peut être considéré comme une supposition assez restrictive. D'un autre côté, on peut se demander l'importance ou la pertinence de permettre une structure de dépendance plus générale sur le bruit, car supposer seulement un bruit blanc non-corrélé est tellement général que cela complique l'étude asymptotique (l'hypothèse d'indépendance simplifie l'obtention des résultats théoriques) ainsi que la mise en œuvre pratique (des échantillons plus grands sont nécessaires afin d'estimer les différentes quantités avec précision). Cependant, les applications économiques suggèrent que plusieurs séries chronologiques financières ou économiques comme les rendements des actions et les taux de change ne sont pas corrélées, mais ne sont pas indépendantes.

De nombreux modèles de séries chronologiques populaires en économétrie et en finance satisfont la propriété de différence de martingales. C'est le cas des modèles autorégressifs conditionnellement hétéroscédastiques (ARCH) ou leur version généralisée (notamment les GARCH, ou ARCH généralisés, et autres variantes). Formellement, une différence de martingales est, conditionnellement au passé, de moyenne nulle, mais la variance conditionnelle peut être fonction du temps. De ce point de vue, il semble approprié de présumer que les processus ARMA ont un terme d'erreur satisfaisant la propriété différence de martingales. Pour les raisons mentionnées précédemment, Durlauf (1991) considère des tests basés sur des spectres pour les corrélations sérielles sous l'hypothèse d'une différence de

martingales conditionnellement homogène, au lieu de présumer l'hypothèse d'indépendance au terme d'erreur. Néanmoins, la plupart des séries financières et économiques qui sont sous l'hypothèse d'une différence de martingales manifestent de l'hétéroscédasticité conditionnelle.

Ainsi, supposer que le terme d'erreur soit une différence de martingales permettant de l'hétéroscédasticité conditionnelle semble être un cadre réaliste. De plus, compte tenu de la popularité des modèles de types ARCH et GARCH, il ne semble pas trop restrictif de présumer que le terme d'erreur soit dans une classe qui inclue ces modèles. Deo (2000) a étudié les tests de corrélation sérielle sous une différence de martingales conditionnellement hétéroscédastique. Il a obtenu la distribution asymptotique de la fonction de distribution spectrale standardisée d'un échantillon sous cette hypothèse. Lobato et al. (2001, 2002) proposent quant à eux un test portemanteau modifié pour inclure des dépendances non-linéaires. Horowitz et al. (2006) considèrent un test de Box-Pierce robuste sous des hypothèses faibles et générales. Escanciano et Lobato (2009) introduisent un test de Box-Pierce qui repose sur les données directement observées (et donc qui ne peut être directement appliqué sur les résidus des modèles) pour corrélation sérielle sous une différence de martingales ergodique strictement stationnaire.

Puisque tous les tests mentionnés précédemment dépendent d'hypothèses techniques sur la manière dont les données sont générées, où certains modèles GARCH communs sont exclus, Shao (2011) considère les tests de bruits blancs sous des hypothèses plus faibles qu'une différence de martingales. Puisque ces hypothèses de bruits blancs faibles sont plus générales, il a utilisé une procédure bootstrap pour estimer la distribution asymptotique sous l'hypothèse nulle. Les tests mentionnés précédemment sont construits directement sur les données brutes, ce qui implique que ce ne sont pas des résidus de modèles qui reposent sur une phase d'estimation. Ainsi, ces tests ne peuvent être utilisés directement sur les résidus des modèles ARMA avec innovations dépendantes, car il est possible que l'étape d'estimation, et l'utilisation de résidus, affectent la distribution asymptotique des tests justifiés avec des données brutes.

Un modèle qui n'ajuste pas adéquatement les données devrait afficher des corrélations résiduelles significativement différentes de zéro. Ainsi, le test pour la qualité d'ajustement du modèle peut être effectué pour tester la non-corrélation des résidus. Les tests de corrélation sérielle existants peuvent être essentiellement classés en deux catégories, soit dans le domaine temporel avec des tests basés sur les corrélations ou soit dans le domaine fréquentiel avec des tests basés sur les estimateurs de la densité spectrale. Les méthodes dans le domaine fréquentiel reçoivent une attention assez soutenue ces dernières années parce que ces tests sont typiquement convergents pour une grande classe d'hypothèses alternatives. Habituellement, les procédures de test dans le domaine des fréquences dépendent des distances normalisées entre des estimateurs de la densité spectrale à base de noyaux et la densité spectrale sous l'hypothèse nulle de non-corrélation sérielle, voir, par exemple, Hong (1996).

Les méthodes d'ondelettes représentent une approche alternative aux estimateurs de la densité spectrale à base de noyaux. Elles utilisent l'expansion d'ondelettes de la densité spectrale. Lee et Hong (2001) ont construit des estimateurs de densité spectrale à base d'ondelettes et ils ont obtenu des statistiques de test convergentes pour la corrélation sérielle en utilisant des mesures quadratiques intégrées. Les méthodes à base d'ondelettes ont généré une littérature appréciable dans l'étude des modèles de séries chronologiques, avec applications en économétrie et en finance. Lorsque l'intérêt est la description de la dépendance, l'engouement des méthodes d'ondelettes peut s'expliquer car elles fournissent des estimateurs de la densité spectrale s'adaptant bien à des situations où il y a moins de régularité ou plus d'inhomogénéité. Ces situations surviennent lorsque les processus sous-jacents possèdent des composantes saisonnières ou périodiques persistantes.

Plusieurs statistiques de test à base d'ondelettes ont été proposées dans différents contextes. Par exemple, Hong et Kao (2004) ont proposé des statistiques de test à base d'ondelettes pour la corrélation sérielle de forme inconnue dans des modèles de séries chronologiques pour données de panel. Li, Yao et Duchesne (2014) ont proposé des tests convergents pour la corrélation sérielle basés sur les coefficients d'ondelettes empiriques et la méthode adaptative de Neyman étudiée dans Fan (1996). Gençay et Signori (2015) ont étudié une transformée en ondelettes pour introduire une nouvelle famille de tests

portemanteaux pour corrélation sérielle dont les tests sont parfois plus puissants que les statistiques de test de Box-Pierce-Ljung.

Dans ce mémoire, suivant le cadre théorique de Francq et al. (2005), nous vérifions la qualité d'ajustement d'un modèle ARMA quand l'innovation est une suite de variables aléatoires de type différence de martingales (ou modèle ARMA semi-fort) et que les tests reposent sur les résidus du modèle ajusté. L'analyse asymptotique tient compte de ce contexte. Pour un modèle que l'on suppose correctement spécifié, nous montrons que les autocovariances empiriques des résidus suivent conjointement une distribution normale et que les coefficients d'ondelettes empiriques de l'expansion d'ondelettes de la densité spectrale pour l'innovation suivent aussi asymptotiquement une distribution normale. Les résultats établis dans Li, Yao et Duchesne (2014) reposaient sur une hypothèse de séries chronologiques iid sous l'hypothèse nulle. Contrairement à leurs résultats, établis dans ce travail en construisant des tests sur des résidus de modèles ARMA, nous montrons que la variance et covariance asymptotiques pour les coefficients d'ondelettes empiriques dépendent autant des coefficients du modèle ARMA que de la structure de covariance du moment d'ordre deux des innovations. De ces résultats, nous proposons de nouveaux tests reposant sur les coefficients d'ondelettes et démontrons qu'ils suivent asymptotiquement des distributions de type chi-carré. Les résultats de nos simulations de Monte Carlo suggèrent que les nouvelles statistiques de test possèdent des avantages en matière de puissance sur les statistiques de test classiques, au moins pour certains modèles.

L'organisation du mémoire se présente comme suit. Dans le chapitre 2, nous fournissons des concepts préliminaires, entre autres, sur les paramètres estimés dans l'ajustement du modèle ARMA, la distribution limite de l'autocovariance résiduelle, les ondelettes et les estimateurs de densité spectrale à base d'ondelettes. La distribution normale asymptotique des coefficients d'ondelettes empiriques est aussi obtenue. Les processus autorégressifs conditionnellement hétéroscédastiques généralisés (GARCH) sont aussi présentés. Dans le chapitre 3, certains tests portemanteaux sont présentés et les nouvelles statistiques de test sont proposées basées sur les résultats asymptotiques. Dans le chapitre 4, des simulations sont fournies et nous illustrons que les nouveaux tests sont comparables aux tests courants en terme de niveaux et plus puissants pour certains modèles. Dans le chapitre 5, un

exemple avec des données réelles est présenté. Nous concluons avec des remarques et toutes les preuves techniques des résultats principaux sont fournies en annexe.

Le mémoire s'articule autour d'un article écrit en collaboration avec Linyuan Li du département de mathématiques et statistique de *l'University of New Hampshire*. Comme co-auteur, j'ai vérifié certains résultats théoriques et j'ai apporté des précisions. J'ai réalisé l'ensemble des simulations. L'analyse de données réelles provient de mes recherches (il n'est pas certain que cette application sera incluse dans l'article final).

Chapitre 2

Préliminaires

2.1. Processus stationnaires

Intuitivement, un processus est stationnaire si les propriétés probabilistes du processus, par exemple sa moyenne et sa variance, ne varient pas avec le temps. De plus, on peut définir le concept d'autocorrélation. Avant d'effectuer l'analyse d'une série chronologique, il s'avère utile de s'assurer que le processus soit stationnaire, car ceci rend possible l'inférence statistique et alors sous des conditions assez générales on peut estimer de façon appropriée la moyenne du processus, sa variance ainsi que les corrélations par des versions empiriques. Dans notre contexte, la définition qui nous intéresse est celle de la stationnarité au second ordre, appelée aussi stationnarité au sens large. Elle fait l'objet de la définition suivante.

Définition 2.1.1. *Un processus $\{X_t\}$ est défini comme stationnaire au second ordre (ou stationnaire au sens large) si:*

- (i) $E(X_t) = \mu_X$, pour tout $t \in \mathbb{Z}$, où μ_X est une constante finie,
- (ii) $E(X_t^2) < \infty$, $\forall t \in \mathbb{Z}$,
- (iii) $\text{Cov}(X_t, X_{t+l}) = c_X(l)$, $\forall t, l \in \mathbb{Z}$;

on dira que $c_X(l)$ est la fonction d'autocovariance du processus $\{X_t\}$, indexée par l .

La fonction d'autocovariance repose sur l , ce que nous appelons le délai. On note que $c_X(\cdot)$ est symétrique comme fonction du délai, c'est-à-dire $c_X(l) = c_X(-l)$, $l \in \mathbb{Z}$. On définit la fonction d'autocorrélation de délai l par $\rho_X(l) = c_X(l)/c_X(0)$, $l \in \mathbb{Z}$.

Les moments d'ordre un et deux d'un processus stationnaire au second ordre ne dépendent pas du temps, ce qui permet de faire de l'inférence sur une réalisation finie, ce que nous appelons une série chronologique. Cette définition de la stationnarité diffère de la définition de la stationnarité au sens strict, qui suppose que la distribution conjointe d'un processus ne change pas dans le temps, c'est-à-dire, que la distribution conjointe de X_1, \dots, X_n est la même que pour X_{1+l}, \dots, X_{n+l} , pour tous $n \geq 1$ et $l \in \mathbb{Z}$.

Basé sur une série chronologique X_1, \dots, X_n , un estimateur naturel pour μ_X est

$$\hat{\mu}_X = n^{-1} \sum_{t=1}^n X_t, \quad (2.1.1)$$

alors que l'on peut estimer $c_X(l)$ comme suit:

$$\hat{c}_X(l) = \begin{cases} n^{-1} \sum_{t=1}^{n-l} (X_t - \hat{\mu}_X)(X_{t+l} - \hat{\mu}_X), & l = 0, \dots, n-1, \\ \hat{c}_X(-l), & l = -1, \dots, -n+1. \end{cases} \quad (2.1.2)$$

La fonction $\hat{c}_X(l)$, $|l| < n$, est appelée fonction d'autocovariance échantillonnale. Utilisant cette dernière, on peut estimer l'autocorrélation échantillonnale de délai l :

$$\hat{\rho}_X(l) = \hat{c}_X(l)/\hat{c}_X(0), \quad |l| < n. \quad (2.1.3)$$

Les fonctions d'autocovariance et d'autocorrélation, et leurs versions échantillonnales, sont particulièrement importantes dans l'étude des tests portemanteaux. Dans les applications, il a été constaté qu'elles contiennent des informations utiles sur la dépendance résiduelle. Sans trop se tromper, on peut affirmer que les tests portemanteaux et leur utilisation font partie d'une analyse complète dans la formulation d'un modèle de séries chronologiques.

2.2. Processus ARMA et ARIMA

Un processus stationnaire au second ordre suit un modèle ARMA s'il est défini comme dans la définition 2.2.1.

Définition 2.2.1. *Un processus stationnaire au second ordre $\{X_t\}$ suit un modèle auto-régressif moyenne mobile d'ordres $p \geq 0$ et $q \geq 0$, noté ARMA(p, q), si:*

$$X_t - a_1 X_{t-1} - \dots - a_p X_{t-p} = \varepsilon_t - b_1 \varepsilon_{t-1} - \dots - b_q \varepsilon_{t-q}, \quad (2.2.1)$$

avec $a_p \neq 0$ et $b_q \neq 0$, et où le processus $\{\varepsilon_t\}$ est un bruit blanc de moyenne 0 et de variance σ_ε^2 . De plus, $\{X_t\}$ suit un modèle ARMA(p,q) de moyenne μ_X si le processus $\{X_t - \mu_X\}$ suit un modèle ARMA(p,q).

Nous définissons l'opérateur retard B tel que $BX_t = X_{t-1}$, $Bt = t - 1$ et $Bc = c$, où t est l'indice temporel et c une constante. En utilisant l'opérateur retard sur t , le modèle défini par (2.2.1) peut être réécrit de façon élégante comme suit:

$$\phi(B)X_t = \psi(B)\varepsilon_t, \quad (2.2.2)$$

où les polynômes $\phi(\cdot)$ et $\psi(\cdot)$ sont définis par:

$$\begin{aligned} \phi(z) &= 1 - a_1z - a_2z^2 - \dots - a_pz^p, \\ \psi(z) &= 1 - b_1z - b_2z^2 - \dots - b_qz^q. \end{aligned} \quad (2.2.3)$$

Afin d'éviter des redondances et autres problèmes techniques, une caractéristique dont il faut s'assurer dans l'écriture du processus $\{X_t\}$ comme un modèle ARMA(p,q) est l'absence de racines communes dans les polynômes $\phi(z)$ et $\psi(z)$. L'importance de cette hypothèse et une discussion portant sur ces considérations se trouvent dans Brockwell et Davis (1991). Sous des conditions sur les racines de $\phi(z)$ et $\psi(z)$, on peut obtenir la solution stationnaire aux équations définissant un processus ARMA. De plus, nous pouvons obtenir la représentation inversible. Ceci fait l'objet de la propriété suivante.

Propriété 2.2.1. *Soit $\{X_t\}$ un processus stationnaire au second ordre qui suit un modèle ARMA(p,q) satisfaisant l'équation stochastique (2.2.1). La représentation stationnaire (ou causale) est:*

$$X_t = \phi^{-1}(B)\psi(B)\varepsilon_t, \quad (2.2.4)$$

et la représentation inversible est:

$$\psi^{-1}(B)\phi(B)X_t = \varepsilon_t. \quad (2.2.5)$$

Les conditions d'existence de la représentation stationnaire et l'inversibilité reposent sur l'étude des racines de $\phi(\cdot)$ et $\psi(\cdot)$, respectivement. Elles doivent être toutes à l'extérieur du disque unitaire. Autrement dit, $\phi(z) = 0$ et $\psi(z) = 0$, pour $|z| > 1$, où z peut être un nombre complexe, et alors $|z|$ dénote son module.

Comme mentionné, une hypothèse importante est que $\phi(z)$ et $\psi(z)$ n'ont aucune racine commune. De plus, la propriété précédente stipule que les racines doivent toutes être à l'extérieur du disque unitaire. Ces conditions garantissent que les représentations stationnaire et inversible existent. Pour fixer les idées, prenons par exemple un modèle ARMA(1,1). Ainsi, nous supposons que le modèle est de la forme:

$$(1 - aB)X_t = (1 - bB)\varepsilon_t,$$

où $a \neq 0$ et $b \neq 0$.

Si $a = b$, alors nous pourrions avoir que $X_t = \varepsilon_t$, un bruit blanc, et en multipliant de chaque côté par $1 - aB = 1 - bB$, nous obtenons un ARMA(1,1) qui n'est pas identifiable. De manière générale, sous les conditions d'inversibilité et de stationnarité, si un modèle ARMA(p,q) est tel que $\phi(\cdot)$ et $\psi(\cdot)$ ont au moins une racine commune, alors le modèle peut être simplifié par un modèle ARMA(p',q'), avec $p' < p$ et $q' < q$. Ceci suggère l'importance de cette hypothèse.

Si $a \neq b$, et si $|a| < 1$ et $|b| < 1$, les représentations stationnaire (2.2.4) et inversible (2.2.5) sont telles que:

$$\begin{aligned} X_t &= (1 - aB)^{-1}(1 - bB)\varepsilon_t = \sum_{l \geq 0} a^l B^l (1 - bB)\varepsilon_t, \\ \varepsilon_t &= (1 - bB)^{-1}(1 - aB)X_t = \sum_{l \geq 0} b^l B^l (1 - aB)X_t, \end{aligned}$$

où nous faisons appel aux développements en séries géométriques de $(1 - az)^{-1}$ et $(1 - bz)^{-1}$. Ces séries de puissance en z existent lorsque $|a| < 1$ et $|b| < 1$, puisque les racines de $1 - az = 0$ et $1 - bz = 0$ sont $1/a$ et $1/b$, respectivement.

Plus généralement, un polynôme quelconque $\phi(z)$ d'ordre p peut toujours être écrit sous la forme:

$$\phi(z) = (1 - R_1 z) \cdots (1 - R_p z),$$

et les racines sont donc R_i^{-1} , $i = 1, \dots, p$. Ainsi, il est raisonnable que la condition de stationnarité repose sur des racines qui sont toutes à l'extérieur du disque unité, afin de trouver la représentation stationnaire. Ceci ne constitue pas une preuve cependant. Des développements plus formels se trouvent dans Shumway et Stoffer (2011, Annexe B.2).

Plusieurs séries chronologiques rencontrées en pratique ne sont pas stationnaires. En effet, une première cause courante de non-stationnarité est la présence d'une tendance. Dans un tel cas, la fonction moyenne $E(X_t)$ est exprimée en fonction du temps. Si la tendance est linéaire, on pourrait penser qu'un modèle possible pour $\{X_t\}$ est de la forme:

$$X_t = \beta_0 + \beta_1 t + Y_t, \quad (2.2.6)$$

où $\{Y_t\}$ serait cependant stationnaire. Une autre cause de non-stationnarité est l'hétéroscédasticité, lorsque $\text{Var}(X_t)$ est fonction du temps.

Dans les applications économiques et financières, une variable d'intérêt pourrait porter sur la valeur d'un portefeuille, ou encore plus spécifiquement sur la valeur d'un titre boursier ou d'une action. Les fluctuations dans ce type de séries chronologiques sont souvent compatibles avec la présence de tendances: en période de croissance, la valeur du titre peut augmenter, et souvent ce type de séries chronologiques affiche également une volatilité croissante, le concept de volatilité étant essentiellement le concept de variances conditionnelles.

On peut penser aux titres associés à des entreprises comme Apple, McDonald ou Nike, pour ne citer que quelques exemples. Ces actions sont populaires parmi les investisseurs dû à l'augmentation continue de leur valeur moyenne et les transactions se font quotidiennement à très grand volume. Ceci explique en partie l'augmentation de la volatilité avec le temps: plus les actions prennent de la valeur, plus elles deviennent populaires, les volumes de transactions augmentent et, par l'offre et la demande, les valeurs des actions varient fortement en fonction de l'augmentation du volume de transactions parce que les investisseurs suivent les tendances de transactions.

Lorsqu'une série chronologique n'est pas stationnaire, une stratégie possible consiste à la transformer afin de rendre l'hypothèse de stationnarité plausible. Par exemple, en présence de tendances, la différentiation définie par $\nabla X_t = X_t - X_{t-1} = (1 - B)X_t$ retire une tendance linéaire. Pour s'en convaincre, considérons la série (2.2.6). On aura alors que $\nabla X_t = \beta_1 + \nabla Y_t$. Si $\{Y_t\}$ est stationnaire, il en sera de même pour $\{\nabla Y_t\}$, et $\{\nabla X_t\}$ sera stationnaire. En présence de tendance quadratique, cubique ou plus généralement polynomiale d'ordre k , disons, différencier un nombre suffisant de fois permet de retirer la tendance polynomiale afin de rendre la série stationnaire.

Le concept de différenciation est à ce point important qu'il suggère une nouvelle classe de modèles ARMA, appelés modèles ARMA intégrés ou ARIMA: ils sont tels que la $d^{\text{ième}}$ différence respecte les conditions d'un modèle ARMA. Ces modèles font l'objet de la définition suivante.

Définition 2.2.2. *Un processus $\{X_t\}$ suit un modèle autorégressif intégré moyenne mobile avec ordres $p \geq 0$, $d > 0$ et $q \geq 0$, noté ARIMA(p,d,q) si:*

$$(1 - B)^d \phi(B) X_t = \psi(B) \varepsilon_t, \quad (2.2.7)$$

où le processus $\{\varepsilon_t\}$ est un bruit blanc de moyenne 0 et de variance σ_ε^2 .

Donc si nous posons $W_t = (1 - B)^d X_t$, pour tout t , alors le processus $\{W_t\}$ suit un modèle ARMA(p,q). Ainsi, lorsque la non-stationnarité est liée à la présence de tendances, la stratégie consiste dans un premier temps à différencier un assez grand nombre de fois afin d'obtenir une série stationnaire. Ensuite, nous faisons l'inférence sur la série transformée W_t , $t = 1, \dots, n$. On note que les prévisions de W_t dans la formulation ARMA doivent être retransformées afin de proposer des prévisions qui s'exprimeront en fonction des X_t , $t = 1, \dots, n$.

Pour le problème d'hétéroscédasticité, une stratégie peut consister à stabiliser la variance par une transformation logarithmique ou racine. Nous pouvons justifier de telles transformations à partir de l'expansion de Taylor de premier ordre. Soit un processus $\{X_t\}$. En pratique, il est commun que la variance au temps t soit une fonction de la moyenne du processus au temps t , donc il est raisonnable de supposer $\text{Var}(X_t) = cf(\mu_t) > 0$, où c est une constante positive et $f(\mu_t) > 0$, pour tout t . Soit une transformation $h(X_t)$ telle que sa variance $\text{Var}\{h(X_t)\}$ soit constante. Par une expansion de Taylor de premier ordre, nous avons $h(X_t) \approx h(\mu_t) + (X_t - \mu_t)h'(\mu_t)$. En prenant la variance de chaque côté de l'équation, nous obtenons $\text{Var}\{h(X_t)\} \approx \{h'(\mu_t)\}^2 \text{Var}(X_t)$ et en remplaçant $\text{Var}(X_t)$ par $cf(\mu_t)$, nous obtenons $\text{Var}\{h(X_t)\} \approx \{h'(\mu_t)\}^2 cf(\mu_t)$.

Donc, afin que la variance du processus transformé soit constante, il est suffisant que $h'(\mu_t) = \{f(\mu_t)\}^{-1/2}$. Si la variance du processus est proportionnel à la moyenne, alors $f(\mu_t) = \mu_t$ et $h(\mu_t) = \sqrt{\mu_t}$, $h(\cdot)$ est donc une transformation racine. Si l'écart-type du processus est proportionnel à la moyenne, alors $f(\mu_t) = \mu_t^2$ et $h(\mu_t) = \log(\mu_t)$, $h(\cdot)$ est donc une transformation logarithmique.

De manière plus générale, la méthodologie de Box-Cox peut être utilisée afin de trouver une transformation de la forme:

$$y_t^{(\lambda)} = \begin{cases} \frac{y_t^\lambda - 1}{\lambda}, & \lambda \neq 0, \\ \log(y_t), & \lambda = 0, \end{cases}$$

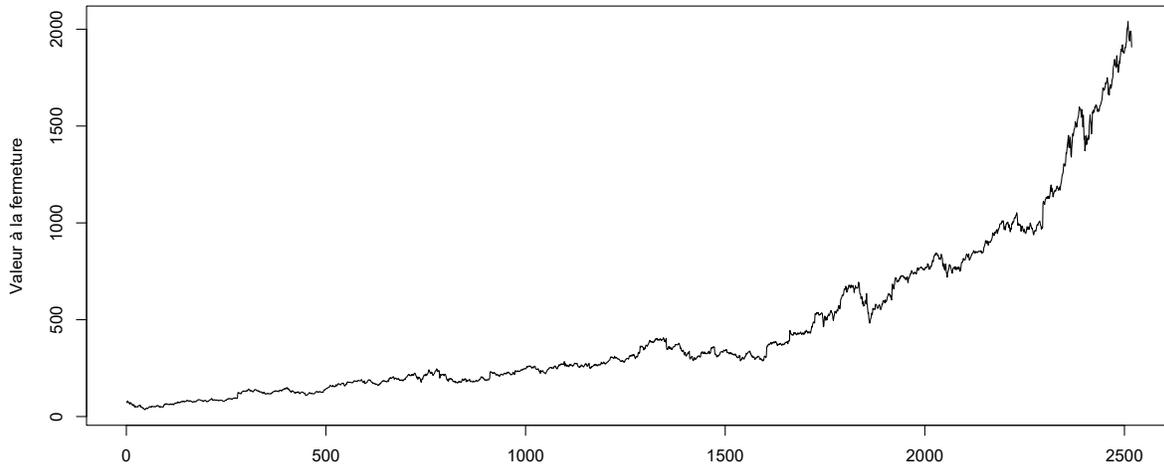
où $\{y_t\}$ est un processus tel que $y_t > 0$ pour tout t et λ est le paramètre de transformation qui peut être choisi en effectuant un graphique de la fonction de vraisemblance gaussienne en fonction de λ .

Comme exemple, considérons $\{X_t\}$ les données quotidiennes de fermeture des actions d'Amazon en date du 17 septembre 2008 au 17 septembre 2018, pour un total de 2518 données. Les données ont été téléchargées du site web *Yahoo! Finance* et sont accessibles à l'adresse:

<https://finance.yahoo.com/quote/AMZN/history?p=AMZN>.

Le graphique des données en fonction du temps est illustré à la figure 2.1. Nous constatons

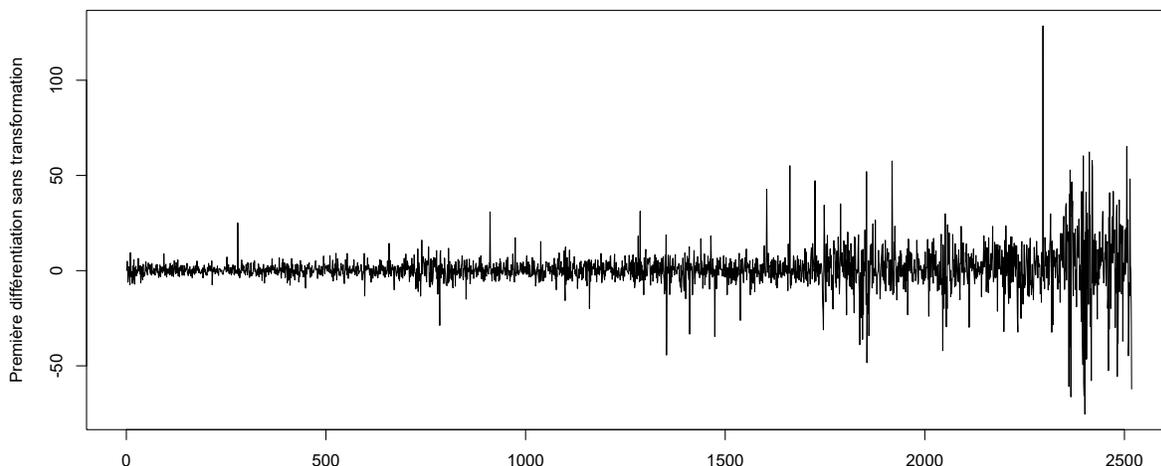
Figure 2.1. Données quotidiennes à la fermeture d'Amazon du 17 septembre 2008 jusqu'au 17 septembre 2018.



qu'il y a une tendance positive, de sorte que la moyenne augmente avec le temps. Soit $\{Y_t = X_t - X_{t-1}\}$, la série après une première différentiation. Cette série est représentée

par la figure 2.2. Nous obtenons un processus dont la tendance a été retirée et la moyenne ne semble plus varier avec le temps.

Figure 2.2. Données quotidiennes à la fermeture d’Amazon du 17 septembre 2008 jusqu’au 17 septembre 2018 après une première différentiation.

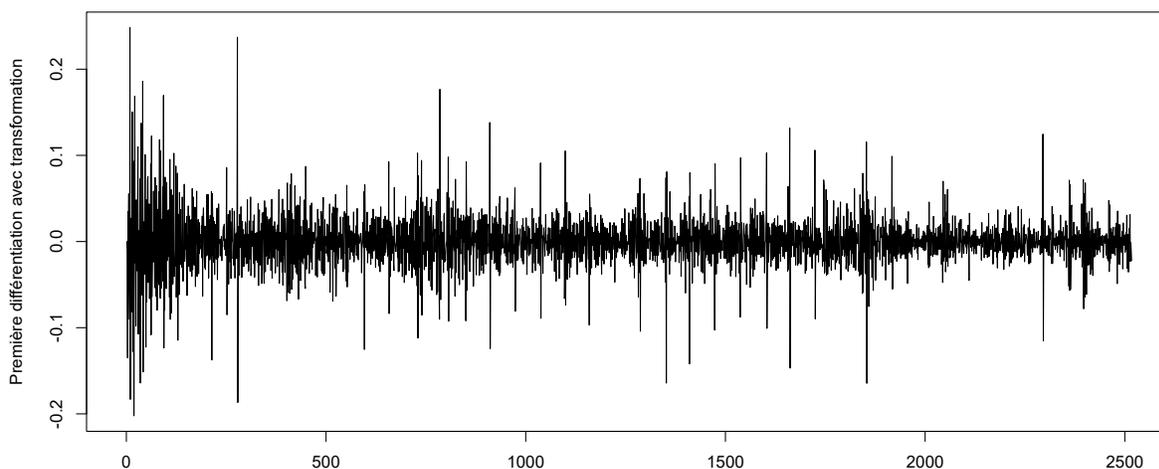


Cependant, nous constatons que la variabilité ne semble pas constante et, dans les dernières années, la variabilité est jusqu’à dix fois plus importante qu’au début. En effectuant d’abord une transformation logarithmique afin de stabiliser la variance et ensuite une première différentiation pour enlever la tendance, nous obtenons la série représentée dans la figure 2.3. La moyenne semble constante et la variabilité semble être stabilisée. L’inférence peut donc être effectuée sur la nouvelle série $\{Z_t = \log(X_t) - \log(X_{t-1})\}$ et après avoir ajusté un modèle sur celle-ci (par exemple de la forme $\phi(B)Z_t = \psi(B)\varepsilon_t$), nous pouvons retransformer le modèle afin d’obtenir une inférence sur la série originale (par exemple, de la forme $\phi(B)\{\log(X_t) - \log(X_{t-1})\} = \psi(B)\varepsilon_t$).

Remarque 2.2.1. Cette transformation est en fait le calcul du rendement. Si p_t dénote le prix au temps t , alors le rendement se calcule comme suit: $\log(p_t/p_{t-1}) = \log(p_t) - \log(p_{t-1}) \approx (p_t - p_{t-1})/p_{t-1}$.

En pratique, les données sont recueillies à des intervalles de temps réguliers, et puisque les événements économiques se répètent à des moments spécifiques à chaque année, des

Figure 2.3. Données quotidiennes à la fermeture d’Amazon du 17 septembre 2008 jusqu’au 17 septembre 2018 après une transformation logarithmique et une première différenciation.



effets saisonniers peuvent être observés et il faut en tenir compte dans l’inférence. Pour les données économiques mensuelles, des dépendances par multiple de 12 seraient présentes dû à la ressemblance des situations économiques selon chaque mois de l’année. Par exemple, novembre est associé à une augmentation des ventes de dindes dû à l’Action de Grâce et octobre est caractérisé par une augmentation des friandises dû à l’Halloween. Les phénomènes naturels comme l’hiver glacial au Québec ou les canicules en été sont aussi des exemples de fluctuations saisonnières dans les données trimestrielles.

En incluant un assez grand nombre de paramètres, les modèles ARIMA peuvent modéliser les séries chronologiques saisonnières. Cependant, ils ne sont typiquement pas très parcimonieux. Afin de tenir compte efficacement et avec parcimonie des effets saisonniers, il est possible de considérer des modèles ARIMA saisonniers purs, que l’on peut appeler modèles SARIMA. Ils sont des types particuliers de modèles ARIMA. Ils s’expriment en fonction d’opérateurs autorégressifs et moyennes-mobiles saisonniers, s’exprimant comme des polynômes en B^s , où s est un paramètre supposé connu lié à la période saisonnière. Par exemple, dans le cas de données mensuelles, il est habituellement postulé $s = 12$, alors

que des données trimestrielles exigeraient typiquement que $s = 4$. Ces modèles ARIMA saisonniers purs font l'objet de la définition suivante.

Définition 2.2.3. *Un processus stationnaire au second ordre $\{X_t\}$ suit un modèle ARIMA saisonnier pur d'ordres $P \geq 0$, $D \geq 0$, $Q \geq 0$ et de période saisonnière $s \geq 1$, noté $\text{ARIMA}(P,D,Q)_s$, si:*

$$\Phi(B^s)(1 - B^s)^D X_t = \Psi(B^s)\varepsilon_t, \quad (2.2.8)$$

où le processus $\{\varepsilon_t\}$ est un bruit blanc de moyenne 0 et de variance σ_ε^2 .

Les polynômes dans (2.2.8) sont:

$$\begin{aligned} \Phi(B^s) &= 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps}, \\ \Psi(B^s) &= 1 - \Psi_1 B^s - \Psi_2 B^{2s} - \dots - \Psi_Q B^{Qs}. \end{aligned} \quad (2.2.9)$$

Les conditions de stationnarité et d'inversibilité sont les mêmes que dans le cas non-saisonnier, c'est-à-dire, les racines des polynômes $\Phi(B^s)$ et $\Psi(B^s)$ doivent être à l'extérieur du disque unité. Les mêmes arguments et preuves utilisés précédemment peuvent être utilisés pour cette situation également.

La différentiation saisonnière retire la non-stationnarité saisonnière. Par exemple, pour le mois de janvier, la consommation moyenne en électricité se ressemblerait fortement d'année en année, celle du mois de février se ressemblerait d'année en année, et ainsi de suite. Nous pouvons modéliser cette situation de la manière suivante:

$$X_t = X_{t-12} + Y_t,$$

où $\{Y_t\}$ représente l'évolution de la consommation moyenne en électricité sans tenir compte des composantes saisonnières. Par une différentiation, nous avons que:

$$(1 - B^{12})X_t = X_t - X_{t-12} = Y_t,$$

et il ne nous reste que $\{Y_t\}$. En pratique, la première différentiation est suffisante.

Dans les séries chronologiques réelles, il est peu plausible que toute la dépendance soit expliquée par des effets saisonniers. En effet, il est vraisemblable qu'en plus d'effets saisonniers, une dépendance locale expliquée par des autocorrélations non-nulles de bas ordre soit également présente. Ceci nous amène à considérer des modèles multiplicatifs

saisonniers. Ils incluent des opérateurs autorégressifs et moyennes-mobiles réguliers, permettant de modéliser une dépendance locale de bas ordre. Les modèles SARIMA, notés $\text{ARIMA}(p,d,q) \times (P,D,Q)_s$ sont introduits dans la définition suivante.

Définition 2.2.4. *Un processus stationnaire au second ordre $\{X_t\}$ suit un modèle ARIMA saisonnier multiplicatif, noté $\text{ARIMA}(p,d,q) \times (P,D,Q)_s$, si:*

$$\Phi(B^s)\phi(B)(1-B)^d(1-B^s)^D X_t = \Psi(B^s)\psi(B)\varepsilon_t, \quad (2.2.10)$$

où le processus $\{\varepsilon_t\}$ est un bruit blanc de moyenne 0 et de variance σ_ε^2 .

Les polynômes de la définition 2.2.4 ont été présentés précédemment, c'est-à-dire, les polynômes des ARIMA non-saisonniers $\phi(B)$ et $\psi(B)$ sont les mêmes que dans la définition 2.2.2 et les polynômes des ARIMA saisonniers purs $\Phi(B^s)$ et $\Psi(B^s)$ sont les mêmes que dans la définition 2.2.3.

Un exemple avec les données de l'indice de production de la Réserve fédérale américaine se trouve dans Shumway et Stoffer (2011, pp. 159–162). Les auteurs illustrent aussi les comportements de la fonction d'autocorrélation et de la fonction d'autocorrélation partielle, et ce à chaque étape du processus d'inférence. Leurs résultats peuvent être reproduits en utilisant le code R fourni et en obtenant les données de la librairie R se nommant `astsa`.

2.3. Processus ARMA semi-fort

Dans la définition 2.2.1, nous avons défini un processus $\{X_t\}$ qui était généré selon des équations menant aux modèles ARMA. Un examen de la définition révèle que le terme de bruit blanc, que l'on avait noté $\{\varepsilon_t\}$, est constitué de variables aléatoires non-corrélées. Il est possible d'être un peu plus précis sur les hypothèses sur le bruit blanc. Ceci nous amène à la définition d'une différence de martingales.

Définition 2.3.1. *Un processus $\{\varepsilon_t\}$ est dit être une différence de martingales si $E|\varepsilon_t| < \infty$ et si:*

$$E(\varepsilon_t | \mathcal{F}_{t-1}) = 0,$$

où \mathcal{F}_t désigne l'information disponible jusqu'au moment t inclusivement.

Il est sous-entendu dans la définition 2.3.1 que les \mathcal{F}_t sont des σ -algèbres satisfaisant $\mathcal{F}_t \subset \mathcal{F}_{t-1}$. De plus, la variable aléatoire ε_t est \mathcal{F}_t -mesurable. Une définition plus rigoureuse se trouve dans Francq et Zakoïan (2009, Annexe A). Selon ces derniers, quand un processus $\{\varepsilon_t\}$ satisfait la définition 2.3.1, il est appelé bruit semi-fort. Ainsi, lorsque $\{\varepsilon_t\}$ est iid, de moyenne nulle, ce processus satisfait facilement la définition 2.3.1, puisque par indépendance $E(\varepsilon_t|\mathcal{F}_{t-1}) = E(\varepsilon_t) = 0$, où $\mathcal{F}_t = \sigma(\varepsilon_t, \varepsilon_{t-1}, \dots)$ désigne la σ -algèbre générée par la suite de variables aléatoires $\{\varepsilon_t, \varepsilon_{t-1}, \dots\}$. À partir de $\{\varepsilon_t\}$ supposé iid, on peut construire aisément des différences de martingale, en considérant par exemple $u_t = \varepsilon_t \varepsilon_{t-1}$. En effet, par le théorème de la double espérance, $E(u_t) = E(\varepsilon_t \varepsilon_{t-1}) = E\{E(\varepsilon_t \varepsilon_{t-1}|\mathcal{F}_{t-1})\}$ et de plus:

$$\begin{aligned} E(u_t|\mathcal{F}_{t-1}) &= E(\varepsilon_t \varepsilon_{t-1}|\mathcal{F}_{t-1}), \\ &= \varepsilon_{t-1} E(\varepsilon_t|\mathcal{F}_{t-1}), \\ &= \varepsilon_{t-1} E(\varepsilon_t), \\ &= 0. \end{aligned}$$

On verra plus loin dans ce chapitre que les modèles autorégressifs conditionnellement hétéroscédastiques (ARCH) sont également des différences de martingale.

On remarque qu'un processus différence de martingales est également un bruit blanc si on suppose qu'il est de variance finie. En effet, si $\{\varepsilon_t\}$ satisfait la définition 2.3.1 et si $E(\varepsilon_t^2) = \sigma_\varepsilon^2 < \infty$, on aura pour $k > 0$ que $E(\varepsilon_t) = E\{E(\varepsilon_t|\mathcal{F}_{t-1})\} = 0$ et:

$$\begin{aligned} \text{Cov}(\varepsilon_t, \varepsilon_{t-k}) &= E(\varepsilon_t \varepsilon_{t-k}) - E(\varepsilon_t)E(\varepsilon_{t-k}), \\ &= E\{E(\varepsilon_t \varepsilon_{t-k}|\mathcal{F}_{t-1})\}, \\ &= E\{\varepsilon_{t-k} E(\varepsilon_t|\mathcal{F}_{t-1})\}, \\ &= E(\varepsilon_{t-k} \cdot 0), \\ &= 0. \end{aligned}$$

Il ressort de cette discussion qu'une différence de martingales est un bruit blanc particulier. De ce point de vue, il est quelque peu naturel de considérer des modèles avec ce bruit comme terme d'erreur, qui s'avère plus général que de présumer l'indépendance, mais plus restrictif que de présumer seulement l'absence de corrélation.

Dans les développements qui suivent, nous devons présumer que $\{\varepsilon_t\}$ est une différence de martingales particulière, satisfaisant des conditions techniques permettant une analyse asymptotique.

Hypothèse A. Soit $\{\varepsilon_t\}$ une différence de martingales satisfaisant la définition 2.3.1 avec $\mathcal{F}_t = \sigma\{\varepsilon_t, \varepsilon_{t-1}, \dots\}$ désignant la σ -algèbre générée par l'ensemble de variables aléatoires $\{\varepsilon_t, \varepsilon_{t-1}, \dots\}$, et de variance finie. Ainsi:

$$(i) \ E(\varepsilon_t | \mathcal{F}_{t-1}) = 0 \text{ presque sûrement pour tout } t;$$

$$(ii) \ E(\varepsilon_t^2) = \sigma_\varepsilon^2 := \sigma^2 < \infty, \text{ pour tout } t;$$

De plus, des hypothèses additionnelles sont supposées:

$$(iii) \ \lim_{n \rightarrow \infty} n^{-1} \sum_{t=1}^n E(\varepsilon_t^2 | \mathcal{F}_{t-1}) = \sigma_\varepsilon^2 > 0 \text{ presque sûrement};$$

$$(iv) \ \text{Il existe une variable aléatoire } U \text{ avec } E(U^2) < \infty \text{ telle que } P(|\varepsilon_t| > u) \leq cP(|U| > u), \text{ pour une constante } 0 < c < \infty \text{ et tout } t, \text{ tout } u \geq 0;$$

$$(v) \ E(\varepsilon_t^2 \varepsilon_{t-r} \varepsilon_{t-s}) = \sigma_\varepsilon^4 \tau_{rs} \text{ est finie et uniformément bornée pour tout } t, r \geq 1, s \geq 1;$$

$$(vi) \ \lim_{n \rightarrow \infty} n^{-1} \sum_{t=1}^n \varepsilon_{t-r} \varepsilon_{t-s} E(\varepsilon_t^2 | \mathcal{F}_{t-1}) = \sigma_\varepsilon^4 \tau_{rs} \text{ presque sûrement pour tout } r \geq 1, s \geq 1.$$

L'hypothèse A a déjà été considérée dans la littérature. Premièrement, on note que (i) et (ii) ne sont que la présupposition que $\{\varepsilon_t\}$ est une différence de martingales avec une variance supposée finie. Les conditions (iii) et (iv) se trouvent dans Hannan et Heyde (1972, p. 2059) qui utilisent ces conditions afin d'obtenir des propriétés de convergence presque sûre pour les autocorrélations échantillonnales. Les conditions (v) et (vi) s'ajoutent afin d'obtenir un théorème limite central pour les autocorrélations (voir le théorème 2 de Hannan et Heyde (1972, p. 2061)). Ces conditions ont été supposées dans Deo (2000, p. 293). Cependant, puisque Deo (2000) développe des tests spectraux pour l'hypothèse de martingale en présence d'hétéroscédasticité conditionnelle, il doit supposer des conditions additionnelles, comme l'existence d'un huitième moment pour ε_t , ce que nous n'avons pas besoin dans notre contexte. On note que l'hypothèse (iii) permet l'hétéroscédasticité conditionnelle. Si nous comparons avec les hypothèses qui se trouvent dans Escanciano et Lobato (2009), malgré que nous ayons besoin de conditions additionnelles sur les moments, notons que nous n'avons pas besoin de supposer la stationnarité au sens strict pour le processus sous-jacent.

Dans le mémoire, nous nous intéressons à la qualité d'ajustement des modèles ARMA avec des innovations $\{\varepsilon_t\}$ satisfaisant une condition de différence de martingales. Plus précisément, les hypothèses nulle et alternative sont formulées comme suit:

H_0 : Le processus $\{X_t\}$ satisfait un modèle ARMA(p, q) comme dans l'équation (2.2.1) et $\{\varepsilon_t\}$ satisfait l'hypothèse A, (2.3.1)

H_1 : Le processus n'admet pas une représentation ARMA, ou il admet une représentation ARMA(p', q') telle que $p' > p$ ou $q' > q$.

Ces hypothèses se trouvent dans Francq et al. (2005). Une façon plus précise de formuler H_0 concerne l'existence des polynômes autorégressifs et moyennes mobiles, d'ordres p et q , respectivement, avec racines à l'extérieur du disque unité et sans racines communes, et de sorte que $\psi^{-1}(B)\phi(B)X_t$ soit de moyenne nulle et que le processus $\{\psi^{-1}(B)\phi(B)X_t\}$ définisse une suite de variables aléatoires non-corrélées.

Sous l'hypothèse alternative H_1 , il est sous-entendu qu'une représentation ARMA comme sous H_0 n'existe pas, à moins qu'un modèle ARMA existe avec des ordres autorégressifs ou moyennes mobiles plus grands.

Ainsi, la qualité de l'ajustement se fera par une étude des résidus. Plus précisément, il sera vérifié que les résidus du modèle ARMA(p, q) seront compatibles avec l'hypothèse de bruit blanc semi-fort. Sous H_0 , les résidus présenteraient des corrélations faibles, de sorte que les tests portemanteaux prendraient de petites valeurs. En revanche, si le modèle ARMA(p, q) spécifié ne permet pas de décrire toute la dépendance, les ordres p et q ne seraient pas assez grands, et les résidus du modèle ARMA ajusté devraient afficher de la corrélation. Les tests portemanteaux devraient avoir également de la puissance pour les alternatives qui ne respectent pas l'hypothèse A, à savoir un bruit blanc qui n'est pas semi-fort. Ainsi, sous ces alternatives, les tests portemanteaux prendront de grandes valeurs et devraient entraîner un rejet de H_0 .

On remarque que la distribution limite des autocovariances résiduelles est obtenue sous des hypothèses précises et un écart de ce contexte devrait se refléter dans les statistiques de test. Bien entendu, comme mentionné dans Francq et al. (2005), les tests portemanteaux reposent sur un nombre fini de délais et par conséquent ils ne peuvent détecter tous les

écarts à l'hypothèse nulle H_0 . Aussi, en pratique, l'utilisateur doit calculer généralement les tests pour plusieurs valeurs de m .

2.4. Distribution limite dans les modèles ARMA

Nous débutons cette section en étudiant la fonction d'autocovariance empirique du bruit inobservable $\{\varepsilon_t\}$. Plus précisément, on considère

$$\hat{c}_\varepsilon(l) = \begin{cases} n^{-1} \sum_{t=1}^{n-l} \varepsilon_t \varepsilon_{t+l}, & l \geq 0, \\ \hat{c}_\varepsilon(-l), & l < 0. \end{cases}$$

Afin de simplifier les notations, on pose $r(l) = \hat{c}_\varepsilon(l)$, $l \in \mathbb{Z}$, et on introduit le vecteur aléatoire suivant:

$$\mathbf{r}_m = (r(1), r(2), \dots, r(m))^\top, \quad m \geq 1.$$

L'entier m est supposé connu et fixé par l'expérimentateur. On notera dans ce qui suit V^\top pour la transposée de la matrice ou du vecteur V .

Le prochain lemme sera utile comme résultat intermédiaire dans les développements qui suivent. En invoquant des arguments similaires à ceux dans la preuve du théorème 2 de Hannan et Heyde (1972, pp. 2062–2063), on obtient la distribution asymptotique des autocovariances empiriques.

Lemme 2.4.1. *Sous l'hypothèse A, le vecteur d'autocovariances empiriques du bruit \mathbf{r}_m satisfait, pour tout $m \geq 1$ fixé:*

$$\sqrt{n} \mathbf{r}_m \longrightarrow_d \mathcal{N}_m(\mathbf{0}_m, \mathbf{\Gamma}_{m,m}), \quad (2.4.1)$$

où $\mathbf{\Gamma}_{m,m'} = (\Gamma(l, l'))_{1 \leq l \leq m, 1 \leq l' \leq m'}$ avec $\Gamma(l, l') = \sigma^4 \tau_{ll'}$ pour $l, l' \geq 1$ et $\tau_{ll'}$ est fourni dans l'hypothèse A(v).

Dans le lemme 2.4.1, $\mathbf{0}_m$ dénote un vecteur de taille m , dont tous les éléments sont des zéros. Puisque l'on présume que la propriété 2.2.1 est remplie, nous exprimons $\phi^{-1}(z)$ et $\psi^{-1}(z)$ comme suit:

$$\begin{aligned} \phi^{-1}(z) &= \sum_{i=0}^{\infty} \phi_i^* z^i, \\ \psi^{-1}(z) &= \sum_{i=0}^{\infty} \psi_i^* z^i. \end{aligned} \quad (2.4.2)$$

Les expansions en (2.4.2) sont définies pour $|z| \leq 1$, voir Francq et al. (2005). Il sera commode d'avoir des coefficients ϕ_i^* et ψ_i^* définis pour tous les indices $i \in \mathbb{Z}$. Ainsi, nous posons $\phi_i^* = \psi_i^* = 0$, pour $i < 0$.

Définissons les vecteurs de $(p + q)$ éléments:

$$\boldsymbol{\lambda}_i = (-\phi_{i-1}^*, -\phi_{i-2}^*, \dots, -\phi_{i-p}^*, \psi_{i-1}^*, \psi_{i-2}^*, \dots, \psi_{i-q}^*)^\top, \quad (2.4.3)$$

et la matrice $(p + q) \times m$:

$$\boldsymbol{\Lambda}_m = (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \dots, \boldsymbol{\lambda}_m) = \begin{pmatrix} -1 & -\phi_1^* & \cdots & -\phi_{m-1}^* \\ 0 & -1 & \cdots & -\phi_{m-2}^* \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\phi_{m-p}^* \\ 1 & \psi_1^* & \cdots & \psi_{m-1}^* \\ 0 & 1 & \cdots & \psi_{m-2}^* \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \psi_{m-q}^* \end{pmatrix}. \quad (2.4.4)$$

À partir de $\boldsymbol{\lambda}_i$ et $\boldsymbol{\Lambda}_m$ définis par (2.4.3) et (2.4.4), nous introduisons les matrices suivantes:

$$\begin{aligned} \boldsymbol{\Lambda}_\infty \boldsymbol{\Lambda}_\infty^\top &= \sum_{i=1}^{\infty} \boldsymbol{\lambda}_i \boldsymbol{\lambda}_i^\top, \\ \boldsymbol{\Lambda}_\infty \boldsymbol{\Gamma}_{\infty, \infty} \boldsymbol{\Lambda}_\infty^\top &= \sum_{l=1}^{\infty} \sum_{l'=1}^{\infty} \boldsymbol{\lambda}_l \Gamma(l, l') \boldsymbol{\lambda}_{l'}^\top, \\ \boldsymbol{\Lambda}_\infty \boldsymbol{\Gamma}_{\infty, m} &= \left(\sum_{i=1}^{\infty} \Gamma(i, 1) \boldsymbol{\lambda}_i, \dots, \sum_{i=1}^{\infty} \Gamma(i, m) \boldsymbol{\lambda}_i \right), \end{aligned}$$

où $\Gamma(l, l')$ est défini dans le lemme 2.4.1. On note que $\boldsymbol{\Lambda}_\infty \boldsymbol{\Lambda}_\infty^\top$ et $\boldsymbol{\Lambda}_\infty \boldsymbol{\Gamma}_{\infty, \infty} \boldsymbol{\Lambda}_\infty^\top$ sont de dimensions $(p + q) \times (p + q)$, alors que $\boldsymbol{\Lambda}_\infty \boldsymbol{\Gamma}_{\infty, m}$ est de dimension $(p + q) \times m$.

Les paramètres du modèle (2.2.1) doivent être estimés. Une technique d'estimation possible est la méthode des moindres carrés. Quelques notations s'imposent. Dans ce qui suit, le cadre de travail est similaire à ce qui se trouve dans Francq et al. (2005). On rappelle que si $\{X_t\}$ est comme dans la définition 2.2.1, on aura que $a_p \neq 0$ et $b_q \neq 0$, et adoptons la convention que $a_0 = b_0 = 1$.

Soit $\boldsymbol{\theta}_0 = (a_1, a_2, \dots, a_p, b_1, b_2, \dots, b_q)^\top$, $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_p, \theta_{p+1}, \dots, \theta_{p+q})^\top$ et dénotons par $\Theta = \{\boldsymbol{\theta} \in \mathbb{R}^{p+q}\}$ l'espace de paramètres. Pour tout $\boldsymbol{\theta} \in \Theta$, il sera commode d'introduire le processus suivant:

$$\varepsilon_t(\boldsymbol{\theta}) = \psi^{-1}(B)\phi(B)X_t.$$

Le processus $\{\varepsilon_t(\boldsymbol{\theta})\}$ définit un processus résiduel et nous notons que $\varepsilon_t(\boldsymbol{\theta}_0) = \varepsilon_t$, presque sûrement. Étant donné les observations X_1, X_2, \dots, X_n , le processus résiduel $\{\varepsilon_t(\boldsymbol{\theta})\}$ peut être approximé par $e_t(\boldsymbol{\theta})$, défini par:

$$e_t(\boldsymbol{\theta}) = X_t - \sum_{i=1}^p \theta_i X_{t-i} + \sum_{i=1}^q \theta_{p+i} e_{t-i}(\boldsymbol{\theta}), \quad (2.4.5)$$

où les valeurs initiales satisfont $e_0(\boldsymbol{\theta}) = e_{-1}(\boldsymbol{\theta}) = \dots = e_{-q+1}(\boldsymbol{\theta}) = X_0 = X_{-1} = \dots = X_{-p+1} = 0$.

Soit Θ^* un sous-ensemble compact de Θ tel que $\boldsymbol{\theta}_0$ est à l'intérieur de Θ^* . L'estimateur des moindres carrés de $\hat{\boldsymbol{\theta}}$ est défini par:

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta} \in \Theta^*}{\operatorname{argmin}} S_n(\boldsymbol{\theta}), \quad (2.4.6)$$

où $S_n(\boldsymbol{\theta}) = n^{-1} \sum_{t=1}^n e_t^2(\boldsymbol{\theta})$. L'hypothèse que $\boldsymbol{\theta} \in \Theta^*$ dans (2.4.6) est une hypothèse technique sous laquelle l'estimateur des moindres carrés est convergent et asymptotiquement normal sous des conditions générales. Voir par exemple Francq et Zakoïan (1998).

Le prochain lemme donne une représentation de l'estimateur des moindres carrés.

Lemme 2.4.2. *Supposons que $\{X_t\}$ suit un modèle ARMA (2.2.1) et que l'innovation $\{\varepsilon_t\}$ satisfait l'hypothèse A, alors l'estimateur $\hat{\boldsymbol{\theta}}$ défini dans (2.4.6) satisfait:*

$$\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 = \mathbf{J}^{-1} \mathbf{Y}_n + \mathbf{O}_p(1/n), \quad (2.4.7)$$

où $\mathbf{J} = \sigma^2 \sum_{i=1}^{\infty} \boldsymbol{\lambda}_i \boldsymbol{\lambda}_i^\top = \sigma^2 \boldsymbol{\Lambda}_\infty \boldsymbol{\Lambda}_\infty^\top$ et $\mathbf{Y}_n = -n^{-1} \sum_{t=1}^n \sum_{i \geq 1} \varepsilon_t \varepsilon_{t-i} \boldsymbol{\lambda}_i$.

Dans l'expression (2.4.7), $\mathbf{O}_p(a_n)$ signifie bornée en probabilité. Voir Serfling (1980) pour une définition précise.

De plus, nous avons la distribution conjointe pour $\hat{\boldsymbol{\theta}}$ et le vecteur d'autocovariances empiriques du bruit \mathbf{r}_m . Le résultat fait l'objet du lemme 2.4.3.

Lemme 2.4.3. *Supposons que $\{X_t\}$ suit un modèle ARMA (2.2.1) et que l'innovation $\{\varepsilon_t\}$ satisfait l'hypothèse A, alors l'estimateur $\hat{\boldsymbol{\theta}}$ et le vecteur d'autocovariances empiriques du bruit \mathbf{r}_m satisfont:*

$$\sqrt{n} \left((\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^\top, \mathbf{r}_m^\top \right)^\top \longrightarrow_d \mathcal{N}_{p+q+m}(\mathbf{0}_{p+q+m}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}, \mathbf{r}_m}), \quad (2.4.8)$$

où la matrice des variances et des covariances est:

$$\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}, \mathbf{r}_m} = \begin{pmatrix} \mathbf{J}^{-1} \boldsymbol{\Lambda}_\infty \boldsymbol{\Gamma}_{\infty, \infty} \boldsymbol{\Lambda}_\infty^\top \mathbf{J}^{-1} & -\mathbf{J}^{-1} \boldsymbol{\Lambda}_\infty \boldsymbol{\Gamma}_{\infty, m} \\ -\boldsymbol{\Gamma}_{m, \infty} \boldsymbol{\Lambda}_\infty^\top \mathbf{J}^{-1} & \boldsymbol{\Gamma}_{m, m} \end{pmatrix}. \quad (2.4.9)$$

Pour les résidus, soit $\hat{\varepsilon}_t = e_t(\hat{\boldsymbol{\theta}})$ quand $p > 0$ ou $q > 0$, et soit $\hat{\varepsilon}_t = e_t = X_t$ quand $p = q = 0$. Les autocovariances résiduelles sont définies par la fonction $\hat{r}(l) = \hat{c}_\varepsilon(l) = n^{-1} \sum_{t=1}^{n-l} \hat{\varepsilon}_t \hat{\varepsilon}_{t+l}$, pour tout $l \geq 0$. Soit le vecteur $\hat{\mathbf{r}}_m = (\hat{r}(1), \hat{r}(2), \dots, \hat{r}(m))^\top$, pour tout $m \geq 1$. Le prochain lemme contient le résultat le plus important de cette section. Il donne la distribution asymptotique des autocovariances résiduelles sous l'hypothèse A. Ceci nous permettra de développer plus tard nos tests.

Lemme 2.4.4. *Supposons que $\{X_t\}$ suit un modèle ARMA (2.2.1) et que l'innovation $\{\varepsilon_t\}$ satisfait l'hypothèse A, alors le vecteur d'autocovariances résiduelles $\hat{\mathbf{r}}_m$ satisfait, pour tout $m \geq 1$ fixé:*

$$\sqrt{n} \hat{\mathbf{r}}_m \longrightarrow_d \mathcal{N}_m(\mathbf{0}_m, \boldsymbol{\Sigma}_{\hat{\mathbf{r}}_m}), \quad (2.4.10)$$

où $\boldsymbol{\Sigma}_{\hat{\mathbf{r}}_m} = \boldsymbol{\Gamma}_{m, m}$ quand $p = q = 0$, et quand $p > 0$ ou $q > 0$:

$$\begin{aligned} \boldsymbol{\Sigma}_{\hat{\mathbf{r}}_m} = & \boldsymbol{\Gamma}_{m, m} + \boldsymbol{\Lambda}_m^\top (\boldsymbol{\Lambda}_\infty \boldsymbol{\Lambda}_\infty^\top)^{-1} \boldsymbol{\Lambda}_\infty \boldsymbol{\Gamma}_{\infty, \infty} \boldsymbol{\Lambda}_\infty^\top (\boldsymbol{\Lambda}_\infty \boldsymbol{\Lambda}_\infty^\top)^{-1} \boldsymbol{\Lambda}_m \\ & - \boldsymbol{\Lambda}_m^\top (\boldsymbol{\Lambda}_\infty \boldsymbol{\Lambda}_\infty^\top)^{-1} \boldsymbol{\Lambda}_\infty \boldsymbol{\Gamma}_{\infty, m} - \boldsymbol{\Gamma}_{m, \infty} \boldsymbol{\Lambda}_\infty^\top (\boldsymbol{\Lambda}_\infty \boldsymbol{\Lambda}_\infty^\top)^{-1} \boldsymbol{\Lambda}_m. \end{aligned} \quad (2.4.11)$$

Les preuves des lemmes 2.4.2, 2.4.3 et 2.4.4 sont fournies dans l'annexe. Les tests proposés sont basés sur les méthodes d'ondelettes, plus particulièrement, sur les coefficients d'ondelettes empiriques de l'expansion d'ondelette de la densité spectrale pour l'innovation. Nous introduisons les méthodes d'ondelettes dans la section suivante.

2.5. Introduction aux ondelettes

La théorie des ondelettes est la partie des mathématiques qui étudie comment une fonction peut s'exprimer dans une base particulière de fonctions. Comme théorie alternative à l'analyse de Fourier, c'est une branche des mathématiques qui est active en soi. Les applications en statistiques sont apparues plus tard.

Rappelons que l'analyse de Fourier porte sur la décomposition d'une fonction continue et périodique en une base de fonctions trigonométriques et les coefficients de la représentation sont appelés coefficients de Fourier. Plus précisément, les séries trigonométriques de Fourier étudient l'approximation d'une fonction f continue et périodique par des combinaisons linéaires de la forme:

$$S_N(x) = \frac{a_0}{2} + \sum_{j=1}^N \{a_j \cos(jwx) + b_j \sin(jwx)\},$$

où $\Delta = 2\pi/w$ est appelée la période. En considérant a_0 , a_j et b_j , $j = 1, \dots, N$, les coefficients dits de Fourier, il peut être montré que $S_N(x)$ converge vers $f(x)$ lorsque $N \rightarrow \infty$.

Les bases d'ondelettes servent également dans l'approximation des fonctions. Il semble que la première base d'ondelettes soit attribuée au mathématicien Alfred Haar. Essentiellement, l'approximation proposée par Haar repose sur des fonctions étagées dont les valeurs sont des moyennes de la fonction sur des intervalles dyadiques appropriés.

Plus généralement, une ondelette est une fonction $\eta \in L_2(\mathbb{R})$, où $L_2(\mathbb{R})$ dénote l'espace des fonctions de carré intégrable sur \mathbb{R} , qui est telle que les translations dyadiques de η , $\eta_{jk}(x) = 2^{j/2}\eta(2^jx - k)$, $j, k \in \mathbb{Z}$, constituent une base orthonormale de $L_2(\mathbb{R})$. Intuitivement, l'entier k représente un paramètre de translation, alors que j quant à lui correspond à la résolution. On peut consulter Vidakovic (1999) qui présente l'historique du développement des techniques d'ondelettes.

La présentation des ondelettes qui suit est inspirée de Lee et Hong (2001). On note que toutes les fonctions $f(x) \in L_2(\mathbb{R})$ peuvent être exprimées par une base d'ondelettes. Lorsqu'un processus stochastique $\{X_t\}$ est stationnaire au second ordre, de fonction d'autocovariances $\{c_X(l)\}$ absolument sommable, c'est-à-dire lorsque $\sum_{l=-\infty}^{\infty} |c_X(l)| < \infty$, on

peut définir la densité spectrale associée à $\{X_t\}$:

$$f_X(w) = (2\pi)^{-1} \sum_{h=-\infty}^{\infty} c_X(h) e^{-ihw}, \quad w \in [-\pi, \pi],$$

voir Fuller (1996). En utilisant la relation $e^{-ihw} = \cos(hw) - \mathbf{i} \sin(hw)$, où $\mathbf{i} = \sqrt{-1}$ dénote le nombre imaginaire, et en utilisant le fait que $c_X(l) = c_X(-l)$, on montre que $f_X(w)$ est une fonction réelle telle que:

$$f_X(w) = \frac{c_X(0)}{2\pi} + \frac{1}{\pi} \sum_{h=1}^{\infty} c_X(h) \cos(hw).$$

On se propose d'utiliser la théorie des ondelettes afin d'obtenir une représentation alternative de la densité spectrale.

Une approche est l'analyse multirésolution, qui repose sur l'ondelette mère η et l'ondelette père ξ . L'ondelette mère est elle-même fonction de l'ondelette père qui joue le rôle d'une fonction d'échelle.

Une motivation pour l'utilisation des ondelettes est la capacité d'adaptation des ondelettes afin de décrire des résolutions fines avec peu de coefficients. L'analyse multirésolution est une approche qui semble dominante. On présente la définition suivante.

Définition 2.5.1. *La multirésolution est une séquence de sous-espaces $\{V_j, j \in \mathbb{Z}\}$ de $L_2(\mathbb{R})$ satisfaisant les conditions suivantes:*

- (i) $V_j \subset V_{j+1}$ et $\cap V_j = \{0\}$, $\overline{\cap V_j} = L_2(\mathbb{R})$,
- (ii) $f(x) \in V_j$ si et seulement si $f(2x) \in V_{j+1}$,
- (iii) $f(x) \in V_j$ si et seulement si $f(x - k) \in V_j$, pour tout $k \in \mathbb{Z}$,
- (iv) V_0 possède une base orthonormale $\{\xi(\cdot - k), k \in \mathbb{Z}\}$, où l'ondelette père $\xi : \mathbb{R} \rightarrow \mathbb{R}$ est telle que $\int_{-\infty}^{\infty} \xi(x) dx = 1$.

De la structure imbriquée de sous-espaces V_j , le complément orthogonal W_j de V_j peut être défini par $V_j \oplus W_j = V_{j+1}$, où \oplus est la somme orthogonale. Soit V_0 le sous-espace initial. Alors:

$$V_0 \oplus \sum_{j=0}^{\infty} W_j = L_2(\mathbb{R}).$$

Par la dilatation et la translation, les séquences $\{\xi_{jk}(x) = 2^{j/2} \xi(2^j x - k)\}$ et $\{\eta_{jk}(x) = 2^{j/2} \eta(2^j x - k)\}$ constituent une base orthonormale complète de V_j et W_j , respectivement. Chaque sous-espace V_j contient les informations du signal de niveau de résolution j et

chaque sous-espace W_j contient les détails qui sont la différence d'informations entre deux résolutions, donc entre V_j et V_{j+1} . Les signaux de niveau j peuvent être combinés avec les détails de niveau j afin d'obtenir des signaux de niveau $j + 1$.

Intuitivement, une petite valeur de j permet de décrire des composantes plus lisses des signaux tandis qu'une grande valeur de j permet la capture des composantes très variables des signaux. De ce point de vue, une petite valeur de j est associée à une petite résolution alors qu'une plus grande valeur explique une plus grande résolution. De plus, pour chaque résolution j , le paramètre de translation k permet d'expliquer localement les traits des signaux.

Dans notre application, la fonction d'échelle $\xi(\cdot)$ décrit les parties les plus lisses de la densité spectrale et les ondelettes $\{\eta(\cdot)\}$ capturent la différence telle que les pics et les sommets. La condition (i) implique, d'une part, que le signal observé à un certain niveau de résolution contient toutes les informations des signaux à des niveaux de résolution plus grands et, d'autre part, qu'une fonction de $L_2(\mathbb{R})$ peut être approximée relativement bien par une résolution plus fine (c'est-à-dire par une valeur de j plus élevée). Les conditions (ii) et (iii) dans la définition 2.5.1 représentent l'invariance de l'échelle et de la translation, respectivement.

Dans la littérature, plusieurs ondelettes ont été développées. Un exemple classique est l'ondelette de Haar, définie par $\xi_H(x) = 1, x \in [0, 1]$ et par $\eta_H(x)$ satisfaisant:

$$\eta_H(x) = \begin{cases} 1, & x \in [0, 1/2), \\ -1, & x \in [1/2, 1), \\ 0, & \text{sinon.} \end{cases} \quad (2.5.1)$$

Un autre exemple est l'ondelette de Shannon qui est définie par $\xi_S(x) = \sin(\pi x)/(\pi x)$ et dont la fonction $\eta_S(x)$ satisfait:

$$\eta_S(x) = -2 \left\{ \frac{\sin(2\pi x) + \cos(\pi x)}{\pi(2x + 1)} \right\}, \quad x \in \mathbb{R}. \quad (2.5.2)$$

Il est à noter que l'ondelette de Haar est bornée et donc représente une onde qui apparaît et qui disparaît tandis que l'ondelette de Shannon perdure infiniment, mais l'intensité diminue rapidement avec le temps.

Par définition, une ondelette doit respecter les conditions suivantes:

- (i) $\int_{-\infty}^{\infty} \eta(x) dx = 0$,
- (ii) $\int_{-\infty}^{\infty} |\eta(x)| dx < \infty$,
- (iii) $\int_{-\infty}^{\infty} \eta(x)\eta(x-k) dx = 0$, pour tout $k \in \mathbb{Z}$, $k \neq 0$,
- (iv) $\int_{-\infty}^{\infty} \eta^2(x) dx = 1$.

Les ondelettes de Haar et de Shannon respectent ces conditions. L'orthonormalité de $\eta(\cdot)$ implique que la séquence doublement infinie $\{\eta_{jk}(\cdot)\}$ constitue une base orthonormale de $L_2(\mathbb{R})$. Ainsi, nous avons les relations suivantes:

$$\int_{-\infty}^{\infty} \eta_{jk}(x)\eta_{lm}(x) dx = \delta_{jl}\delta_{km}, \quad j, l, k, m \in \mathbb{Z},$$

où δ_{jl} est la fonction δ définie par $\delta_{jl} = 1$ si $j = l$ et $\delta_{jl} = 0$, sinon.

La transformée de Fourier joue un rôle important dans l'étude des ondelettes. Elle est définie par la relation:

$$\hat{\eta}(w) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} \eta(x) e^{-iwx} dx, \quad w \in \mathbb{R}. \quad (2.5.3)$$

Les conditions mentionnées précédemment assurent son existence et on note également que $\hat{\eta}(w)$ définit une fonction continue. En statistique, pour un paramètre θ , on dénote $\hat{\theta}$ son estimateur. On remarque qu'en (2.5.3) on adopte la notation fréquemment rencontrée en notant $\hat{\eta}$ la transformée de Fourier de η . Ceci ne cause pas de soucis dans l'exposition de nos résultats.

Une classe importante d'ondelettes est constituée des ondelettes splines. Elles sont définies à partir de la transformée de Fourier. Lorsque m est un entier pair, $\eta_s^{(m)}$ peut s'exprimer comme:

$$\hat{\eta}_s^{(m)}(2x) = \frac{-\mathbf{i} \exp(\mathbf{i}x) \{\sin(x/2)\}^{2m+2}}{(2\pi)^{1/2} (x/2)^{m+1}} \left\{ \frac{P_{2m+1}(x/2 + \pi/2)}{P_{2m+1}(x)P_{2m+1}(x/2)} \right\}^{1/2}. \quad (2.5.4)$$

Une expression pour m impair se trouve dans Hernández et Weiss (1996), et:

$$P_n(x) = \sum_{k \in \mathbb{Z}} \frac{\{2 \sin(x)\}^{n+1}}{(2x + 2k\pi)^{n+1}}.$$

On vérifie que:

$$\begin{aligned}
 P_1(x) &= 1, \\
 P_2(x) &= \cos(x), \\
 P_3(x) &= 1 - \frac{2}{3} \sin^2(x), \\
 P_4(x) &= \frac{1}{3} \cos^3(x) + \frac{2}{3} \cos(x),
 \end{aligned}$$

voir Hernández et Weiss (1996, p. 157).

On se propose de calculer la transformée de Fourier de l'ondelette de Haar. Un calcul direct donne:

$$\begin{aligned}
 \hat{\eta}_H(w) &= (2\pi)^{-1/2} \int_{-\infty}^{\infty} \eta_H(x) e^{-iwx} dx, \\
 &= (2\pi)^{-1/2} \left\{ \int_0^{1/2} e^{-iwx} dx - \int_{1/2}^1 e^{-iwx} dx \right\}, \\
 &= \frac{\mathbf{i}}{w} (2\pi)^{-1/2} \left\{ 2 \exp\left(\frac{-\mathbf{i}w}{2}\right) - \exp(-\mathbf{i}w) - 1 \right\}, \\
 &= \frac{\mathbf{i}}{w} (2\pi)^{-1/2} \exp\left(\frac{-\mathbf{i}w}{2}\right) \left\{ 2 - \exp\left(\frac{-\mathbf{i}w}{2}\right) - \exp\left(\frac{\mathbf{i}w}{2}\right) \right\},
 \end{aligned}$$

où l'on utilise que $1/\mathbf{i} = -\mathbf{i}$. Utilisant les relations $\exp(\mathbf{i}\theta) = \cos \theta + \mathbf{i} \sin \theta$ et $2 \sin^2(\theta) = 1 - \cos(2\theta)$, on trouve que:

$$\hat{\eta}_H(w) = \mathbf{i} (2\pi)^{-1/2} \exp\left(\frac{-\mathbf{i}w}{2}\right) \frac{\sin^2(w/4)}{w/4}, \quad (2.5.5)$$

et ainsi en posant $m = 0$ dans (2.5.4) on constate que l'ondelette de Haar correspond essentiellement à l'ondelette spline de degré 0. On constate en étudiant (2.5.5) que $|\hat{\eta}_H(w)|$ tend vers 0 quand $|w| \rightarrow \infty$ à la vitesse $|w|^{-1}$, ce qui est considéré relativement lent, et même quelque peu problématique dans nos analyses mathématiques.

En fait, pour des raisons techniques dans les preuves des résultats techniques, l'ondelette $\eta(w)$ doit être telle que la transformée de Fourier tend plus rapidement que $|w|^{-1}$. Plus précisément, cette hypothèse est nécessaire dans l'obtention de la distribution limite des coefficients empiriques d'ondelettes. En fait, nous aurons besoin de supposer la

condition suivante lorsque w est assez grand:

$$|\hat{\eta}(w)| \leq C(1 + |w|)^{-\alpha}, \quad (2.5.6)$$

pour $\alpha > 1$ et une constante C . Une question naturelle porte sur l'existence d'ondelettes telles que (2.5.6) est remplie. Heureusement, plusieurs ondelettes satisfont cette condition, notamment les ondelettes de Meyer et plus généralement les ondelettes splines avec des ordres strictement positifs.

Dans l'élaboration de tests portemanteaux utilisant la théorie des ondelettes, les résultats disponibles, à la fois théoriques et empiriques suggèrent que le choix de l'ondelette joue un rôle quelque peu secondaire. Voir par exemple Lee et Hong (2001) ou encore Duchesne, Li et Vandermeerschen (2010). Pour cette raison, nous nous limitons à l'ondelette spline de degré un, appelée également ondelette de Franklin. Dans le domaine des fréquences, l'ondelette de Franklin admet les transformées de Fourier suivantes:

$$\begin{aligned} \hat{\xi}(w) &= (2\pi)^{-1/2} \frac{\sin^2(w/2)}{(w/2)^2} \{1 - (2/3) \sin^2(w/2)\}^{-1/2}, \\ \hat{\eta}(w) &= e^{iw/2} (2\pi)^{-1/2} \frac{\sin^4(w/4)}{(w/4)^2} \left[\frac{1 - (2/3) \cos^2(w/4)}{\{1 - (2/3) \sin^2(w/2)\} \{1 - (2/3) \sin^2(w/4)\}} \right]^{1/2}. \end{aligned} \quad (2.5.7)$$

L'ondelette de Franklin est construite à l'aide de fonctions linéaires par morceaux. Elle a un support compact dans le domaine temporel et une décroissance exponentielle dans le domaine de fréquences. Pour plus de détails, voir Hernández et Weiss (1996, p. 148). Donc, nous avons $|\hat{\eta}(w)| \leq C(1 + |w|)^{-2}$ quand $|w| \rightarrow \infty$ pour l'ondelette de Franklin, ce qui respecte la condition (2.5.6) avec $\alpha = 2$. Pour les autres ondelettes et leurs propriétés, voir Hernández et Weiss (1996), Vidakovic (1999), Daubechies (1992), entre autres.

2.6. Ondelettes et estimateurs de densité spectrale à base d'ondelettes

Dans cette section, on se propose de développer des procédures de test afin de confronter les hypothèses énoncées en (2.3.1). On rappelle que l'hypothèse nulle correspond à l'absence de corrélation dans le terme d'erreur:

$$c_\varepsilon(h) = \text{Cov}(\varepsilon_t, \varepsilon_{t+h}) = 0, \quad \forall h \neq 0,$$

alors que sous l'alternative, de la corrélation est présente pour au moins un délai h , et donc on suppose que l'alternative prend la forme:

$$c_\varepsilon(h) \neq 0,$$

pour au moins un $h \neq 0$.

Dans la section 2.5, on a déjà introduit la densité spectrale associée à un processus stationnaire au sens large qui admet une fonction d'autocovariance absolument sommable. Lorsque l'on calcule la densité spectrale du bruit $\{\varepsilon_t\}$ dans le modèle ARMA formulé en (2.2.1), on trouve que la densité spectrale du terme d'erreur est :

$$f_\varepsilon(w) = (2\pi)^{-1} \sum_{h=-\infty}^{\infty} c_\varepsilon(h) e^{-ihw}, \quad w \in [-\pi, \pi]. \quad (2.6.1)$$

On peut donc reformuler les hypothèses H_0 et H_1 en fonction de la densité spectrale de $\{\varepsilon_t\}$. Sous l'hypothèse nulle, la fonction d'autocorrélation est nulle et la densité spectrale devient une constante:

$$f_\varepsilon(w) = \frac{\sigma_\varepsilon^2}{2\pi}, \quad w \in [-\pi, \pi]. \quad (2.6.2)$$

Cependant, sous l'hypothèse alternative, la méthodologie proposée devrait avoir de la puissance lorsque de la corrélation est présente pour au moins un délai, de sorte que la densité spectrale ne devrait pas être constante pour $w \in [-\pi, \pi]$, ce qui implique dans de tels cas que sous H_1 :

$$f_\varepsilon(w) \neq \frac{\sigma_\varepsilon^2}{2\pi}, \quad w \in [-\pi, \pi].$$

Le test proposé devra comparer selon une certaine mesure de distance globale un estimateur de la densité spectrale calculé avec les données, à la densité spectrale sous l'hypothèse nulle fournie par l'expression (2.6.2).

On se propose de ré-écrire la densité spectrale du processus $\{\varepsilon_t\}$ en utilisant une base d'ondelettes. On utilise ici l'approche mise de l'avant dans Lee et Hong (2001). On note qu'en général la densité spectrale de $\{\varepsilon_t\}$ (sous H_0 et surtout sous H_1) est une fonction périodique de période 2π sur \mathbb{R} . Il est alors souhaitable de choisir une base d'ondelettes qui tient compte de cette particularité.

Posons $\Pi = [-\pi, \pi]$ l'intervalle naturel de définition de la densité spectrale. On considère une base d'ondelettes notée $\{\Phi_{jk}(\cdot), \Psi_{jk}(\cdot)\}$, qui est définie sur $L_2(\Pi)$, l'espace des

fonctions de carré intégrable sur l'intervalle Π . Par la suite, on fait appel à la technique dite de périodisation.

La technique de périodisation est décrite en détail dans Mallat (1999, pp. 282–284). Voir aussi Hernández et Weiss (1996, Section 4.5, pp. 186–193). Supposons que l'on dispose d'une base orthonormale d'ondelettes $\{\xi_{jk}(\cdot), \eta_{jk}(\cdot)\}$ sur $L_2(\mathbb{R})$. La base d'ondelettes de $L_2(\Pi)$ obtenue par périodisation de $\{\xi_{jk}(\cdot), \eta_{jk}(\cdot)\}$ est obtenue selon les expressions suivantes :

$$\begin{aligned}\Phi_{jk}(w) &= \frac{1}{(2\pi)^{1/2}} \sum_{m=-\infty}^{\infty} \xi_{jk} \left(\frac{w}{2\pi} + m \right), \\ \Psi_{jk}(w) &= \frac{1}{(2\pi)^{1/2}} \sum_{m=-\infty}^{\infty} \eta_{jk} \left(\frac{w}{2\pi} + m \right),\end{aligned}\tag{2.6.3}$$

où $-\infty < w < \infty$. Dans l'expression (2.6.3), on vérifie facilement que $\Phi_{jk}(w)$ et $\Psi_{jk}(w)$ sont maintenant des fonctions périodiques à valeurs réelles et de période 2π .

Mallat (1999, p. 283) montre que les fonctions $\{\Phi_{jk}(\cdot), 1 \leq k \leq 2^{j_0}, \Psi_{jk}(\cdot), j \geq j_0, 1 \leq k \leq 2^j\}$ forment une base orthonormale de l'espace $L_2(\Pi)$, et ce quelque soit $j_0 \geq 0$. On note que les notations de Mallat (1999) doivent être quelques peu adaptées (car il considère des fonctions sur $L_2([0, 1])$), et que de plus l'indice k est tel que $1 \leq k \leq 2^j$, plutôt que $0 \leq k \leq 2^j - 1$, ce qui semble souvent considéré dans la littérature.

Ainsi, en posant $j_0 = 0$, on obtient que la densité spectrale peut s'exprimer comme suit:

$$f_\varepsilon(w) = \alpha_{00}\Phi_{00}(w) + \sum_{j=0}^{\infty} \sum_{k=1}^{2^j} \beta_{jk}\Psi_{jk}(w), \quad w \in [-\pi, \pi],\tag{2.6.4}$$

où

$$\begin{aligned}\alpha_{00} &= \int_{-\pi}^{\pi} f_\varepsilon(w)\Phi_{00}(w)dw, \\ \beta_{jk} &= \int_{-\pi}^{\pi} f_\varepsilon(w)\Psi_{jk}(w)dw, \quad j \geq 0, k = 1, \dots, 2^j.\end{aligned}$$

Les coefficients d'ondelettes β_{jk} définis dans (2.6.4) ont des propriétés utiles.

Lemme 2.6.1. *Supposons que l'innovation $\{\varepsilon_t\}$ satisfait l'hypothèse A. Supposons aussi que la fonction de l'ondelette de Franklin $\eta(\cdot)$ dans (2.5.7) est adoptée. Alors, nous avons que les coefficients d'ondelettes β_{jk} définis par (2.6.4) satisfont la relation:*

$$\beta_{jk_1} = \beta_{jk_2}, \quad \text{si } k_1 + k_2 = 2^j + 1,$$

pour tout $j = 1, 2, \dots$ et $k_1, k_2 = 1, 2, \dots, 2^j$.

Du lemme 2.6.1, dont les idées principales de la preuve se trouvent dans Duchesne, Li et Vandermeersch (2010), il ressort qu'au plus la moitié des coefficients d'ondelettes β_{jk} , $k = 1, 2, \dots, 2^{j-1}$ sont nécessaires.

Soient $\hat{\Phi}_{jk}(\cdot)$ et $\hat{\xi}_{jk}(\cdot)$ les transformées de Fourier de $\Phi_{jk}(\cdot)$ et $\xi_{jk}(\cdot)$, respectivement. Nous avons alors que:

$$\begin{aligned}
\hat{\Phi}_{jk}(h) &= \frac{1}{(2\pi)^{1/2}} \int_{-\pi}^{\pi} \Phi_{jk}(w) e^{-iwh} dw, \\
&= \frac{1}{(2\pi)^{1/2}} \int_{-\pi}^{\pi} \left\{ \frac{1}{(2\pi)^{1/2}} \sum_{m=-\infty}^{\infty} \xi_{jk} \left(\frac{w}{2\pi} + m \right) \right\} e^{-iwh} dw, \\
&= \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \int_{-\pi}^{\pi} \xi_{jk} \left(\frac{w}{2\pi} + m \right) e^{-iwh} dw, \\
&= \sum_{m=-\infty}^{\infty} \int_{m-1/2}^{m+1/2} \xi_{jk}(x) e^{-ix(2\pi h)} dx, \\
&= \int_{-\infty}^{\infty} \xi_{jk}(x) e^{-ix(2\pi h)} dx, \\
&= \sqrt{2\pi} \hat{\xi}_{jk}(2\pi h), \\
&= e^{-i2\pi h k / 2^j} \left(\frac{2\pi}{2^j} \right)^{1/2} \hat{\xi} \left(\frac{2\pi h}{2^j} \right).
\end{aligned}$$

Ainsi, pour l'ondelette de Franklin, on trouve que:

$$\hat{\Phi}_{00}(h) = \sqrt{2\pi} \hat{\xi}(2\pi h) = \begin{cases} 1, & h = 0, \\ 0, & h \neq 0. \end{cases} \quad (2.6.5)$$

De plus:

$$\Phi_{00}(w) = (2\pi)^{-1/2} \sum_{h=-\infty}^{\infty} \hat{\Phi}_{00}(h) e^{iwh} = (2\pi)^{-1/2},$$

pour tout $w \in [-\pi, \pi]$. On a alors que:

$$\begin{aligned}
\alpha_{00} &= \int_{-\pi}^{\pi} f_{\varepsilon}(w) \Phi_{00}(w) dw, \\
&= \frac{1}{(2\pi)^{1/2}} \int_{-\pi}^{\pi} f_{\varepsilon}(w) dw, \\
&= \frac{\sigma_{\varepsilon}^2}{(2\pi)^{1/2}},
\end{aligned}$$

puisque:

$$\begin{aligned}\int_{-\pi}^{\pi} f_{\varepsilon}(w)dw &= \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} c_{\varepsilon}(h) \int_{-\pi}^{\pi} e^{-ihw} dw, \\ &= c_{\varepsilon}(0), \\ &= \sigma_{\varepsilon}^2,\end{aligned}$$

utilisant la relation:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ihw} dw = \begin{cases} 1, & h = 0, \\ 0, & \text{sinon,} \end{cases}$$

que l'on retrouve par exemple dans Brockwell et Davis (1991, p. 120). On remarque également que $c_{\varepsilon}(h) = \int_{-\pi}^{\pi} e^{ihw} f_{\varepsilon}(w)dw$, et on obtient également $c_{\varepsilon}(0) = \int_{-\pi}^{\pi} f_{\varepsilon}(w)dw$. Ainsi, la densité spectrale peut s'écrire de manière alternative de la manière suivante:

$$f_{\varepsilon}(w) = \frac{\sigma_{\varepsilon}^2}{2\pi} + \sum_{j=0}^{\infty} \sum_{k=1}^{2^j} \beta_{jk} \Psi_{jk}(w), \quad w \in [-\pi, \pi]. \quad (2.6.6)$$

On note que l'argumentation précédente utilise les propriétés de l'ondelette de Franklin. La représentation (2.6.6) est valide plus généralement pour des ondelettes satisfaisant certaines conditions de régularité. On peut consulter Lee et Hong (2001).

Sous l'hypothèse nulle H_0 , les coefficients d'ondelettes β_{jk} s'annulent. En effet:

$$\begin{aligned}\beta_{jk} &= \int_{-\pi}^{\pi} f_{\varepsilon}(w) \Psi_{jk}(w) dw, \\ &= \frac{\sigma_{\varepsilon}^2}{2\pi} \int_{-\pi}^{\pi} \Psi_{jk}(w) dw, \\ &= 0,\end{aligned}$$

pour tout $j \geq 0$, $k = 1, \dots, 2^j$. On note que l'on utilise encore l'argument suivant:

$$\begin{aligned}\int_{-\pi}^{\pi} \Psi_{jk}(w) dw &= \frac{1}{(2\pi)^{1/2}} \sum_{m=-\infty}^{\infty} \left\{ \int_{-\pi}^{\pi} \eta_{jk} \left(\frac{w}{2\pi} + w \right) dw \right\}, \\ &= \frac{1}{(2\pi)^{1/2}} \sum_{m=-\infty}^{\infty} \int_{m-1/2}^{m+1/2} \eta_{jk}(x) dx, \\ &= \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\infty} \eta_{jk}(x) dx, \\ &= 0.\end{aligned}$$

Ainsi la densité spectrale devient sous l'hypothèse nulle $f_\varepsilon(w) \equiv f_0(w) = \sigma_\varepsilon^2/(2\pi)$, $w \in [-\pi, \pi]$, ce qui est en accord avec ce que l'on a déjà établi précédemment.

Ceci implique que les hypothèses originales peuvent être reformulées en utilisant les coefficients d'ondelettes β_{jk} , $j, k \in \mathbb{Z}$. En effet, les hypothèses (2.3.1) peuvent s'exprimer comme suit:

$$\begin{aligned} H_0: \beta_{jk} &= 0, \quad j \geq 0, k = 1, \dots, 2^j, \\ H_1: \beta_{jk} &\neq 0, \text{ pour au moins une paire } (j, k). \end{aligned} \tag{2.6.7}$$

Afin de développer notre test, nous devons commencer par proposer un estimateur de la densité spectrale exploitant la représentation alternative (2.6.6). On commence par exprimer les coefficients β_{jk} en fonction de la fonction d'autocovariance du processus. Un calcul direct donne:

$$\begin{aligned} \beta_{jk} &= \int_{-\pi}^{\pi} \left\{ \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} c_\varepsilon(h) e^{-ihw} \right\} \Psi_{jk}(w) dw, \\ &= \frac{1}{(2\pi)^{1/2}} \sum_{h=-\infty}^{\infty} c_\varepsilon(h) \hat{\Psi}_{jk}(h), \\ &= \sum_{h=-\infty}^{\infty} c_\varepsilon(h) \hat{\eta}_{jk}(2\pi h), \end{aligned}$$

où l'on utilise la définition de la densité spectrale pour $\{\varepsilon_t\}$ donnée par (2.6.1), et où l'on utilise la relation $\hat{\Psi}_{jk}(h) = (2\pi)^{1/2} \hat{\eta}_{jk}(2\pi h)$, qui s'établit en utilisant une démarche similaire à celle fournissant la relation $\hat{\Phi}_{jk}(h) = (2\pi)^{1/2} \hat{\xi}_{jk}(2\pi h)$.

On note en étudiant (2.6.4) que $f_\varepsilon(w)$ repose sur une infinité de coefficients d'ondelettes β_{jk} . Il faudra que l'estimateur repose sur un nombre fini de résolutions, ce que nous appelons J_n . Intuitivement, J_n représente l'échelle la plus fine possible. Ceci mène à l'estimateur suivant:

$$\hat{f}_{\varepsilon, J_n}(w) = \frac{\hat{\sigma}_\varepsilon^2}{2\pi} + \sum_{j=0}^{J_n} \sum_{k=1}^{2^j} \bar{\beta}_{jk} \Psi_{jk}(w), \quad w \in [-\pi, \pi]. \tag{2.6.8}$$

Basé sur le processus non observable $\{\varepsilon_t\}$, les coefficients d'ondelettes empiriques $\bar{\beta}_{jk}$ sont définis par:

$$\bar{\beta}_{jk} = \sum_{h=-(n-1)}^{n-1} \hat{c}_\varepsilon(h) \hat{\eta}_{jk}(2\pi h) = \sum_{h=1}^{n-1} \hat{c}_\varepsilon(h) \{ \hat{\eta}_{jk}(2\pi h) + \hat{\eta}_{jk}(-2\pi h) \}, \tag{2.6.9}$$

où $\hat{c}_\varepsilon(h) = r_\varepsilon(h) = n^{-1} \sum_{t=1}^{n-h} \varepsilon_t \varepsilon_{t+h}$. Puisque $\hat{\eta}_{jk}(0) = 0$, pour tout $j = 0, 1, \dots, J_n$ et $k = 1, 2, \dots, 2^j$, les coefficients d'ondelettes empiriques peuvent être réécrits:

$$\bar{\beta}_{jk} = \sum_{h=1}^{n-1} r_\varepsilon(h) b_{jk}(h), \quad (2.6.10)$$

où $b_{jk}(h) = \hat{\eta}_{jk}(2\pi h) + \hat{\eta}_{jk}(-2\pi h)$.

Les coefficients d'ondelettes empiriques ne représentent qu'une première définition, puisque ces derniers reposent sur le bruit non observable $\{\varepsilon_t\}$. De façon naturelle, l'autocovariance empirique calculée avec $\{\varepsilon_t\}$ est elle-même estimée en remplaçant les ε_t par les résidus, de sorte que nous remplaçons $r_\varepsilon(h)$ par l'autocovariance résiduelle, c'est-à-dire que nous considérons:

$$\hat{\beta}_{jk} = \sum_{h=1}^{n-1} \hat{r}(h) b_{jk}(h), \quad (2.6.11)$$

où $\hat{r}(h) = \hat{c}_\varepsilon(h) = n^{-1} \sum_{t=1}^{n-h} \hat{\varepsilon}_t \hat{\varepsilon}_{t+h}$.

Des résultats similaires à ce qui est évoqué dans le lemme 2.6.1 sont aussi valables pour les coefficients d'ondelettes empiriques. Ceci fait l'objet du lemme suivant:

Lemme 2.6.2. *Supposons que $\{X_t\}$ suit un modèle ARMA avec un processus d'innovation satisfaisant l'hypothèse A. Supposons comme dans le lemme 2.6.1 que l'ondelette de Franklin est adoptée. Considérons les coefficients d'ondelettes empiriques calculés avec les résidus du modèle ARMA. Alors les coefficients $\hat{\beta}_{jk}$ définis par (2.6.11) satisfont:*

$$\hat{\beta}_{jk_1} = \hat{\beta}_{jk_2}, \quad \text{si } k_1 + k_2 = 2^j + 1,$$

pour tout $j = 1, 2, \dots, J_n$ et $k_1, k_2 = 1, 2, \dots, 2^j$. Le nombre J_n est le plus grand entier satisfaisant $2^{J_n+1} \leq n$.

Le lemme 2.6.2 se trouve dans Duchesne, Li et Vandermeerschen (2010). Ainsi, pour construire les statistiques de test, à chaque résolution j , $j = 1, \dots, J_n$, la moitié des coefficients d'ondelettes empiriques sont nécessaires. Pour une discussion plus approfondie sur les coefficients d'ondelettes $\hat{\beta}_{jk}$ et β_{jk} utilisant d'autres ondelettes, voir Lee et Hong (2001) et Duchesne, Li et Vandermeerschen (2010).

De la définition de $b_{jk}(h)$ dans (2.6.10) et en utilisant l'ondelette de Franklin $\eta(\cdot)$ dans (2.5.7), nous obtenons:

$$b_{jk}(h) = \frac{2^{(3j/2)+5}}{\sqrt{2\pi}} \cos \left\{ \frac{2\pi h(k-1/2)}{2^j} \right\} \frac{\sin^4(2\pi h/2^{j+2})}{(2\pi h)^2} \\ \times \left[\frac{1 - (2/3) \cos^2(2\pi h/2^{j+2})}{\{1 - (2/3) \sin^2(2\pi h/2^{j+1})\} \{1 - (2/3) \sin^2(2\pi h/2^{j+2})\}} \right]^{1/2}. \quad (2.6.12)$$

Et donc de la définition de $\hat{\beta}_{jk}$ dans (2.6.11), en utilisant encore la relation $\hat{\eta}_{jk}(2\pi h) = e^{-i2\pi h k/2^j} 2^{-j/2} \hat{\eta}(2\pi h/2^j)$, la transformée de Fourier de l'ondelette de Franklin et un peu d'algèbre, nous avons l'expression suivante:

$$\hat{\beta}_{jk} = \frac{2^{(3j/2)+5}}{\sqrt{2\pi}} \sum_{h=1}^{n-1} \hat{r}(h) \cos \left\{ \frac{2\pi h(k-1/2)}{2^j} \right\} \frac{\sin^4(2\pi h/2^{j+2})}{(2\pi h)^2} \\ \times \left[\frac{1 - (2/3) \cos^2(2\pi h/2^{j+2})}{\{1 - (2/3) \sin^2(2\pi h/2^{j+1})\} \{1 - (2/3) \sin^2(2\pi h/2^{j+2})\}} \right]^{1/2}. \quad (2.6.13)$$

Malgré l'allure imposante de l'expression de $\hat{\beta}_{jk}$ donnée par (2.6.13), les seules variables aléatoires sont les autocovariances résiduelles $\hat{r}(h)$ et les $b_{jk}(h)$ peuvent être interprétés comme des pondérations des autocovariances résiduelles qui dépendent de l'ondelette choisie (dans notre cas, l'ondelette de Franklin).

2.7. Distributions asymptotiques des coefficients d'ondelettes empiriques

L'objectif de cette section est de présenter la distribution asymptotique des coefficients d'ondelettes $\hat{\beta}_{jk}$ définis par l'expression (2.6.11). Les résultats sont obtenus sous l'hypothèse que le modèle ARMA est correctement spécifié.

Dénotons le vecteur de m éléments $\mathbf{B}_{jk,m} = (b_{jk}(1), b_{jk}(2), \dots, b_{jk}(m))^\top$ et à partir des matrices introduites à la section 2.4, nous introduisons:

$$\begin{aligned}\mathbf{B}_{jk,\infty}^\top \mathbf{\Gamma}_{\infty,\infty} \mathbf{B}_{jk,\infty} &= \sum_{l=1}^{\infty} \sum_{l'=1}^{\infty} b_{jk}(l) \Gamma(l, l'), \\ \mathbf{\Lambda}_{\infty} \mathbf{B}_{jk,\infty} &= \sum_{i=1}^{\infty} b_{jk}(i) \boldsymbol{\lambda}_i, \\ \mathbf{\Lambda}_{\infty} \mathbf{\Gamma}_{\infty,\infty} \mathbf{B}_{jk,\infty} &= \sum_{i=1}^{\infty} \sum_{l=1}^{\infty} \boldsymbol{\lambda}_i \Gamma(i, l) b_{jk}(l),\end{aligned}$$

où $\boldsymbol{\lambda}_i$ est défini par l'expression (2.4.3), et où $\mathbf{B}_{jk,\infty}^\top \mathbf{\Gamma}_{\infty,\infty} \mathbf{B}_{jk,\infty}$ est un scalaire, alors que $\mathbf{\Lambda}_{\infty} \mathbf{B}_{jk,\infty}$ et $\mathbf{\Lambda}_{\infty} \mathbf{\Gamma}_{\infty,\infty} \mathbf{B}_{jk,\infty}$ sont des vecteurs de $(p+q)$ éléments. De plus, notons $\mathbf{B}_{jk,\infty}^\top \mathbf{\Lambda}_{\infty}^\top$ et $\mathbf{B}_{jk,\infty}^\top \mathbf{\Gamma}_{\infty,\infty} \mathbf{\Lambda}_{\infty}^\top$ les transposées de $\mathbf{\Lambda}_{\infty} \mathbf{B}_{jk,\infty}$ et $\mathbf{\Lambda}_{\infty} \mathbf{\Gamma}_{\infty,\infty} \mathbf{B}_{jk,\infty}$, respectivement.

Les distributions asymptotiques des coefficients d'ondelettes sont énoncées dans le théorème 2.7.1.

Théorème 2.7.1. *Supposons que $\{X_t\}$ soit généré par le modèle ARMA (2.2.1) et l'innovation $\{\varepsilon_t\}$ satisfait l'hypothèse A. Soit J ($1 \leq J < J_n$) un entier fixé et supposons que l'ondelette de Franklin $\eta(\cdot)$ soit adoptée. Alors, nous avons, pour tout $0 \leq j, j_1, j_2 \leq J$ et tout $k = 1, \dots, 2^{j-1}$, $k_1 = 1, \dots, 2^{j_1-1}$, $k_2 = 1, \dots, 2^{j_2-1}$ (pour le cas particulier $j = 0$, $k = 1$, nous traitons $2^{0-1} = 1$ afin que l'expression soit la même que dans les autres cas):*

$$\begin{aligned}\sqrt{n} \hat{\beta}_{jk} &\longrightarrow_d \mathcal{N}(0, \sigma_{jk}^2), \\ \text{Cov}(\sqrt{n} \hat{\beta}_{j_1 k_1}, \sqrt{n} \hat{\beta}_{j_2 k_2}) &\longrightarrow \sigma_{j_1 k_1, j_2 k_2},\end{aligned}\tag{2.7.1}$$

où $\sigma_{jk}^2 = \mathbf{B}_{jk,\infty}^\top \mathbf{\Gamma}_{\infty,\infty} \mathbf{B}_{jk,\infty}$ et $\sigma_{j_1 k_1, j_2 k_2} = \mathbf{B}_{j_1 k_1, \infty}^\top \mathbf{\Gamma}_{\infty,\infty} \mathbf{B}_{j_2 k_2, \infty} = \sum_{l, l'} b_{j_1 k_1}(l) \Gamma(l, l') b_{j_2 k_2}(l')$ quand $p = q = 0$, et quand $p > 0$ ou $q > 0$:

$$\begin{aligned}\sigma_{jk}^2 &= \mathbf{B}_{jk,\infty}^\top \mathbf{\Gamma}_{\infty,\infty} \mathbf{B}_{jk,\infty} \\ &+ \mathbf{B}_{jk,\infty}^\top \mathbf{\Lambda}_{\infty}^\top (\mathbf{\Lambda}_{\infty} \mathbf{\Lambda}_{\infty}^\top)^{-1} \mathbf{\Lambda}_{\infty} \mathbf{\Gamma}_{\infty,\infty} \mathbf{\Lambda}_{\infty}^\top (\mathbf{\Lambda}_{\infty} \mathbf{\Lambda}_{\infty}^\top)^{-1} \mathbf{\Lambda}_{\infty} \mathbf{B}_{jk,\infty} \\ &- \mathbf{B}_{jk,\infty}^\top \mathbf{\Lambda}_{\infty}^\top (\mathbf{\Lambda}_{\infty} \mathbf{\Lambda}_{\infty}^\top)^{-1} \mathbf{\Lambda}_{\infty} \mathbf{\Gamma}_{\infty,\infty} \mathbf{B}_{jk,\infty} \\ &- \mathbf{B}_{jk,\infty}^\top \mathbf{\Gamma}_{\infty,\infty} \mathbf{\Lambda}_{\infty}^\top (\mathbf{\Lambda}_{\infty} \mathbf{\Lambda}_{\infty}^\top)^{-1} \mathbf{\Lambda}_{\infty} \mathbf{B}_{jk,\infty},\end{aligned}$$

$$\begin{aligned}
\sigma_{j_1 k_1, j_2 k_2} &= \mathbf{B}_{j_1 k_1, \infty}^\top \boldsymbol{\Gamma}_{\infty, \infty} \mathbf{B}_{j_2 k_2, \infty} \\
&+ \mathbf{B}_{j_1 k_1, \infty}^\top \boldsymbol{\Lambda}_\infty^\top (\boldsymbol{\Lambda}_\infty \boldsymbol{\Lambda}_\infty^\top)^{-1} \boldsymbol{\Lambda}_\infty \boldsymbol{\Gamma}_{\infty, \infty} \boldsymbol{\Lambda}_\infty^\top (\boldsymbol{\Lambda}_\infty \boldsymbol{\Lambda}_\infty^\top)^{-1} \boldsymbol{\Lambda}_\infty \mathbf{B}_{j_2 k_2, \infty} \\
&- \mathbf{B}_{j_1 k_1, \infty}^\top \boldsymbol{\Lambda}_\infty^\top (\boldsymbol{\Lambda}_\infty \boldsymbol{\Lambda}_\infty^\top)^{-1} \boldsymbol{\Lambda}_\infty \boldsymbol{\Gamma}_{\infty, \infty} \mathbf{B}_{j_2 k_2, \infty} \\
&- \mathbf{B}_{j_1 k_1, \infty}^\top \boldsymbol{\Gamma}_{\infty, \infty} \boldsymbol{\Lambda}_\infty^\top (\boldsymbol{\Lambda}_\infty \boldsymbol{\Lambda}_\infty^\top)^{-1} \boldsymbol{\Lambda}_\infty \mathbf{B}_{j_2 k_2, \infty}.
\end{aligned}$$

Les calculs des expressions de σ_{jk}^2 et de $\sigma_{j_1 k_1, j_2 k_2}$ dans le théorème 2.7.1 sont similaires à celle de $\boldsymbol{\Sigma}_{\hat{\mathbf{r}}_m}$ du lemme 2.4.4, avec $m \rightarrow \infty$ puisque les seules variables aléatoires dans l'expression de $\hat{\beta}_{jk}$ dans (2.6.13) sont les autocovariances résiduelles.

Une esquisse de la preuve du théorème 2.7.1 se trouve dans l'annexe. Sans surprise, une expansion de Taylor et le lemme 2.4.4 sont nécessaires. Puisque le résultat est obtenu sous l'hypothèse que le modèle ARMA est correctement spécifié, on pourra utiliser la distribution limite afin de tester l'hypothèse nulle d'adéquation, voir les énoncés (2.3.1) et (2.6.7). Les tests portemanteaux proposés dans la section 3.3 reposent sur le théorème 2.7.1.

Un objectif est de développer les tests portemanteaux lorsque le bruit blanc du modèle ARMA est plus général qu'un bruit blanc fort. On a mentionné que l'hypothèse différence de martingales comme dans la définition 2.3.1 pouvait être assez générale et intéressante.

L'objectif de la prochaine section est d'introduire la classe des modèles ARCH ainsi que les modèles ARCH généralisés (GARCH), qui est une classe de modèles particulièrement naturels dans notre contexte.

2.8. Modèle GARCH

La classe des modèles ARCH et GARCH contient des modèles de séries chronologiques fort populaires dans les applications impliquant des séries chronologiques financières. Notre présentation s'inspire de Brockwell et Davis (2016, Chapitre 7). Les caractéristiques reconnues des séries chronologiques provenant de la finance incluent notamment des ailes plus lourdes comparativement à la distribution normale, des phénomènes de volatilité par grappes (*volatility clustering* en anglais), et ce qui nous intéresse plus particulièrement ici, l'absence de corrélation, mais pas de dépendance.

Un exemple classique porte sur la série de prix d'un actif, disons $\{P_t, t \in \mathbb{Z}\}$. Bien que la série chronologique des prix soit typiquement incompatible avec l'hypothèse de stationnarité, les financiers et économètres s'intéressent plus particulièrement à ce qui est appelé les rendements, définis par $r_t = \log(P_t/P_{t-1})$, que l'on a déjà discutés dans la remarque 2.2.1. L'hypothèse de stationnarité de $\{r_t\}$ est typiquement raisonnable, et il est alors possible de modéliser les rendements par des modèles ARMA, voir la section 2.2.

Dans la définition suivante, on introduit la classe des modèles GARCH.

Définition 2.8.1. *Soit $\{\eta_t\}$ un bruit blanc fort, ainsi constitué de variables aléatoires indépendantes et satisfaisant:*

$$\begin{aligned} \mathbb{E}(\eta_t) &= 0, \\ \text{Var}(\eta_t) &= 1. \end{aligned}$$

Le processus stochastique $\{\varepsilon_t\}$ est dit provenir d'un modèle autorégressif conditionnellement hétéroscédastique généralisé d'ordre $r \geq 0$ et $s \geq 0$, noté GARCH(r,s), si:

$$\begin{aligned} \varepsilon_t &= \sigma_t \eta_t, \\ \sigma_t^2 &= \omega + \sum_{i=1}^r \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2, \end{aligned} \tag{2.8.1}$$

où $\omega > 0$, $\alpha_i \geq 0$, $i = 1, \dots, r$ et $\beta_j \geq 0$, $j = 1, \dots, s$.

Ainsi, si $r = s = 0$ dans la définition 2.8.1, on trouve que $\{\varepsilon_t\}$ est un bruit blanc fort tel que $\mathbb{V}(\varepsilon_t) = \omega > 0$. L'astuce déterminante dans les modèles GARCH est l'inclusion de σ_t^2 , qui est appelée la volatilité. On voit aisément que la volatilité n'est en fait rien d'autre que la variance conditionnelle de ε_t . Souvent, $\{\eta_t\}$ est un bruit blanc Gaussien, bien que ce ne soit pas absolument nécessaire.

Nous pouvons réécrire l'équation (2.8.1) en utilisant l'opérateur retard:

$$\sigma_t^2 = \omega + \alpha(B)\varepsilon_t^2 + \beta(B)\sigma_t^2, \tag{2.8.2}$$

où $\alpha(\cdot)$ et $\beta(\cdot)$ sont les polynômes suivants:

$$\begin{aligned} \alpha(z) &= \alpha_1 z + \alpha_2 z^2 + \dots + \alpha_r z^r, \\ \beta(z) &= \beta_1 z + \beta_2 z^2 + \dots + \beta_s z^s. \end{aligned}$$

Si $\beta(B) \equiv 0$, alors la volatilité se réduit à:

$$\sigma_t^2 = \omega + \sum_{i=1}^r \alpha_i \varepsilon_{t-i}^2,$$

qui est appelé également un modèle ARCH(r).

L'étude de la stationnarité dans les modèles ARCH et GARCH est abordée dans Francq et Zakoïan (2010). Sous des conditions d'existence d'une solution au GARCH, il peut être montré que la condition nécessaire et suffisante de stationnarité pour un GARCH(r,s) est:

$$\sum_{i=1}^r \alpha_i + \sum_{j=1}^s \beta_j < 1,$$

que l'on peut écrire également de façon succincte comme:

$$\alpha(1) + \beta(1) < 1.$$

S'il existe une solution stationnaire au GARCH, on note alors que $\{\varepsilon_t\}$ est une différence de martingales au sens de la définition 2.3.1, car σ_t^2 est alors fonction de $\{\varepsilon_s, s < t\}$ et:

$$\begin{aligned} \mathbf{E}(\varepsilon_t | \mathcal{F}_{t-1}) &= \sigma_t \mathbf{E}(\eta_t | \mathcal{F}_{t-1}), \\ &= \sigma_t \mathbf{E}(\eta_t), \\ &= 0, \end{aligned}$$

et on montre que $\{\varepsilon_t\}$ est un bruit blanc. Donc:

$$c_\varepsilon(h) = 0, \quad \forall h \neq 0.$$

Comme pour le modèle ARMA, nous supposons que les racines de $\alpha(\cdot)$ et $\beta(\cdot)$ sont toutes à l'extérieur du disque unitaire, ce qui nous permet d'écrire:

$$\begin{aligned} \sigma_t^2 &= \omega + \alpha(B)\varepsilon_t^2 + \beta(B)\sigma_t^2, \\ \{1 - \beta(B)\}\sigma_t^2 &= \omega + \alpha(B)\varepsilon_t^2, \\ \sigma_t^2 &= \frac{\omega}{1 - \beta(B)} + \frac{\alpha(B)}{1 - \beta(B)}\varepsilon_t^2, \\ \sigma_t^2 &= \omega^* + \sum_{i=1}^{\infty} \delta_i \varepsilon_{t-i}^2, \end{aligned}$$

où $\omega^* = \omega/\{1 - \beta(1)\}$ et $\delta_i = \alpha(B)/\{1 - \beta(B)\}$. Nous pouvons donc réécrire un GARCH(r,s) comme un ARCH(∞).

Par définition, l'innovation du processus $\{\varepsilon_t\}$ est $\{v_t = \varepsilon_t^2 - \sigma_t^2\}$. En remplaçant les σ_t^2 par $\varepsilon_t^2 - v_t$ dans (2.8.2), nous pouvons calculer:

$$\begin{aligned}\varepsilon_t^2 - v_t &= \omega + \alpha(B)\varepsilon_t^2 + \beta(B)(\varepsilon_t^2 - v_t), \\ \varepsilon_t^2 &= \omega + \{\alpha(B) + \beta(B)\}\varepsilon_t^2 + v_t - \beta(B)v_t, \\ \varepsilon_t^2 &= \omega + \sum_{i=1}^m (\alpha_i + \beta_i)\varepsilon_{t-i}^2 + v_t - \sum_{j=1}^s \beta_j v_{t-j},\end{aligned}$$

où $m = \max(r,s)$ avec, par convention, $\alpha_i = 0$ et $\beta_j = 0$ pour $i > r$ et $j > s$, respectivement. Cette équation possède la structure d'un modèle ARMA. Plus précisément, si le processus $\{\varepsilon_t\}$ est de modèle GARCH(r,s) et que le processus $\{\varepsilon_t^2\}$ est stationnaire de second ordre, alors $\{\varepsilon_t^2\}$ admet une représentation ARMA(m,s). De même, si $\{\varepsilon_t\}$ admet une représentation ARCH(s), alors $\{\varepsilon_t^2\}$ admet une représentation AR(s). Ces représentations sous la forme d'un modèle ARMA sont utiles dans l'estimation et l'identification des modèles GARCH.

Dans le cas particulier du modèle GARCH(1,1), l'équation (2.8.1) peut s'écrire:

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2. \quad (2.8.3)$$

avec $\omega > 0$, $\alpha \geq 0$ et $\beta \geq 0$. La condition de stationnarité devient donc $\alpha + \beta < 1$. Plus α est proche de 1, plus la volatilité est grande tandis que plus β est proche de 1, plus la volatilité est persistante. Dans les applications, β est souvent près de 1.

Dans ce chapitre, nous avons présenté les préliminaires qui seront utiles dans le prochain chapitre. Nous avons en effet discuté des modèles ARMA, comment les estimer et nous avons présenté l'expansion de la densité spectrale en fonction d'une base d'ondelettes. Nous avons montré comment estimer ces coefficients et nous avons énoncé la distribution asymptotique de ces quantités. Nous avons également abordé une classe de modèles dont les éléments sont des différences de martingale. Dans le prochain chapitre, nous présenterons quelques tests portemanteaux connus et nous introduirons la nouvelle méthodologie.

Chapitre 3

Statistiques de test

Comme mentionné dans l'introduction, il existe une littérature très vaste portant sur l'élaboration de statistiques de test visant à vérifier l'ajustement de modèles de séries chronologiques. L'objectif de ce chapitre est de présenter quelques méthodes parmi les plus populaires et les plus utilisées. Les méthodes présentées sont valides sous certaines conditions et reposent sur des résidus de modèles. Parmi ces dernières, on retrouve les méthodes Box-Pierce-Ljung (BPL) élaborées par Box et Pierce (1970) et Ljung et Box (1978) pour vérifier la qualité d'ajustement d'un modèle $\text{ARMA}(p,q)$. Nous abordons également les tests modifiés de Francq et al. (2005) que l'on notera tests modifiés FRZ. Après avoir introduit les méthodologies existantes, nous expliquerons en détail la nouvelle méthodologie et les nouveaux tests.

3.1. Tests de Box-Pierce-Ljung

Les statistiques de test de Box-Pierce-Ljung ont été développées sous le modèle ARMA fort. Ainsi, dans l'étude des méthodes, les termes d'erreur, ou les innovations, sont supposés des variables aléatoires indépendantes. Afin de motiver ces statistiques de test, il est utile de considérer en premier temps l'étude de la fonction d'autocorrélation empirique reposant sur le bruit non observable. Ainsi, si un modèle $\text{ARMA}(p,q)$ est correctement ajusté au processus $\{X_t\}$, alors la fonction d'autocorrélation empirique $\rho(l) = \hat{\rho}_\varepsilon(l) = r(l)/r(0)$ ne sera pas statistiquement différente de la valeur 0 pour $l > 0$.

Pour n grand et un m grand et fixé, nous avons que $\boldsymbol{\rho}_m = (\rho(1), \rho(2), \dots, \rho(m))^\top$ s'approche d'une loi normale multivariée de moyenne nulle, avec $\text{Var}\{\rho(l)\} = (n-l)/\{n(n+$

2)}, pour $l > 0$. Donc, la statistique $n(n+2) \sum_{l=1}^m \rho^2(l)/(n-l)$ converge asymptotiquement vers une distribution de χ_m^2 . Notons aussi que pour $n \gg m$, nous avons que $(n-l)/\{n(n+2)\} \approx 1/n$ et donc la statistique $n \sum_{l=1}^m \rho^2(l)$ converge aussi asymptotiquement vers une distribution de χ_m^2 . On souligne que le nombre de degrés de liberté est m , car puisque les statistiques de test reposent sur $\rho(l)$, il n'y a pas d'impact compte tenu de la phase d'estimation.

Après l'ajustement d'un modèle et en estimant les paramètres θ par $\hat{\theta}$, les résidus e_t et la variance résiduelle $\hat{\sigma}^2$, le calcul de la statistique de test précédente perd $p+q$ degrés de libertés. Ainsi, nous avons les distributions approximatives suivantes:

$$\tilde{Q}_m = n(n+2) \sum_{l=1}^m \frac{\hat{\rho}^2(l)}{n-l} \approx \chi_{m-p-q}^2,$$

$$Q_m = n \sum_{l=1}^m \hat{\rho}^2(l) \approx \chi_{m-p-q}^2,$$

où $\hat{\rho}(l) = \hat{\rho}_\varepsilon(l) = \hat{r}(l)/\hat{\sigma}^2$ est la fonction d'autocorrélation résiduelle. Il est à noter que contrairement à une croyance populaire dans la littérature des séries chronologiques, les statistiques \tilde{Q}_m et Q_m n'admettent pas des lois χ_{m-p-q}^2 asymptotiquement. Ceci a été noté par Ljung (1986). Cependant l'approximation des lois exactes de \tilde{Q}_m et Q_m par une loi χ_{m-p-q}^2 est typiquement satisfaisante en pratique. Les statistiques de test \tilde{Q}_m et Q_m sont parfois notées les statistiques de LB et BP, respectivement. La statistique de LB a la réputation d'être plus performante pour des tailles d'échantillon plus petites (voir Ljung et Box (1978) pour plus de détails). Pour indiquer ces méthodes on utilisera parfois l'acronyme BPL.

Les tests de LB peuvent être effectués en suivant les étapes que nous nous proposons de détailler.

Étape 1. Ajuster un modèle et obtenir les estimations des paramètres θ par $\hat{\theta}$, les résidus e_t et la variance résiduelle $\hat{\sigma}^2$.

Étape 2. Calculer les autocorrélations résiduelles:

$$\hat{\rho}(l) = \frac{\hat{r}(l)}{\hat{\sigma}^2} = n^{-1} \sum_{t=1}^{n-l} \frac{e_t e_{t+l}}{\hat{\sigma}^2},$$

pour $1 \leq l \leq m$.

Étape 3. Calculer les statistiques de test de LB:

$$\tilde{Q}_m = n(n+2) \sum_{l=1}^m \frac{\hat{\rho}^2(l)}{n-l}.$$

Étape 4. Rejeter l'hypothèse nulle, H_0 de (2.3.1), si $\tilde{Q}_m > \chi_{m-p-q}^2(1-\alpha)$ où $\chi_{m-p-q}^2(1-\alpha)$ est le quantile $1-\alpha$ d'une distribution χ_{m-p-q}^2 .

Il faut choisir $m > p + q$ afin de pouvoir calculer le quantile. En pratique, le calcul de la statistique doit être fait pour plusieurs valeurs de m afin de pouvoir faire un choix éclairé. Par exemple, prendre une valeur de m petite permet de vérifier la dépendance de bas ordre, alors qu'un m plus grand permet de détecter les effets saisonniers.

3.2. Tests de Francq-Roy-Zakoïan

La méthodologie de Francq et al. (2005) généralise les tests de Box-Pierce-Ljung lorsque l'innovation d'un modèle ARMA comme défini en (2.2.1) n'est plus présumée indépendante. Comme on a déjà vu, les propriétés fondamentales des équations ARMA ne reposent que sur l'hypothèse d'absence de corrélation du bruit et Francq et al. (2005) ne s'en tiennent qu'à cette hypothèse particulièrement générale. Compte tenu de l'importance des tests de Box-Pierce-Ljung en pratique, c'est la statistique qui est étudiée sous des hypothèses plus faibles et les lois des tests LB et BP sont obtenues sous l'hypothèse d'un modèle ARMA avec innovations faibles correctement ajusté, menant à ce qu'on l'on appellera les tests FRZ.

À la manière des résultats élaborés dans la section 2.4, Francq et al. (2005) obtiennent la distribution asymptotique d'un vecteur d'autocorrélations résiduelles et détaillent la structure de la matrice de covariance asymptotique lorsque le modèle est correctement spécifié. De ce résultat, la loi des tests BP et LB sont déduites, qui sont cependant passablement plus complexe que dans le cas d'un modèle ARMA fort.

Afin d'obtenir la distribution asymptotique des autocorrélations, nous devons invoquer des théorèmes limites, comme le théorème limite central, valide sous des hypothèses plus générales que le théorème limite central classique qui lui repose sur l'hypothèse d'indépendance.

Les résultats de Francq et al. (2005) font appel au concept probabiliste de mélange. Les hypothèses de mélange sont des idées remontant au probabiliste Rosenblatt (1956), et sont exposées dans Francq et Zakoïan (2009, Annexe A.3). Essentiellement, les coefficients de mélange mesurent en quelque sorte la distance entre le passé et le futur d'un processus. Il existe les coefficients de α -mélange et les coefficients de β -mélange. Le théorème limite utilisé par Francq et al. (2005) utilise les coefficients de α -mélange, appelés également coefficients de mélange fort.

On présente les idées de base des coefficients de α -mélange. Pour un processus $\{X_t, t \in \mathbb{Z}\}$, pas forcément stationnaire, on définit le coefficient $\alpha_X(h)$ ainsi:

$$\alpha_X(h) = \sup_t \alpha \{ \sigma(X_u, u \leq t), \sigma(X_u, u \geq t+h) \},$$

où $\sigma(X_u, u \leq t)$ dénote la σ -algèbre engendrée par les variables aléatoires $\{X_u, u \leq t\}$, et similairement $\sigma(X_u, u \geq t+h)$ dénote la σ -algèbre engendrée par le « futur » de X_t , c'est-à-dire les variables $\{X_u, u \geq t+h\}$, et:

$$\alpha(\mathcal{A}, \mathcal{B}) = \sup_{A \in \mathcal{A}, B \in \mathcal{B}} |P(A \cap B) - P(A)P(B)|,$$

pour deux tribus \mathcal{A} et \mathcal{B} . On note que si $\{X_t\}$ est constitué de variables aléatoires iid, alors $\alpha_X(h) = 0$. De plus, on peut apprécier que $\alpha(\mathcal{A}, \mathcal{B})$ représente une mesure quantitative à l'indépendance, puisque dans le cas d'évènements indépendants on a bien évidemment $P(A \cap B) = P(A)P(B)$. Un processus $\{X_t\}$ sera dit α -mélangeant si son coefficient de mélange $\alpha_X(h) \rightarrow 0$ lorsque $h \rightarrow \infty$. Selon Francq et Zakoïan (2009, p. 424), un processus α -mélangeant est également appelé fortement mélangeant. Comme mentionné plus haut, un intérêt de l'hypothèse qu'un processus soit α -mélangeant est que si $\{X_t\}$ satisfait certaines conditions, alors il existe un théorème limite central pour $n^{-1/2} \sum_{t=1}^n X_t$, où l'on présume que $E(X_t) = 0$. Plus précisément, le théorème de Herrndorf (1984) stipule que pour un processus $\{X_t\}$ centré, c'est-à-dire que $E(X_t) = 0$, si:

$$\sup_t \{E(X_t^{2+v})\}^{1/(2+v)} < \infty,$$

et

$$\sum_{h=0}^{\infty} \{\alpha_X(h)\}^{v/(2+v)} < \infty,$$

pour un certain $v > 0$, et si:

$$\sigma^2 = \lim_{n \rightarrow \infty} \text{Var} \left\{ n^{-1/2} \sum_{t=1}^n X_t \right\} < \infty,$$

avec $\sigma^2 > 0$, alors:

$$n^{-1/2} \sum_{t=1}^n X_t \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2).$$

Ainsi présumant un processus strictement stationnaire, si $\{X_t\}$ satisfait les équations d'un modèle ARMA (2.2.1), avec les conditions $E|X_t|^{4+2v} < \infty$ et $\sum_{k=0}^{\infty} \{\alpha_X(k)\}^{v/(2+v)} < \infty$, pour un certain $v > 0$, Francq et al. (2005) ont montré que:

$$\sqrt{n} \hat{\mathbf{r}}_m \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\hat{\rho}_m}),$$

où $\boldsymbol{\Sigma}_{\hat{\rho}_m} = \boldsymbol{\Sigma}_{\hat{\mathbf{r}}_m} / \sigma_\varepsilon^4$ et $\boldsymbol{\Sigma}_{\hat{\mathbf{r}}_m}$ représente la matrice des variances et des covariances asymptotiques des composantes du vecteur $\sqrt{n} \hat{\mathbf{r}}_m$.

De ces résultats, la distribution asymptotique de la statistique de test de LB sous un modèle ARMA faible peut être obtenue. La statistique converge en distribution vers:

$$\mathbf{Z}_m(\boldsymbol{\xi}_m) := \sum_{i=1}^m \xi_{i,m} Z_i^2,$$

où $\boldsymbol{\xi}_m = (\xi_{1,m}, \dots, \xi_{m,m})^\top$ est le vecteur de valeurs propres de la matrice $\boldsymbol{\Sigma}_{\hat{\rho}_m}$ et les Z_1, \dots, Z_m sont des variables indépendantes $\mathcal{N}(0,1)$. Donc, $\mathbf{Z}_m(\boldsymbol{\xi}_m)$ est une somme de variables de loi χ_1^2 pondérées par les poids constituant $\boldsymbol{\xi}_m$.

En pratique, nous devons estimer $\boldsymbol{\Sigma}_{\hat{\rho}_m}$ et donc estimer $\boldsymbol{\xi}_m$. Pour ce faire, Francq et al. (2005) ont suggéré une estimation convergeante de la matrice $\boldsymbol{\Gamma}_{m,m}$ par une extension du travail de Berk (1974) sur les processus multivariés linéaires. Ils ont posé $\boldsymbol{\Gamma}_{m,m} = \sum_{h=-\infty}^{+\infty} \text{Cov}(\mathbf{Y}_t, \mathbf{Y}_{t+h})$ où $\mathbf{Y}_t = \varepsilon_t(\varepsilon_{t+1}, \dots, \varepsilon_{t+m})^\top$ et ont présumé que le processus \mathbf{Y}_t admet une représentation AR(∞) de la forme $\mathcal{A}(B)\mathbf{Y}_t := \mathbf{Y}_t - \sum_{i=1}^{\infty} A_i \mathbf{Y}_{t-i} = \mathbf{u}_t$, où $\boldsymbol{\Sigma}_{u,u} := \text{Var}(\mathbf{u}_t)$ n'est pas singulière. Par l'interprétation de $(2\pi)^{-1} \boldsymbol{\Gamma}_{m,m}$ comme la densité spectrale de $\{\mathbf{Y}_t\}$ évaluée à la fréquence 0 (voir Brockwell et Davis (1991, p. 459)), ils obtiennent $\boldsymbol{\Gamma}_{m,m} = \mathcal{A}(1)^{-1} \boldsymbol{\Sigma}_{u,u} \mathcal{A}^\top(1)^{-1}$.

Dans le cadre de processus linéaires univariés avec des innovations indépendantes, Berk (1974) montre que la densité spectrale peut être estimée de manière convergeante en ajustant un modèle autorégressif d'ordre $r = r(n)$ avec $r \rightarrow \infty$ mais tel que $r^3/n \rightarrow 0$ quand

$n \rightarrow \infty$. Francq et al. (2005) ont construit la statistique $\hat{\mathbf{Y}}_t = e_t(e_{t+1}, e_{t+2}, \dots, e_{t+m})^\top$ afin d'approximer \mathbf{Y}_t .

Soit $\hat{\mathcal{A}}_r(z) = \mathbf{I}_m - \sum_{i=1}^r \hat{\mathbf{A}}_{r,i} z^i$, où $\hat{\mathbf{A}}_{r,1}, \dots, \hat{\mathbf{A}}_{r,r}$ sont les coefficients de la régression multivariée de $\hat{\mathbf{Y}}_t$ sur $\hat{\mathbf{Y}}_{t-1}, \dots, \hat{\mathbf{Y}}_{t-r}$ pour $t = 1, \dots, n - m$ (par convention, $\hat{\mathbf{Y}}_t = \mathbf{0}$, quand $t \leq 0$). La matrice $\hat{\Sigma}_{\hat{u}_r, \hat{u}_r}$ dénote l'estimateur empirique de la variance des termes d'erreur de cette régression. Les matrices $\hat{\mathbf{A}}_{r,1}, \dots, \hat{\mathbf{A}}_{r,r}$ et $\hat{\Sigma}_{\hat{u}_r, \hat{u}_r}$ peuvent être calculées récursivement par l'algorithme donné par Whittle (1963). Francq et al. (2005) ont montré que sous des hypothèses faibles sur la représentation AR(∞) et supposant que $E|\varepsilon_t|^{8+4v} < \infty$ et $\sum_{k=0}^{\infty} \{\alpha_\varepsilon(k)\}^{v/(2+v)} < \infty$, pour $v > 0$:

$$\hat{\Gamma}_{m,m} = \hat{\mathcal{A}}_r(1)^{-1} \hat{\Sigma}_{\hat{u}_r, \hat{u}_r} \hat{\mathcal{A}}_r^\top(1)^{-1} \xrightarrow{p} \Gamma_{m,m} = \mathcal{A}_r(1)^{-1} \Sigma_{u_r, u_r} \mathcal{A}_r^\top(1)^{-1}, \quad (3.2.1)$$

quand $r \rightarrow \infty$ mais tel que $r^3/n \rightarrow 0$ quand $n \rightarrow \infty$.

Il est utile de considérer également:

$$\Gamma(-j) = \sum_{k=1}^r \tilde{\mathbf{A}}_{rk} \Gamma(k-j), \quad j = 1, \dots, r.$$

Posons:

$$\begin{aligned} \Sigma_{u_r, u_r} &= \Gamma(0) - \mathbf{A}_{r1} \Gamma(-1) - \dots - \mathbf{A}_{rr} \Gamma(-r), \\ \tilde{\Sigma}_{u_r, u_r} &= \Gamma(0) - \tilde{\mathbf{A}}_{r1} \Gamma(1) - \dots - \tilde{\mathbf{A}}_{rr} \Gamma(r), \\ \Delta_r &= \Gamma(r+1) - \mathbf{A}_{r1} \Gamma(r) - \dots - \mathbf{A}_{rr} \Gamma(1), \\ \tilde{\Delta}_r &= \Gamma(-r-1) - \tilde{\mathbf{A}}_{r1} \Gamma(-r) - \dots - \tilde{\mathbf{A}}_{rr} \Gamma(-1), \end{aligned} \quad (3.2.2)$$

où $\Gamma(-l) = \Gamma^\top(l)$ et les valeurs initiales sont:

$$\begin{aligned} \Sigma_{u_0, u_0}^{-1} &= \tilde{\Sigma}_{u_0, u_0}^{-1} = \Gamma(0), \\ \Delta_0 &= \tilde{\Delta}_0^\top = \Gamma(1). \end{aligned}$$

On a alors:

$$\begin{aligned} \mathbf{A}_{rr} &= \Delta_{r-1} \tilde{\Sigma}_{u_{r-1}, u_{r-1}}^{-1}, \\ \tilde{\mathbf{A}}_{rr} &= \tilde{\Delta}_{r-1} \Sigma_{u_{r-1}, u_{r-1}}^{-1}, \\ \mathbf{A}_{rk} &= \mathbf{A}_{r-1,k} - \mathbf{A}_{rr} \tilde{\mathbf{A}}_{r-1, r-k}, \\ \tilde{\mathbf{A}}_{rk} &= \tilde{\mathbf{A}}_{r-1,k} - \tilde{\mathbf{A}}_{rr} \mathbf{A}_{r-1, r-k}, \end{aligned} \quad (3.2.3)$$

Contrairement au cas univarié, les solutions de la prédiction avant et de la prédiction arrière sont différentes et doivent être trouvées simultanément dans le cas multivarié, car $\Gamma(l) \neq \Gamma(-l)$.

En appliquant l'algorithme de Whittle sur le processus empirique $\{\hat{\mathbf{Y}}_t\}$, nous pouvons calculer $\hat{\mathcal{A}}_r(1)$ et $\hat{\Sigma}_{\hat{u}_r, \hat{u}_r}$ et ensuite estimer $\Gamma_{m,m}$ par (3.2.1). Le choix de r peut se faire par la minimisation d'un critère.

Dans Francq et al. (2005), les auteurs considèrent le critère d'information de Bayes (BIC) et le critère AIC de Akaike. Ces critères de sélection de modèles sont discutés dans Brockwell et Davis (2016). Il est bien connu qu'un critère comme le AIC a tendance à prendre des modèles avec trop de paramètres dans les modèles autorégressifs, voir Brockwell et Davis (2016, p. 152). Le critère BIC est connu comme un critère convergent. Francq et al. (2005) ont comparé les deux critères et affirment qu'ils donnent des résultats équivalents. Dans nos propres expériences empiriques présentées plus loin, nous avons obtenu de meilleurs résultats avec le critère BIC. C'est donc le critère que nous recommandons pour la suite.

Considérons un entier M tel que $M > m$ et posons aussi $\hat{\Lambda}_\infty = (\hat{\Lambda}_M, \hat{\Lambda}_{M+})$ où $\hat{\Lambda}_\infty$ est l'estimation de Λ_∞ obtenu des estimateurs des paramètres $\hat{\theta}$. Il faut calculer M_Λ tel que $\|\hat{\Lambda}_{M_\Lambda+}\| < \tau^2$ selon une tolérance τ qui est une mesure de précision, et qui est choisie relativement petite.

La matrice $\Sigma_{\hat{r}_m}$ est estimée par:

$$\begin{aligned} \hat{\Sigma}_{\hat{r}_m} = & \hat{\Gamma}_{m,m} + \hat{\Lambda}_m^\top (\hat{\Lambda}_M \hat{\Lambda}_M^\top)^{-1} \hat{\Lambda}_M \hat{\Gamma}_{M,M} \hat{\Lambda}_M^\top (\hat{\Lambda}_M \hat{\Lambda}_M^\top)^{-1} \hat{\Lambda}_m \\ & - \hat{\Lambda}_m^\top (\hat{\Lambda}_M \hat{\Lambda}_M^\top)^{-1} \hat{\Lambda}_M \hat{\Gamma}_{M,m} - \hat{\Gamma}_{m,M} \hat{\Lambda}_M^\top (\hat{\Lambda}_M \hat{\Lambda}_M^\top)^{-1} \hat{\Lambda}_m, \end{aligned} \quad (3.2.4)$$

où les matrices $\Gamma_{m,m}$, $\Gamma_{m,M}$ et $\Gamma_{M,m}$ sont des sous-matrices de $\Gamma_{M,M}$. L'estimation de la matrice $\Sigma_{\hat{\rho}_m}$ est donnée par $\hat{\Sigma}_{\hat{\rho}_m} = \hat{\Sigma}_{\hat{r}_m} / \hat{\sigma}^4$. Nous pouvons obtenir $\hat{\xi}_m$, l'estimateur de ξ_m , en calculant les valeurs propres de la matrice $\hat{\Sigma}_{\hat{\rho}_m}$. Une région critique asymptotique pour les statistiques de test \tilde{Q}_m pour les modèles ARMA faibles peut être obtenue. En fait, il peut être montré que:

$$\lim_{n \rightarrow \infty} P \left\{ \tilde{Q}_m > z_m(1 - \alpha) \right\} = \alpha,$$

où $\mathbf{z}_m(1 - \alpha)$ est tel que $P\{\mathbf{Z}_m(\hat{\boldsymbol{\xi}}_m) > \mathbf{z}_m(1 - \alpha)\} = \alpha$.

L'évaluation des points critiques de $\mathbf{z}_m(1 - \alpha)$ peut être effectuée avec l'algorithme d'Imhof (voir Imhof (1961)).

Pour effectuer les tests FRZ, il faut effectuer les trois premières étapes des tests de Box-Pierce-Ljung. Elles sont complétées avec les étapes suivantes:

Étape 4. Calculer $\hat{\boldsymbol{\Lambda}}_\infty$ à partir des paramètres estimés $\hat{\boldsymbol{\theta}}$, choisir une valeur de τ appropriée, calculer $M = \max(m, M_\Lambda)$ et obtenir $\hat{\boldsymbol{\Lambda}}_m$ et $\hat{\boldsymbol{\Lambda}}_M$.

Étape 5. Choisir r_{\max} , effectuer l'ajustement du processus multivarié $\{\hat{\mathbf{Y}}_t\}$ à l'aide de l'algorithme de Whittle pour $r \in \{0, 1, \dots, r_{\max}\}$, utiliser le critère BIC afin de sélectionner $r_0 \in \{0, 1, \dots, r_{\max}\}$, obtenir $\hat{\mathbf{A}}_{r_0}(1)$ et $\hat{\boldsymbol{\Sigma}}_{\hat{u}_{r_0}, \hat{u}_{r_0}}$ et calculer $\hat{\boldsymbol{\Gamma}}_{M, M}$.

Étape 6. Obtenir $\hat{\boldsymbol{\Gamma}}_{m, M}$, $\hat{\boldsymbol{\Gamma}}_{M, m}$ et $\hat{\boldsymbol{\Gamma}}_{m, m}$ à partir de $\hat{\boldsymbol{\Gamma}}_{M, M}$, calculer $\hat{\boldsymbol{\Sigma}}_{\hat{r}_m}$ à partir de (3.2.4) et calculer $\hat{\boldsymbol{\Sigma}}_{\hat{\rho}_m} = \hat{\boldsymbol{\Sigma}}_{\hat{r}_m} / \hat{\sigma}^4$.

Étape 7. Calculer les valeurs propres de $\hat{\boldsymbol{\Sigma}}_{\hat{\rho}_m}$ pour obtenir $\hat{\boldsymbol{\xi}}_m$ et calculer le point critique de $\mathbf{z}_m(1 - \alpha)$ en utilisant l'algorithme de Imhof.

Étape 8. Rejeter l'hypothèse nulle, H_0 de (2.3.1), si $P\{\tilde{Q}_m > \mathbf{z}_m(1 - \alpha)\} < \alpha$.

On note que la statistique \tilde{Q}_m repose sur le choix de m . On peut interpréter m comme le niveau de dépendance maximum que la statistique \tilde{Q}_m peut détecter avec les mesures de corrélation. On remarque que m a été considéré fixé dans les résultats asymptotiques. Habituellement, les valeurs de m petites permettent de mesurer la dépendance de bas ordre, alors que les valeurs m plus grandes permettent de tenir compte d'éventuels effets saisonniers. Dans la version classique des modèles ARMA(p, q), m doit être tel que $m > p + q$ puisque la loi limite est χ_{m-p-q}^2 . Ainsi, si m est trop petit, la loi limite n'est pas définie par une variable aléatoire χ^2 .

Un sérieux avantage de la méthodologie de Francq et al. (2005) est qu'une telle restriction sur m n'existe pas, et la méthodologie est valide quel que soit $m > 0$. Évidemment, le prix de cette généralité est que la loi limite est plus complexe.

On termine cette section en rappelant qu'un avantage indéniable des statistiques de FRZ est qu'elles sont applicables sur les résidus des modèles ARMA. La prochaine section introduit les nouveaux tests reposant sur les coefficients d'ondelettes.

3.3. Nouveaux tests reposant sur les ondelettes

On rappelle que l'hypothèse de départ est telle que formulée dans l'expression (2.3.1). Afin de vérifier (2.3.1), on teste l'absence de corrélation dans le terme d'erreur, ce qui revient à vérifier si la densité spectrale est constante ou non. Cependant, utilisant une expansion avec une base d'ondelettes, on peut reformuler à nouveaux les hypothèses en (2.3.1), par des tests sur les coefficients d'ondelettes comme détaillés dans les expressions (2.6.7).

On remarque cependant que le nombre de coefficients d'ondelettes sous H_0 est infini. Puisque la taille de la série chronologique n est finie, il faudra considérer un nombre fini de coefficients d'ondelettes, et il est également anticipé que ce nombre ne pourra pas être trop grand par rapport à n . Heureusement, la capacité d'adaptation des méthodes d'ondelettes est telle que peu de coefficients sont habituellement nécessaires dans la mise en oeuvre de ces méthodes. Ceci représente en fait une motivation de développer de nouveaux tests suivant l'approche des méthodes d'ondelettes.

Soit donc un entier J , considéré fixé, et satisfaisant la relation:

$$1 \leq J \leq J_n,$$

où l'on rappelle que J_n est tel que:

$$2^{J_n+1} \leq n,$$

voir aussi la section 2.6. On regroupe les coefficients d'ondelettes théoriques et empiriques dans des vecteurs qui sont dénotés par:

$$\begin{aligned} \mathbf{V}_J &= (\beta_{01}, \beta_{11}, \beta_{21}, \beta_{22}, \beta_{31}, \dots, \beta_{34}, \beta_{41}, \dots, \beta_{48}, \dots, \beta_{J,1}, \dots, \beta_{J,2^{J-1}})^\top, \\ \hat{\mathbf{V}}_J &= (\hat{\beta}_{01}, \hat{\beta}_{11}, \hat{\beta}_{21}, \hat{\beta}_{22}, \hat{\beta}_{31}, \dots, \hat{\beta}_{34}, \hat{\beta}_{41}, \dots, \hat{\beta}_{48}, \dots, \hat{\beta}_{J,1}, \dots, \hat{\beta}_{J,2^{J-1}})^\top, \end{aligned} \quad (3.3.1)$$

respectivement. Les vecteurs \mathbf{V}_J et $\hat{\mathbf{V}}_J$ ont chacun $1 + \sum_{j=1}^J 2^{j-1} = 1 + (2^J - 1)/(2 - 1) = 2^J$ éléments. On rappelle que sous l'hypothèse nulle H_0 telle que formulée en (2.6.7), le vecteur \mathbf{V}_J se réduit au vecteur nul. Afin d'étudier la distribution asymptotique de $\hat{\mathbf{V}}_J$ sous H_0 , on fait appel au théorème de Cramér-Wold, tel qu'énoncé dans Serfling (1980). On montre que toutes les combinaisons linéaires de $\hat{\mathbf{V}}_J$ sont asymptotiquement normales. Plus précisément, pour un vecteur arbitraire \mathbf{a} fixe de dimension $2^J \times 1$, il peut être

montré en adoptant une argumentation largement similaire à celle utilisée pour montrer la normalité asymptotique des $\hat{\beta}_{jk}$ que $\mathbf{a}^\top \hat{\mathbf{V}}_J$ est de loi normale. Ainsi, on conclut que $\hat{\mathbf{V}}_J$ est asymptotiquement une distribution normale multivariée. On dénote $\Sigma_{\hat{\mathbf{V}}}$ la matrice des variances et des covariances asymptotiques de $n^{1/2} \hat{\mathbf{V}}_J$. Ainsi, comme montré dans le théorème 2.7.1, on trouve la loi asymptotique de $n^{1/2} \hat{\mathbf{V}}_J$. Ceci fait l'objet du théorème 3.3.1 qui est maintenant énoncé.

Théorème 3.3.1. *Supposons que $\{X_t\}$ suit un modèle ARMA (2.2.1) et que l'innovation $\{\varepsilon_t\}$ satisfait l'hypothèse A. Soit J ($1 \leq J < J_n$) un entier fixé et présumons que l'ondelette de Franklin η soit adoptée. Alors nous avons sous l'hypothèse nulle H_0 que:*

$$\sqrt{n} \hat{\mathbf{V}}_J \longrightarrow_d \mathcal{N}_{2^J}(\mathbf{0}, \Sigma_{\hat{\mathbf{V}}}),$$

où la matrice $2^J \times 2^J$ des variances et des covariances $\Sigma_{\hat{\mathbf{V}}}$ est telle que:

$$\Sigma_{\hat{\mathbf{V}}} = \begin{pmatrix} \sigma_{01}^2 & \sigma_{01,11} & \sigma_{01,21} & \sigma_{01,22} & \sigma_{01,31} & \cdots & \sigma_{01,J,2^{J-1}} \\ \sigma_{01,11} & \sigma_{11}^2 & \sigma_{11,21} & \sigma_{11,22} & \sigma_{11,31} & \cdots & \sigma_{11,J,2^{J-1}} \\ \sigma_{01,21} & \sigma_{11,21} & \sigma_{21}^2 & \sigma_{21,22} & \sigma_{21,31} & \cdots & \sigma_{21,J,2^{J-1}} \\ \sigma_{01,22} & \sigma_{11,22} & \sigma_{21,22} & \sigma_{22}^2 & \sigma_{22,31} & \cdots & \sigma_{22,J,2^{J-1}} \\ \sigma_{01,31} & \sigma_{11,31} & \sigma_{21,31} & \sigma_{22,31} & \sigma_{31}^2 & \cdots & \sigma_{31,J,2^{J-1}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \ddots & \cdots \\ \sigma_{01,J,2^{J-1}} & \sigma_{11,J,2^{J-1}} & \sigma_{21,J,2^{J-1}} & \sigma_{22,J,2^{J-1}} & \sigma_{31,J,2^{J-1}} & \cdots & \sigma_{J,2^{J-1}}^2 \end{pmatrix}. \quad (3.3.2)$$

et ses éléments sont définis dans le théorème 2.7.1.

La statistique de test utilise le résultat énoncé dans le théorème 3.3.1. Plus précisément, supposant que $\Sigma_{\hat{\mathbf{V}}}$ soit inversible, on peut construire la forme quadratique $n \hat{\mathbf{V}}_J^\top \Sigma_{\hat{\mathbf{V}}}^{-1} \hat{\mathbf{V}}_J$. Cependant, la matrice $\Sigma_{\hat{\mathbf{V}}}$ n'est pas connue, et il faut dans un premier temps l'estimer. Heureusement, en étudiant attentivement les expressions (2.7.1) provenant du théorème 2.7.1, on peut proposer un estimateur convergeant de la matrice $\Sigma_{\hat{\mathbf{V}}}$ puisque σ_{jk}^2 et $\sigma_{j_1 k_1, j_2 k_2}$ font intervenir les expressions $\Gamma(l, l')$ telles que définies dans le lemme 2.4.1, voir aussi

l'hypothèse A (v). De plus, il faut estimer σ^2 . Des estimateurs naturels sont:

$$\hat{\Gamma}(l, l') = \frac{1}{n} \sum_{t=1}^n \hat{\varepsilon}_t^2 \hat{\varepsilon}_{t-l} \hat{\varepsilon}_{t-l'}, \quad (3.3.3)$$

$$\hat{\sigma}^2 = \hat{c}_{\varepsilon}(0) = \frac{1}{n} \sum_{t=1}^n \hat{\varepsilon}_t^2. \quad (3.3.4)$$

Ainsi, les expressions σ_{jk}^2 et $\sigma_{j_1 k_1, j_2 k_2}$ sont estimées par $\hat{\sigma}_{jk}^2$ et $\hat{\sigma}_{j_1 k_1, j_2 k_2}$ respectivement, en prenant soin de remplacer $\Gamma(l, l')$ et σ^2 par $\hat{\Gamma}(l, l')$ et $\hat{\sigma}^2$, tels que définis en (3.3.3) et (3.3.4) respectivement. On montre que l'estimateur $\hat{\Sigma}_{\hat{\mathcal{V}}}$ ainsi obtenu est convergent pour $\Sigma_{\hat{\mathcal{V}}}$ sous l'hypothèse nulle H_0 . Ceci fait l'objet du résultat suivant:

Théorème 3.3.2. *Supposons que $\{X_t\}$ suit un modèle ARMA (2.2.1) et l'innovation $\{\varepsilon_t\}$ satisfait l'hypothèse A. Alors l'estimateur $\hat{\Sigma}_{\hat{\mathcal{V}}}$ défini dans (3.3.3) est convergent pour $\Sigma_{\hat{\mathcal{V}}}$ défini dans (3.3.2), c'est-à-dire:*

$$\hat{\Sigma}_{\hat{\mathcal{V}}} \longrightarrow_p \Sigma_{\hat{\mathcal{V}}},$$

quand $n \rightarrow \infty$.

Maintenant, du théorème 3.3.1, nous proposons la statistique de test suivante:

$$W_J = n \hat{\mathbf{V}}_J^\top \hat{\Sigma}_{\hat{\mathcal{V}}}^{-1} \hat{\mathbf{V}}_J, \quad (3.3.5)$$

où J est un entier fixé déterminé par l'analyste.

Des théorèmes 2.7.1, 3.3.1, 3.3.2 et en appliquant le théorème de Slutsky, nous obtenons le résultat suivant:

Théorème 3.3.3. *Supposons que $\{X_t\}$ suit un modèle ARMA (2.2.1) et ses innovations $\{\varepsilon_t\}$ satisfont l'hypothèse A. Soit J ($1 \leq J < J_n$) un entier fixé et l'ondelette de Franklin η est adoptée. Alors nous avons:*

$$W_J \longrightarrow_d \chi_{2J}^2, \quad (3.3.6)$$

quand $n \rightarrow \infty$.

Dans les tests de type BPL ou FRZ, il faut choisir m . Dans notre méthodologie, il faut choisir J . En pratique, nous pouvons considérer des statistiques de test avec des petites valeurs de J telles que 1, 2 ou 3. Pour les densités spectrales irrégulières avec des pics, de la saisonnalité, etc, il existe des coefficients d'ondelettes non-négligeable avec de grands niveaux de résolution j . Dans ces cas, il faut choisir un J relativement grand

pour une meilleure approximation et une meilleure puissance de test. Des travaux sont nécessaires sur la manière de sélectionner J objectivement, telles que AIC, BIC, entre autres, et valider ces techniques dans notre contexte.

Nous résumons maintenant les étapes nécessaires à la mise en oeuvre du test. Comme à la section précédente, considérons un entier M et $\mathbf{B}_{jk,\infty} = (\mathbf{B}_{jk,M}^\top, \mathbf{B}_{jk,M+}^\top)^\top$ pour $1 \leq j \leq J$ et $1 \leq k \leq 2^{j-1}$. Il faut calculer $M_{\mathbf{B}_{jk}}$ tel que $\|\mathbf{B}_{jk,M_{\mathbf{B}_{jk}+}}\| < \tau^2$ et ce pour chaque combinaison de j et de k . En choisissant M égal au plus grand $M_{\mathbf{B}_{jk}}$, $1 \leq j \leq J$ et $1 \leq k \leq 2^{j-1}$, nous pouvons reformuler les estimateurs $\hat{\sigma}_{jk}^2$ et $\hat{\sigma}_{j_1 k_1, j_2 k_2}$ du théorème 2.7.1 en remplaçant les symboles ∞ par M .

Pour effectuer les nouveaux tests, il faut effectuer les étapes suivantes:

Étape 1. Ajuster un modèle de type ARMA et obtenir les estimations des paramètres $\boldsymbol{\theta}$ par $\hat{\boldsymbol{\theta}}$, les résidus e_t et la variance résiduelle $\hat{\sigma}^2$.

Étape 2. Calculer $\hat{\boldsymbol{\Lambda}}_\infty$ à partir des paramètres estimés $\hat{\boldsymbol{\theta}}$.

Étape 3. Calculer les $b_{jk}(h)$ afin d'obtenir les $\mathbf{B}_{jk,\infty}$, pour $1 \leq j \leq J$, $1 \leq k \leq 2^{j-1}$ et $h \geq 1$.

Étape 4. Choisir une valeur de τ appropriée et calculer $M_\mathbf{A}$ et les $M_{\mathbf{B}_{jk}}$.

Étape 5. Calculer les autocovariances résiduelles $\hat{r}(l) = n^{-1} \sum_{t=1}^{n-l} e_t e_{t+l}$, pour $1 \leq l \leq n-1$.

Étape 6. Calculer les $\hat{\beta}_{jk}$ (2.6.13) afin d'obtenir $\hat{\mathbf{V}}_J$.

Étape 7. Déterminer la constante M_{\max} suivante:

$$M_{\max} = \max(M_\mathbf{A}, M_{\mathbf{B}_{11}}, \dots, M_{\mathbf{B}_{J,2^{J-1}}}),$$

et obtenir $\hat{\boldsymbol{\Lambda}}_{M_{\max}}$ et $\mathbf{B}_{jk,M_{\max}}$ pour chaque combinaison de j et de k afin de calculer $\hat{\boldsymbol{\Gamma}}_{M_{\max}, M_{\max}}$.

Étape 8. Calculer chaque élément de $\hat{\boldsymbol{\Sigma}}_{\hat{\mathbf{V}}}$ avec:

$$\begin{aligned} \hat{\sigma}_{jk}^2 &= \mathbf{B}_{jk,M}^\top \hat{\boldsymbol{\Gamma}}_{M,M} \mathbf{B}_{jk,M} \\ &+ \mathbf{B}_{jk,M}^\top \hat{\boldsymbol{\Lambda}}_M^\top (\hat{\boldsymbol{\Lambda}}_M \hat{\boldsymbol{\Lambda}}_M^\top)^{-1} \hat{\boldsymbol{\Lambda}}_M \hat{\boldsymbol{\Gamma}}_{M,M} \hat{\boldsymbol{\Lambda}}_M^\top (\hat{\boldsymbol{\Lambda}}_M \hat{\boldsymbol{\Lambda}}_M^\top)^{-1} \hat{\boldsymbol{\Lambda}}_M \mathbf{B}_{jk,M} \\ &- \mathbf{B}_{jk,M}^\top \hat{\boldsymbol{\Lambda}}_M^\top (\hat{\boldsymbol{\Lambda}}_M \hat{\boldsymbol{\Lambda}}_M^\top)^{-1} \hat{\boldsymbol{\Lambda}}_M \hat{\boldsymbol{\Gamma}}_{M,M} \mathbf{B}_{jk,M} \\ &- \mathbf{B}_{jk,M}^\top \hat{\boldsymbol{\Gamma}}_{M,M} \hat{\boldsymbol{\Lambda}}_M^\top (\hat{\boldsymbol{\Lambda}}_M \hat{\boldsymbol{\Lambda}}_M^\top)^{-1} \hat{\boldsymbol{\Lambda}}_M \mathbf{B}_{jk,M}, \end{aligned}$$

$$\begin{aligned}
\hat{\sigma}_{j_1 k_1, j_2 k_2} &= \mathbf{B}_{j_1 k_1, M}^\top \hat{\mathbf{\Gamma}}_{M, M} \mathbf{B}_{j_2 k_2, M} \\
&+ \mathbf{B}_{j_1 k_1, M}^\top \hat{\mathbf{\Lambda}}_M^\top (\hat{\mathbf{\Lambda}}_M \hat{\mathbf{\Lambda}}_M^\top)^{-1} \hat{\mathbf{\Lambda}}_M \hat{\mathbf{\Gamma}}_{M, M} \hat{\mathbf{\Lambda}}_M^\top (\hat{\mathbf{\Lambda}}_M \hat{\mathbf{\Lambda}}_M^\top)^{-1} \hat{\mathbf{\Lambda}}_M \mathbf{B}_{j_2 k_2, M} \\
&- \mathbf{B}_{j_1 k_1, M}^\top \hat{\mathbf{\Lambda}}_M^\top (\hat{\mathbf{\Lambda}}_M \hat{\mathbf{\Lambda}}_M^\top)^{-1} \hat{\mathbf{\Lambda}}_M \hat{\mathbf{\Gamma}}_{M, M} \mathbf{B}_{j_2 k_2, M} \\
&- \mathbf{B}_{j_1 k_1, M}^\top \hat{\mathbf{\Gamma}}_{M, M} \hat{\mathbf{\Lambda}}_M^\top (\hat{\mathbf{\Lambda}}_M \hat{\mathbf{\Lambda}}_M^\top)^{-1} \hat{\mathbf{\Lambda}}_M \mathbf{B}_{j_2 k_2, M}.
\end{aligned}$$

Étape 9. Calculer les nouvelles statistiques de test $W_J = n \hat{\mathbf{V}}_J^\top \hat{\mathbf{\Sigma}}_{\hat{\mathbf{V}}}^{-1} \hat{\mathbf{V}}_J$.

Étape 10. Rejeter l'hypothèse nulle, H_0 de (2.3.1), si $W_J > \chi_{2J}^2(1 - \alpha)$, où $\chi_{2J}^2(1 - \alpha)$ est le quantile $1 - \alpha$ d'une distribution χ_{2J}^2 .

Dans les applications, le nombre de valeurs de J à calculer devrait être petit. Notons également qu'en pratique J ne peut pas être trop grand, compte tenu de la contrainte $2^{J+1} \leq n$. Il est donc recommandé de calculer la statistique pour $J = 1, 2$, par exemple. Une étude de simulation explore les propriétés de niveaux et de puissances du nouveau test W_J tel que défini par (3.3.5).

Chapitre 4

Expériences empiriques

Au chapitre précédent, de nouvelles statistiques de test à base d'ondelettes ont été proposées afin de diagnostiquer un modèle ARMA lorsque les erreurs sont non-corrélées, mais possiblement dépendantes. Comparativement aux tests existants dans ce contexte, les nouveaux tests devraient posséder des propriétés avantageuses de puissance lorsque la densité spectrale du terme d'erreur affiche un comportement plus irrégulier et non homogène. Dans le théorème 3.3.3, la loi asymptotique de la nouvelle statistique de test a été établie sous l'hypothèse nulle. D'un point de vue pratique, il est pertinent d'étudier les propriétés empiriques des nouveaux tests en terme des niveaux empiriques, qui peuvent être comparés avec ceux des compétiteurs. De plus, il semble particulièrement pertinent d'étudier les puissances exactes avec des simulations de Monte Carlo. Ces expériences empiriques sont abordées dans ce chapitre.

De manière générale, les comparaisons des nouveaux tests avec le test standard de Ljung-Box et les méthodes de Francq et al. (2005) semblent particulièrement pertinentes. Afin d'étudier empiriquement les points soulevés précédemment, nous avons étudié les niveaux exacts et les puissances empiriques des tests statistiques pour plusieurs tailles échantillonales. Plus précisément, les tailles $n = 500, 1000$ ont été considérées. On note que ces tailles sont assez importantes, mais compte tenu des applications économiques et financières, elles demeurent réalistes. Par exemple, au prochain chapitre, notre illustration avec des données réelles a une taille de plus de 7000. De plus, la détection de phénomènes non-corrélés mais dépendants nécessite des tailles de séries chronologiques assez grandes.

Dans les expériences décrites à partir de la prochaine section, nous avons considéré les niveaux nominaux de 5% et de 10%.

Pour l'étude de niveau, nous avons calculé la fréquence empirique du rejet de l'hypothèse nulle lorsqu'elle était vraie. De ces expériences, des points critiques empiriques ont pu être calculés. Ainsi, ces points critiques ont pu être utilisés dans l'étude de puissance, afin que les niveaux soient corrigés. Cette pratique est particulièrement pertinente lorsque les niveaux empiriques sont plus ou moins près des niveaux nominaux, ce qui peut arriver dans notre contexte. En effet, la convergence des tests portemanteaux utilisant des méthodes spectrales peut être plus ou moins lente, et donc pour n relativement petit il peut être pertinent de corriger pour les niveaux.

Toutes les simulations ont été faites avec Rstudio, version 1.1.453. Avant chaque groupe d'expériences, le germe de R a été initialisé à 999 en incluant l'instruction `set.seed(999)`.

4.1. Étude des niveaux empiriques

Les modèles de base pour les processus $\{X_t\}$ sont les processus autorégressifs suivants:

$$AR(1) : (1 - \phi B)X_t = \varepsilon_t, \quad (4.1.1)$$

$$AR(2) : (1 - \phi B^2)X_t = \varepsilon_t, \quad (4.1.2)$$

où $\phi \in \{0,1; 0,2; 0,3\}$ et $\{\varepsilon_t\}$ est de la forme:

$$\varepsilon_t = \sigma_t \eta_t,$$

avec σ_t^2 pouvant s'exprimer comme un GARCH(1,1):

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2. \quad (4.1.3)$$

Dans toutes les expériences, $\omega = 1$. Cependant, plusieurs combinaisons de (α, β) ont été considérées. Le processus $\{\eta_t\}$ est gaussien, constitué de variables normales indépendantes, c'est-à-dire que $\eta_t \sim \mathcal{N}(0,1)$. Les valeurs de α et de β sont telles que $\alpha \geq 0$, $\beta \geq 0$, mais avec $\alpha + \beta < 1$. Intuitivement, une grande valeur de β est associée à une volatilité σ_t^2 persistente, alors que quand α est plus grand c'est l'effet contraire.

Notons que nous avons considéré le cas $\alpha = \beta = 0$, correspondant à la situation où les erreurs $\{\varepsilon_t\}$ sont iid: dans un tel cas, les tests de Ljung-Box classiques utilisant les

points critiques d'une distribution χ^2 sont justifiés. Dans nos expériences, nous avons étudié plusieurs combinaisons de (α, β) , mais afin de limiter la quantité de tableaux et puisque plusieurs résultats étaient finalement assez similaires, nous nous sommes limités à présenter les cas $(\alpha, \beta) = (0, 6; 0, 1)$, $(0, 3; 0, 1)$, $(0, 1; 0, 3)$ et $(0, 1; 0, 6)$.

Notons que les innovations GARCH respectent les conditions d'hétéroscédasticité conditionnelle, et sont également associées à des propriétés d'irrégularité (les grandes valeurs de α présentent des pics localement importants dans la volatilité, alors que lorsque β est plus grand, ces pics semblent plus persistants). Nous avons également généré des innovations ARCH(1), ARCH(4) et même ARCH(12). Bien que la dépendance puisse afficher une certaine forme de saisonnalité, l'irrégularité dans la volatilité n'avantage pas les nouveaux tests par rapport aux tests existants. Puisque les tests performant de façon plutôt similaire, les résultats ne sont pas présentés ici et nous nous limitons au GARCH(1,1).

Les statistiques de test \tilde{Q}_m et Q_m , telles que décrites dans la section 3.1, ont été incluses dans nos expériences empiriques. On note que l'on doit choisir m . Nous avons considéré $m = 2, 3, 4, 6, 12, 24$ et $m_{\max} = 36$. L'inclusion de petites valeurs de m permet de se concentrer sur de la dépendance de bas ordre, c'est-à-dire qui survient pour de petits délais. Cependant, un choix de m trop petit peut manquer de la dépendance saisonnière, et il est recommandé en pratique d'inclure également des valeurs de m plus grandes. Plus précisément, un test avec un m trop petit ne sera pas convergent pour des formes de dépendance plus complexes, qui sont détectées avec davantage de délais. Nous avons utilisé l'approximation de la loi de \tilde{Q}_m par une loi χ^2_{m-p-q} , et nous avons également, pour les mêmes valeurs de m , considéré les tests de FRZ reposant sur \tilde{Q}_m . Les points critiques des tests de Francq et al. (2005) ont été calculés en utilisant la fonction `Imhof()` de la librairie `CompQuadForm` de R. Voir Duchesne et Lafaye De Micheaux (2010) pour la description de cette librairie.

Pour le nouveau test reposant sur les ondelettes, nous avons posé $J = 1, 2, 3$. On remarque que le nombre de coefficients d'ondelettes augmente assez rapidement avec J , quoique 2^J pour les valeurs considérées demeure petit pour les tailles simulées. Pour les modèles présentés dans notre étude, prendre $J \leq 3$ était suffisant quoiqu'en pratique il pourrait être envisagé de prendre J plus grand.

Nous avons utilisé $\tau^2 = 0.00001$ afin de calculer $M_{\mathbf{\Lambda}}$ et $M_{B_{jk}}$, $1 \leq j \leq J$ et $1 \leq k \leq 2^{j-1}$. On note que les coefficients d'ondelettes $\hat{\beta}_{jk}$ définis par (2.6.11) reposent sur les coefficients $b_{jk}(h)$, définis par la formule (2.6.12). Ces coefficients $b_{jk}(h)$ interviennent également dans les calculs des σ_{jk}^2 et $\sigma_{j_1k_1, j_2k_2}$, voir les expressions définies dans le théorème 2.7.1. On note que ces dernières expressions font intervenir des sommes infinies et il n'existe pas des expressions explicites pour ces formules, même dans les cas simples. Donc il faut considérer des points de troncation dans les sommes. Nous avons vérifié que les sommes faisant intervenir les $b_{jk}(h)$ convergeaient plutôt lentement. Ainsi, si la valeur de τ n'est pas assez petite, les niveaux empiriques des nouveaux tests peuvent être éloignés des niveaux nominaux. Ceci n'est pas une défaillance due à l'approximation en échantillons finis de la loi asymptotique, mais bien une considération numérique dans le calcul des σ_{jk}^2 et $\sigma_{j_1k_1, j_2k_2}$. En procédant de façon ad hoc, nous avons déterminé que la valeur de $\tau^2 = 0,00001$ était suffisamment petite pour que les résultats ne soient pas sensibles à l'effet de troncation. Ainsi, les valeurs des $M_{B_{jk}}$ peuvent prendre des valeurs assez élevées. L'approximation asymptotique nécessite tout de même une taille n assez élevée.

Nous fournissons les étapes détaillées afin d'effectuer l'étude de niveau :

Étape 1. Génération des séries chronologiques selon les modèles (4.1.1) ou (4.1.2) avec erreurs GARCH définies par (4.1.3). Afin de limiter l'impact des valeurs initiales, $1000 + n$ valeurs ont été générées et les n dernières constituaient la série chronologique à ajuster par un modèle ARMA.

Étape 2. Ajustement d'un modèle ARMA(p, q), en utilisant les bons ordres, c'est-à-dire $p = 1$, ou $p = 2$, et en posant $q = 0$. Compte tenu des modèles en cause, la méthode des moindres carrés a été utilisée.

Étape 3. Avec les estimateurs $\hat{\boldsymbol{\theta}}$ et $\hat{\sigma}^2$ de l'étape 2, les résidus sont calculés et les autocorrélations résiduelles déterminées.

Étape 4. Effectuer les tests standard LB et les tests FRZ. Pour cette étape, nous avons choisi $M = \max(M_{\mathbf{\Lambda}}, m_{\max})$ et ensuite $M = \min(M, m_{\max} + 20)$ et avons choisi le critère de sélection BIC avec $r_{\max} = 5$ pour la sélection de r_0 . La raison de choisir M de cette manière est d'éviter que $M_{\mathbf{\Lambda}}$ soit trop grand et que le temps d'exécution soit trop long.

Ceci ne pose pas de problèmes puisque les valeurs dans $\hat{\Lambda}_\infty$ deviennent très petites très rapidement.

Étape 5. Effectuer les nouveaux tests. Pour cette étape, nous avons choisi $M_{\max} = \max(M_{\mathbf{A}}, M_{\mathbf{B}_{11}}, \dots, M_{\mathbf{B}_{J,2^{J-1}}})$.

Étape 6. Déterminer les taux de rejet et les points critiques empiriques. Dans les études Monte Carlo, nous avons considéré un total de $N = 5000$ répliques. Ainsi, pour ces expériences Monte Carlo, les erreurs standards associées au niveau nominal des niveaux empiriques sont données par $\sqrt{\alpha_{\text{niv}}(1 - \alpha_{\text{niv}})/N}$, pour les niveaux nominaux α_{niv} de 5% et 10%, respectivement.

Nous décrivons maintenant comment nous avons procédé afin d'effectuer l'étude de puissance. Dans tous les cas, nous avons considéré des situations où les modèles postulés par (4.1.1) et (4.1.2) étaient en fait des modèles autorégressifs multiplicatifs de la forme:

$$\text{AR}(1) \times \text{AR}(4) : (1 - \phi_1 B)(1 - \phi_4 B)X_t = \varepsilon_t, \quad (4.1.4)$$

$$\text{AR}(2) \times \text{AR}(4) : (1 - \phi_2 B^2)(1 - \phi_4 B)X_t = \varepsilon_t, \quad (4.1.5)$$

où ϕ_1 et ϕ_4 appartenait à l'ensemble $\{0,1; 0,2; 0,3\}$. Plus précisément, nous avons ajusté des modèles AR(1) lorsqu'en fait le modèle (4.1.4) était généré, et de manière similaire, des modèles AR(2) lorsqu'en fait le processus sous-jacent était (4.1.5). On note que les modèles (4.1.4) et (4.1.5) sont utiles dans la modélisation de phénomènes saisonniers et que les densités spectrales des processus (4.1.4) et (4.1.5) affichent des comportements plus irréguliers. On note que lorsque $\phi_4 = 0$ on retombe sur les modèles de l'étude de niveau. Nous présentons les résultats de puissance pour les combinaisons suivantes: $(\phi_1, \phi_4), (\phi_2, \phi_4) \in \{(0,2; 0,1); (0,3; 0,1); (0,3; 0,2)\}$, pour les cas (α, β) de l'étude de niveau. Nous avons aussi inclus $(\alpha, \beta) = (0, 0)$ correspondant à des innovations iid. Nous avons étudié empiriquement un nombre de cas beaucoup plus grand, mais les résultats que nous présentons sont représentatifs. De plus, lorsque ϕ_4 était trop grand, les taux de rejet étaient trop élevés et ne permettaient pas de bien distinguer les performances des tests. On note que les taux de rejet sont calculés après l'application des corrections pour le niveau.

4.2. Discussion de l'étude sur le niveau exact

Les résultats de l'étude de niveau pour un processus respectant un modèle AR(1) avec innovations GARCH(1,1) sont présentés dans les tableaux 4.1 et 4.2.

Dans le tableau 4.1, $\beta = 0,1$ dans l'équation GARCH et α varie de 0,3 à 0,6. Si on regarde les résultats du test classique LB, on constate que les taux de rejet sont pratiquement systématiquement à l'extérieur des bandes de signification pour $\alpha = 0,3$, et que de sévères sur-rejets surviennent pour $\alpha = 0,6$, avec des taux de rejet dépassant souvent les 20% alors que le niveau nominal était de 5%. Ceci suggère que la présence de non-corrélation mais de dépendance invalide sévèrement les points critiques usuels. Ces résultats confirment les expériences empiriques de Francq et al. (2005). Dans le tableau 4.2, les tests LB continuent d'afficher certains sur-rejets mais α petit et β grand semblent moins problématiques. On remarque que les taux de rejet augmentent souvent en fonction de n , et s'améliorent à mesure que m augmente.

Si on étudie la performance des tests FRZ, on apprécie qu'ils se comportent bien dans la plupart des situations. Cependant, à mesure que m augmente, on note que les taux de rejet diminuent de sorte qu'une légère sous-estimation des niveaux nominaux est observée. Ceci est particulièrement vrai pour $m = 12, 24, 36$, ce qui peut être considéré décevant, dans la mesure où ces valeurs de m sont pertinentes dans la détection d'effets saisonniers. Cependant, il demeure pertinent de noter que le choix des coefficients α et β importe peu, et les conclusions sont essentiellement les mêmes quel que soit (α, β) .

Pour les nouveaux tests, de manière générale, les taux de rejet pour $J = 1, 2$ sont dans les bandes de signification, avec une légère détérioration pour $\phi = 0,3$. Pour ces deux valeurs de J , que ce soit pour les niveaux 5% ou 10%, les taux de rejet sont dans les bandes de signification ou très proches. Pour $J = 3$, les résultats sont moins bons, particulièrement pour $\phi = 0,2$ et $0,3$.

Les résultats de l'étude de niveau pour un processus respectant un modèle AR(2) avec l'innovation d'un modèle GARCH(1,1) sont présentés dans les tableaux 4.3 et 4.4. Les résultats des tests LB et FRZ ont essentiellement le même comportement que pour l'autorégression AR(1). Les nouveaux tests semblent bien performer pour $J = 1$, et les

Tableau 4.1. Niveaux empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(1) avec des innovations GARCH(1,1), $\phi \in \{0,1; 0,2; 0,3\}$, $\alpha \in \{0,3; 0,6\}$, $\beta = 0,1$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$.

		$\phi = 0,1$				$\phi = 0,2$				$\phi = 0,3$			
		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
$\alpha=0,3$	Ljung-Box												
$\beta=0,1$	$m = 2$	8,38	15,52	8,56	14,80	8,36	15,82	8,78	15,10	8,76	15,86	9,56	15,52
	$m = 3$	7,76	13,48	8,64	14,98	7,86	13,56	9,04	15,32	8,08	14,24	9,54	15,72
	$m = 4$	7,34	13,70	8,50	15,12	7,36	13,68	8,92	15,26	7,66	13,94	9,52	15,62
	$m = 6$	6,66	12,60	8,02	14,00	6,72	12,64	7,98	14,32	6,98	12,90	8,06	14,72
	$m = 12$	6,12	11,28	6,98	12,40	6,00	11,36	6,96	12,48	5,94	11,50	7,14	12,74
	$m = 24$	5,58	10,82	6,12	11,98	5,72	10,76	6,08	11,94	5,88	10,96	6,12	12,06
	$m = 36$	5,72	10,96	6,20	11,82	5,66	10,98	6,32	11,66	5,62	11,14	6,36	11,84
	Francq-Roy-Zakoian												
	$m = 2$	5,04	10,90	5,36	10,02	4,98	10,70	5,24	10,14	5,06	10,42	5,42	10,24
	$m = 3$	4,76	9,52	5,28	10,04	4,50	9,52	5,26	10,24	4,40	9,54	5,36	10,46
	$m = 4$	4,60	9,26	5,42	10,92	4,74	9,30	5,48	11,00	4,70	9,30	5,48	11,04
	$m = 6$	4,72	9,34	5,38	10,30	4,86	9,52	5,30	10,48	4,78	9,62	5,38	10,70
	$m = 12$	4,30	9,34	5,04	10,18	4,32	9,30	5,10	9,96	4,14	9,38	5,08	9,94
	$m = 24$	3,88	8,60	4,46	9,70	3,92	8,78	4,48	9,64	3,92	8,64	4,50	9,52
	$m = 36$	3,46	8,28	4,46	9,58	3,42	8,08	4,60	9,56	3,32	8,02	4,58	9,54
	Nouveau test												
	$J = 1$	5,16	9,90	5,42	10,48	5,14	10,64	5,54	10,78	5,84	11,44	5,64	11,24
	$J = 2$	4,50	9,92	5,28	10,04	5,06	10,92	5,44	11,06	5,82	11,82	6,08	11,74
	$J = 3$	8,44	14,60	7,50	12,94	12,78	18,90	9,88	15,92	14,38	20,38	11,04	16,56
$\alpha=0,6$	Ljung-Box												
$\beta=0,1$	$m = 2$	22,12	29,82	25,50	33,26	22,52	30,56	26,16	33,72	23,10	31,38	26,80	34,26
	$m = 3$	24,08	32,50	29,90	38,40	24,60	33,12	30,26	39,24	25,36	33,96	31,06	40,22
	$m = 4$	24,22	33,18	31,14	39,84	24,66	33,64	31,78	40,52	25,48	34,48	32,24	41,34
	$m = 6$	23,38	31,92	30,58	38,94	23,62	32,18	30,60	38,96	24,06	32,92	30,98	39,74
	$m = 12$	18,78	25,90	25,26	33,60	19,16	26,08	25,62	33,74	19,38	26,78	26,10	33,96
	$m = 24$	14,30	21,52	19,50	26,94	14,54	21,84	20,06	27,22	14,62	21,92	20,00	27,52
	$m = 36$	12,18	18,06	16,54	23,78	12,24	18,40	17,00	24,08	12,44	18,64	17,50	24,32
	Francq-Roy-Zakoian												
	$m = 2$	5,66	10,46	5,24	10,54	5,46	11,06	5,46	10,92	5,52	11,12	5,20	10,90
	$m = 3$	4,76	10,04	5,12	9,78	4,92	10,08	5,00	10,04	4,88	9,88	4,66	10,08
	$m = 4$	4,58	9,56	4,70	10,06	4,80	9,68	4,70	10,04	4,70	9,66	4,70	9,92
	$m = 6$	4,96	9,36	4,28	9,26	4,80	9,42	4,38	9,40	4,64	9,66	4,48	9,22
	$m = 12$	4,62	9,16	4,28	8,52	4,52	9,36	4,46	8,58	4,52	9,30	4,62	8,52
	$m = 24$	3,60	8,38	3,92	8,44	3,74	8,28	3,96	8,48	3,84	8,24	3,90	8,48
	$m = 36$	3,28	6,90	3,50	7,50	3,34	7,00	3,56	7,54	3,32	7,20	3,44	7,86
	Nouveau test												
	$J = 1$	4,02	9,40	4,52	9,98	4,58	10,04	4,90	10,24	5,36	10,98	5,16	11,06
	$J = 2$	4,62	9,50	4,28	10,02	5,16	10,36	4,84	10,44	5,78	11,08	5,38	11,06
	$J = 3$	8,68	14,24	7,16	11,70	11,80	17,72	8,96	14,58	13,06	19,02	10,14	15,76

Tableau 4.2. Niveaux empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(1) avec des innovations GARCH(1,1), $\phi \in \{0,1; 0,2; 0,3\}$, $\alpha = 0,1$, $\beta \in \{0,3; 0,6\}$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$.

		$\phi = 0,1$				$\phi = 0,2$				$\phi = 0,3$			
		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
$\alpha=0,1$	Ljung-Box												
$\beta=0,3$	$m = 2$	6,04	11,58	5,90	10,80	5,86	11,64	5,94	10,84	6,04	11,78	6,12	10,86
	$m = 3$	5,66	11,08	6,12	11,32	5,62	11,04	6,36	11,14	5,68	11,10	6,40	11,22
	$m = 4$	5,34	10,82	6,30	11,42	5,40	11,00	6,48	11,58	5,54	11,00	6,68	11,64
	$m = 6$	5,24	10,42	5,94	11,50	5,34	10,16	5,96	11,30	5,42	10,44	5,98	11,56
	$m = 12$	5,28	9,90	5,72	10,70	5,28	9,84	5,90	10,68	5,12	9,86	5,94	10,76
	$m = 24$	5,26	9,76	5,56	10,68	5,26	9,66	5,68	10,70	5,20	9,72	5,64	10,74
	$m = 36$	5,34	10,10	5,68	11,04	5,30	10,16	5,70	11,10	5,16	10,26	5,74	11,00
	Francq-Roy-Zakoian												
	$m = 2$	5,12	10,40	5,14	9,78	5,30	10,12	5,04	9,74	5,08	10,38	5,04	9,48
	$m = 3$	4,90	9,88	5,34	9,76	5,02	9,88	5,38	9,88	4,98	9,82	5,36	10,00
	$m = 4$	4,62	9,72	5,54	10,44	4,72	9,82	5,58	10,72	4,74	9,78	5,70	10,52
	$m = 6$	4,48	9,84	5,06	10,48	4,46	9,52	5,08	10,52	4,70	9,68	5,16	10,58
	$m = 12$	4,60	9,44	5,16	10,00	4,54	9,36	5,22	10,06	4,42	9,34	5,28	10,14
	$m = 24$	4,32	8,80	4,70	9,70	4,24	8,60	4,68	9,72	4,16	8,74	4,74	9,72
	$m = 36$	3,68	8,42	4,78	9,92	3,58	8,46	4,82	9,98	3,52	8,38	4,70	9,94
	Nouveau test												
	$J = 1$	5,20	10,30	5,40	10,00	5,52	10,74	5,82	10,32	5,86	11,30	5,88	10,68
	$J = 2$	4,86	10,20	5,18	10,50	5,10	11,02	5,54	11,18	5,88	11,46	5,96	11,64
	$J = 3$	8,68	14,54	7,28	13,00	13,46	19,80	10,26	16,36	14,96	21,50	11,46	17,40
$\alpha=0,1$	Ljung-Box												
$\beta=0,6$	$m = 2$	6,68	12,84	6,40	12,04	6,82	12,96	6,66	12,10	6,88	13,08	6,80	12,24
	$m = 3$	6,66	12,70	7,24	12,82	6,76	12,36	7,36	12,96	6,82	12,66	7,50	13,22
	$m = 4$	6,62	12,88	7,76	13,68	6,60	12,96	8,00	13,72	6,80	13,08	7,90	14,02
	$m = 6$	6,56	12,66	7,42	13,86	6,52	12,62	7,58	14,02	6,66	12,54	7,62	14,04
	$m = 12$	5,94	11,54	6,66	12,38	5,98	11,62	6,74	12,28	5,82	11,64	6,90	12,38
	$m = 24$	5,88	10,66	6,10	11,62	5,96	10,56	6,24	11,64	5,82	10,50	6,30	11,62
	$m = 36$	5,88	10,90	6,64	12,36	5,78	10,86	6,58	12,36	5,80	11,04	6,62	12,26
	Francq-Roy-Zakoian												
	$m = 2$	5,14	10,44	5,04	9,76	5,18	10,24	4,90	9,76	5,26	10,22	5,00	9,58
	$m = 3$	5,06	9,64	5,24	10,16	4,88	9,58	5,38	9,94	4,90	9,60	5,50	10,06
	$m = 4$	4,76	9,88	5,68	10,58	4,80	9,96	5,72	10,66	4,84	9,86	5,64	10,82
	$m = 6$	4,68	9,98	5,10	10,48	4,66	9,94	5,16	10,64	4,64	10,00	5,34	10,66
	$m = 12$	4,60	9,32	5,24	10,14	4,70	9,32	5,26	10,16	4,64	9,32	5,30	10,18
	$m = 24$	4,26	8,70	4,84	9,60	4,26	8,64	4,78	9,64	4,20	8,62	4,84	9,74
	$m = 36$	3,56	8,56	4,82	10,06	3,52	8,64	4,92	10,04	3,62	8,54	4,94	10,14
	Nouveau test												
	$J = 1$	5,26	9,96	5,46	10,10	5,32	10,44	5,62	10,24	5,88	11,40	5,76	10,50
	$J = 2$	4,88	10,20	5,26	10,54	5,10	10,88	5,50	11,06	5,80	11,50	6,06	11,44
	$J = 3$	8,46	14,30	7,44	12,84	13,42	19,82	10,28	16,24	14,82	21,20	11,56	17,60

taux de rejet sont légèrement au-dessus des taux nominaux pour $J = 2$ et un peu au-dessus des taux nominaux pour $J = 3$. On note une certaine amélioration quand n augmente.

Les résultats de l'étude de niveau pour les processus générés selon les modèles AR(1) et AR(2) avec des innovations iid sont présentés dans les tableaux 4.5 et 4.6, respectivement. Les tests LB et FRZ performant bien dans les deux cas et pour les nouveaux tests, les résultats sont similaires aux résultats des tableaux 4.1 et 4.3, respectivement.

Les résultats pour les tests de LB justifient la nécessité d'utiliser les tests modifiés de FRZ ou encore d'autres tests lorsque les innovations sont non corrélées mais dépendantes. Ainsi, ceci suggère que la distribution exacte des tests LB s'éloigne passablement de l'approximation asymptotique fournie par la loi χ_{m-p-q}^2 . On note que pour m élevé la situation est moins problématique car les niveaux s'améliorent à mesure que m augmente.

Les tests FRZ font bien en général quelle que soit la situation. On remarque que selon la méthodologie de Francq et al. (2005), il faut choisir des points de troncature car les lois asymptotiques font intervenir des sommes infinies. Cependant la convergence des sommes en cause est rapide (plus rapide que les sommes intervenant dans les nouveaux tests). Ainsi, les valeurs des coefficients autorégressifs dans les modèles AR ou encore les coefficients (α, β) dans le modèle GARCH ont peu d'impact sur les résultats. Cependant, on note une détérioration dans les niveaux comme fonction de m ce qui peut être considéré décevant si des effets saisonniers sont suspectés.

Les nouveaux tests ont affiché de bons comportements pour les niveaux de résolution $J = 1, 2$, mais comme mentionné précédemment les sommes infinies sont plus difficiles à calculer avec précision et la tolérance τ^2 a du être spécifiée assez petite car l'amélioration de l'approximation par la loi asymptotique était perceptible à mesure que τ^2 devenait petit. On a noté une plus grande sensibilité au modèle (en fonction de ϕ dans les modèles AR) que les tests FRZ. On a noté également une amélioration en fonction de la taille échantillonnale n . Les nouveaux tests reposent sur un estimateur de la densité spectrale. Comme mentionné, il est bien connu que de tels estimateurs convergent plus lentement. De ce point de vue, il n'est pas très surprenant que des tailles échantillonnales plus grandes soient nécessaires. Cependant, la détection de phénomènes non corrélés mais dépendants

Tableau 4.3. Niveaux empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(2) avec des innovations GARCH(1,1), $\phi \in \{0,1; 0,2; 0,3\}$, $\alpha \in \{0,3; 0,6\}$, $\beta = 0,1$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$.

		$\phi = 0,1$				$\phi = 0,2$				$\phi = 0,3$			
		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
$\alpha=0,3$	Ljung-Box												
$\beta=0,1$	$m = 2$	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	$m = 3$	5,62	11,12	6,42	12,20	6,04	11,52	7,02	12,46	6,40	12,76	7,64	13,72
	$m = 4$	5,70	10,86	6,66	11,98	5,94	11,54	6,78	12,38	6,70	12,32	7,28	12,98
	$m = 6$	5,28	10,52	6,36	12,04	5,44	10,88	6,60	12,48	5,92	11,74	7,10	12,74
	$m = 12$	4,76	9,68	6,34	10,92	5,00	9,66	6,12	11,36	5,28	10,06	6,22	11,76
	$m = 24$	4,74	9,46	5,64	10,28	4,78	9,40	5,58	10,34	4,92	9,56	5,60	10,60
	$m = 36$	4,74	9,86	5,36	10,66	4,96	9,94	5,46	10,82	5,18	9,76	5,44	10,80
	Francq-Roy-Zakoian												
	$m = 2$	4,72	10,02	5,36	10,36	5,04	10,34	5,22	10,50	5,34	10,50	5,42	10,24
	$m = 3$	4,58	9,52	4,96	10,32	4,82	9,64	5,12	10,26	4,84	9,68	5,34	10,36
	$m = 4$	4,60	9,94	5,40	10,14	5,00	10,22	5,52	9,98	4,94	10,40	5,36	10,16
	$m = 6$	4,42	9,42	5,34	10,50	4,50	9,78	5,40	10,32	4,78	10,00	5,50	10,52
	$m = 12$	4,10	8,74	5,22	9,92	4,16	9,04	5,22	9,88	4,32	9,06	4,88	9,90
	$m = 24$	3,68	8,44	4,34	9,40	3,74	8,30	4,32	9,38	3,78	8,28	4,22	9,22
	$m = 36$	3,10	7,76	4,42	9,58	3,16	7,84	4,42	9,82	3,20	7,78	4,44	9,54
	Nouveau test												
	$J = 1$	5,16	10,46	5,12	10,28	5,26	10,14	5,32	10,10	5,10	10,56	5,78	10,06
	$J = 2$	4,66	10,02	5,18	10,34	5,74	11,34	5,84	11,04	7,94	13,80	6,90	12,88
	$J = 3$	7,56	12,86	6,38	11,46	8,84	14,36	7,28	12,82	10,02	15,26	8,02	13,28
$\alpha=0,6$	Ljung-Box												
$\beta=0,1$	$m = 2$	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	$m = 3$	13,70	21,12	17,44	25,22	14,94	22,28	18,88	26,32	16,68	25,28	20,60	29,20
	$m = 4$	15,66	23,32	21,36	28,38	16,58	24,70	22,20	29,50	18,58	26,68	22,94	30,84
	$m = 6$	15,54	23,48	21,04	29,24	16,36	24,76	22,06	30,22	17,96	25,66	23,44	31,38
	$m = 12$	12,10	19,00	17,68	25,22	12,72	19,80	18,42	26,02	13,80	20,86	19,06	27,22
	$m = 24$	9,12	15,04	13,20	19,68	9,64	15,62	13,40	20,38	10,16	16,68	14,32	21,10
	$m = 36$	8,08	13,50	11,58	17,60	8,58	13,56	12,18	17,88	8,86	13,98	12,50	18,58
	Francq-Roy-Zakoian												
	$m = 2$	5,28	10,86	5,30	9,54	5,24	10,80	5,10	9,92	5,58	10,78	4,92	9,80
	$m = 3$	4,94	9,92	4,86	10,04	4,94	10,28	4,80	9,98	5,06	10,02	4,82	9,68
	$m = 4$	4,38	9,32	4,70	9,52	4,68	9,82	4,90	9,40	5,10	10,10	4,88	9,20
	$m = 6$	4,64	8,76	4,42	9,40	4,80	9,18	4,52	9,12	5,06	9,52	4,40	8,92
	$m = 12$	4,22	8,58	4,08	8,38	4,22	8,52	3,98	8,50	4,38	8,76	3,94	8,44
	$m = 24$	3,52	7,62	3,38	7,64	3,36	7,82	3,52	7,76	3,52	7,94	3,50	7,66
	$m = 36$	3,00	6,10	3,28	7,14	2,88	6,32	3,36	7,30	2,86	6,48	3,52	7,48
	Nouveau test												
	$J = 1$	5,26	11,10	5,04	9,58	5,14	10,22	5,26	10,38	4,86	10,46	4,88	9,78
	$J = 2$	4,64	9,86	4,50	9,74	5,96	11,62	5,80	10,74	8,14	13,60	6,98	12,82
	$J = 3$	9,10	14,84	6,80	11,86	10,04	15,42	7,70	12,82	10,56	16,50	8,16	13,60

Tableau 4.4. Niveaux empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(2) avec des innovations GARCH(1,1), $\phi \in \{0,1; 0,2; 0,3\}$, $\alpha = 0,1$, $\beta \in \{0,3; 0,6\}$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$.

		$\phi = 0,1$				$\phi = 0,2$				$\phi = 0,3$			
		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
$\alpha=0,1$	Ljung-Box												
$\beta=0,3$	$m = 2$	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	$m = 3$	4,94	10,32	5,48	10,36	5,00	10,24	5,42	10,88	5,56	11,18	5,94	11,42
	$m = 4$	4,86	9,92	5,58	10,66	5,02	10,16	5,44	10,64	5,40	10,40	5,60	11,00
	$m = 6$	4,66	9,98	5,74	10,86	4,82	10,04	5,76	10,62	5,08	10,24	5,70	10,88
	$m = 12$	4,78	9,38	5,54	10,44	4,78	9,56	5,48	10,62	4,70	9,68	5,42	10,46
	$m = 24$	4,62	9,36	5,30	10,38	4,54	9,46	5,36	10,34	4,62	9,68	5,36	10,56
	$m = 36$	4,54	9,80	5,50	10,50	4,60	9,84	5,48	10,48	4,70	9,86	5,44	10,58
	Francq-Roy-Zakoian												
	$m = 2$	4,76	9,80	5,24	10,54	4,78	10,20	5,14	10,22	4,86	10,30	5,44	9,90
	$m = 3$	4,54	10,06	5,14	9,84	4,48	10,12	5,14	9,98	4,72	9,84	5,30	10,12
	$m = 4$	4,78	9,96	5,20	10,38	4,60	10,14	5,10	10,26	4,98	10,56	5,12	10,32
	$m = 6$	4,58	9,72	5,32	10,10	4,56	9,86	5,40	10,12	4,80	10,10	5,22	10,36
	$m = 12$	4,38	9,26	5,08	10,16	4,32	9,20	5,10	10,26	4,58	9,34	5,04	10,06
	$m = 24$	3,82	8,62	4,82	9,88	3,88	8,62	4,78	9,76	3,96	8,70	4,66	9,84
	$m = 36$	3,36	8,14	4,66	9,92	3,40	8,16	4,64	9,88	3,38	8,32	4,62	9,88
	Nouveau test												
	$J = 1$	5,08	10,42	4,96	9,62	4,92	9,86	5,22	9,82	5,02	10,16	5,40	10,58
	$J = 2$	4,88	9,40	5,04	10,46	5,32	10,86	5,52	10,56	7,72	13,28	6,94	12,16
	$J = 3$	7,22	12,48	5,96	11,52	8,24	14,06	6,80	12,18	9,20	14,74	7,60	12,88
$\alpha=0,1$	Ljung-Box												
$\beta=0,6$	$m = 2$	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	$m = 3$	5,88	11,10	6,36	11,96	6,04	11,34	6,44	12,40	6,30	12,26	6,86	12,80
	$m = 4$	5,88	11,50	6,98	12,56	6,12	12,00	6,90	12,72	6,44	12,04	6,96	12,48
	$m = 6$	6,00	11,82	6,88	12,52	6,06	11,78	6,94	12,58	6,48	12,18	6,98	12,48
	$m = 12$	5,24	10,66	6,40	11,86	5,18	10,80	6,50	11,94	5,46	10,90	6,46	12,08
	$m = 24$	5,20	10,00	5,90	11,40	5,08	10,08	5,98	11,38	5,16	10,22	5,90	11,50
	$m = 36$	5,22	10,46	6,24	11,50	5,16	10,36	6,08	11,52	5,26	10,20	5,90	11,58
	Francq-Roy-Zakoian												
	$m = 2$	4,94	9,78	5,46	10,32	4,74	10,12	5,28	10,38	5,02	10,44	5,14	10,24
	$m = 3$	4,74	9,78	5,28	9,82	4,64	9,94	5,22	9,98	4,80	9,86	5,16	10,26
	$m = 4$	4,60	10,00	5,46	10,28	4,74	10,42	5,32	10,28	4,84	10,56	5,24	10,36
	$m = 6$	4,76	9,66	5,56	10,20	4,68	9,90	5,68	9,96	4,84	10,10	5,48	10,16
	$m = 12$	4,28	8,96	5,24	10,28	4,40	9,10	5,18	10,22	4,44	9,22	4,98	10,14
	$m = 24$	3,84	8,44	4,72	9,72	3,80	8,34	4,74	9,68	3,94	8,34	4,66	9,74
	$m = 36$	3,34	8,14	4,70	10,10	3,26	8,28	4,80	10,00	3,28	8,26	4,70	9,90
	Nouveau test												
	$J = 1$	5,12	10,42	4,92	10,00	4,82	10,08	5,34	9,90	4,96	10,00	5,56	10,44
	$J = 2$	4,58	9,42	5,14	10,08	5,40	10,90	5,56	10,72	7,72	13,28	6,94	12,40
	$J = 3$	7,44	12,42	6,22	11,28	8,12	14,20	6,90	12,16	9,22	14,54	7,60	12,86

Tableau 4.5. Niveaux empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(1) avec des innovations iid, $\phi \in \{0,1; 0,2; 0,3\}$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$.

	$\phi = 0,1$				$\phi = 0,2$				$\phi = 0,3$			
	$n = 500$		$n = 1000$		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
Ljung-Box												
$m = 2$	5,08	9,76	4,84	9,30	5,14	9,78	4,98	9,38	5,16	10,04	4,98	9,46
$m = 3$	4,78	9,92	5,24	9,58	4,76	9,82	5,18	9,56	4,80	10,00	5,18	9,80
$m = 4$	4,72	9,42	5,28	10,16	4,70	9,46	5,42	10,28	4,60	9,52	5,28	10,30
$m = 6$	4,70	9,60	5,16	10,26	4,70	9,46	5,16	10,44	4,82	9,58	5,16	10,42
$m = 12$	5,12	9,56	5,38	10,26	5,10	9,42	5,46	10,30	4,98	9,38	5,44	10,48
$m = 24$	5,06	9,28	5,42	10,02	4,88	9,26	5,46	10,10	4,90	9,44	5,44	10,30
$m = 36$	4,98	9,78	5,74	10,74	5,02	9,72	5,74	10,84	5,02	9,46	5,78	10,90
Francq-Roy-Zakoian												
$m = 2$	5,16	10,20	4,90	9,50	5,32	10,26	5,10	9,36	5,14	10,16	4,96	9,30
$m = 3$	5,00	9,82	5,06	9,80	5,06	9,86	5,14	9,88	4,98	9,96	5,20	10,02
$m = 4$	4,68	9,54	5,40	10,04	4,68	10,02	5,46	10,16	4,82	9,92	5,46	10,26
$m = 6$	4,56	10,10	5,30	10,38	4,52	9,98	5,32	10,38	4,64	9,84	5,26	10,44
$m = 12$	4,86	9,26	5,18	10,12	4,72	9,30	5,26	10,22	4,54	9,32	5,18	10,32
$m = 24$	4,14	8,94	4,86	9,68	4,22	8,72	4,88	9,72	4,14	8,70	4,84	9,70
$m = 36$	3,86	8,32	4,90	9,92	3,94	8,30	4,84	9,94	3,80	8,14	4,88	10,06
Nouveau test												
$J = 1$	5,28	10,54	5,14	9,68	5,60	10,70	5,62	10,40	6,12	11,36	5,66	10,52
$J = 2$	4,98	10,02	5,08	10,82	5,28	10,76	5,42	10,90	5,86	11,30	5,64	11,46
$J = 3$	8,70	14,40	7,18	12,80	13,32	20,00	10,04	16,50	14,84	21,44	11,12	17,68

exigent des tailles plus élevées, et les applications de nos méthodes impliquant des modèles GARCH proviennent principalement de l'économétrie et la finance. Globalement, si l'utilisateur veut utiliser l'approximation asymptotique, les choix $J = 1, 2$ sont recommandés. Bien que cela dépasse l'objectif principal de ce mémoire, l'utilisation de méthodes de ré-échantillonnage comme le bootstrap pourrait être envisagée si des valeurs supérieures à $J = 2$ étaient requises.

Tableau 4.6. Niveaux empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(2) avec des innovations iid, $\phi \in \{0,1; 0,2; 0,3\}$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$.

	$\phi = 0,1$				$\phi = 0,2$				$\phi = 0,3$			
	$n = 500$		$n = 1000$		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
Ljung-Box												
$m = 2$	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
$m = 3$	4,58	9,78	5,08	9,68	4,80	9,98	5,28	9,98	4,96	10,24	5,32	10,30
$m = 4$	4,54	9,54	5,22	10,24	4,50	9,58	5,22	10,36	4,68	9,86	5,26	10,20
$m = 6$	4,60	9,62	5,24	10,14	4,70	9,76	5,28	10,10	4,94	9,98	5,26	10,42
$m = 12$	4,68	9,28	5,18	10,32	4,68	9,22	5,24	10,24	4,64	9,48	5,12	10,14
$m = 24$	4,78	9,18	5,42	10,18	4,62	9,10	5,42	10,06	4,66	9,18	5,44	10,12
$m = 36$	4,70	9,36	5,76	10,34	4,80	9,36	5,70	10,24	4,70	9,36	5,52	10,28
Francq-Roy-Zakoian												
$m = 2$	4,92	9,74	5,26	10,28	4,80	9,82	4,88	10,24	5,06	10,24	5,18	9,88
$m = 3$	4,56	10,04	5,12	9,82	4,68	10,32	5,12	9,76	4,90	9,98	5,18	9,82
$m = 4$	4,58	9,88	5,16	10,32	4,72	10,08	5,36	10,14	4,76	10,22	5,28	10,24
$m = 6$	4,66	9,88	5,28	10,24	4,70	10,16	5,26	10,14	4,76	10,20	5,08	10,32
$m = 12$	4,48	9,24	4,96	10,00	4,54	9,14	4,94	10,06	4,48	9,20	4,96	10,02
$m = 24$	3,90	8,74	5,02	9,84	3,86	8,74	4,96	9,70	3,84	8,70	4,88	9,74
$m = 36$	3,46	7,82	4,92	9,72	3,50	7,86	4,90	9,54	3,36	7,98	4,80	9,66
Nouveau test												
$J = 1$	5,12	10,32	4,80	9,50	4,82	9,96	5,14	10,12	4,78	10,08	5,42	10,66
$J = 2$	4,74	9,66	5,22	10,00	5,38	10,82	5,50	10,42	7,56	13,32	6,60	12,06
$J = 3$	6,86	12,46	5,82	11,00	8,00	13,72	6,64	12,08	8,88	14,38	7,12	12,82

4.3. Discussion de l'étude sur la puissance

Dans cette section, nous discutons les résultats de puissance lorsque les alternatives sont multiplicatives comme détaillées en (4.1.4) et (4.1.5).

Les résultats lorsque l'alternative est un modèle $AR(1) \times AR(4)$ mais qu'un modèle $AR(1)$ est ajusté, avec innovations GARCH(1,1), sont présentés dans les tableaux 4.7 et 4.8. Puisque la forme du test est essentiellement la même et que des points critiques empiriques sont utilisés, les résultats pour les tests LB et FRZ sont largement similaires. La puissance exacte est optimale pour $m = 4$ et décroît à mesure que m augmente. La puissance pour $m = 2$ et $m = 3$ est beaucoup plus faible. Ceci s'explique par le fait que la dépendance aux délais supérieurs n'est pas prise en compte si $m < 4$ dans les tests LB ou FRZ.

Pour les nouveaux tests, la puissance est maximisée pour $J = 2$ et la puissance pour $J = 1$ et $J = 3$ est beaucoup plus faible. Parfois, la puissance maximale des nouveaux tests est plus grande que celle des autres tests. C'est le cas quand le coefficient α est grand, entre autres, pour $(\alpha, \beta) = (0,6, 0,1)$. Cependant, ce n'est pas le cas quand $(\alpha, \beta) = (0,3, 0,1)$. Ces conclusions sont valides pour toutes les valeurs de n et des valeurs nominales. L'écart augmente quand ϕ_4 augmente.

Les résultats de l'étude de puissance pour un processus généré selon le modèle $AR(2)$ avec l'innovation de modèle GARCH(1,1) sont présentés dans les tableaux 4.9 et 4.10. Encore une fois, les résultats des tests LB et FRZ sont similaires et la puissance semble être maximisée pour $m = 4$ et diminue quand m augmente. La puissance pour $m = 3$ est beaucoup plus faible. Pour les nouveaux tests, la puissance est maximisée pour $J = 1$, plus faible lorsque $J = 3$ et, pour $J = 2$, elle est entre les deux. Nous observons encore des situations où la puissance maximale est plus haute pour les nouveaux tests que pour les autres quand α est grand et l'écart augmente quand ϕ_4 augmente. Ces conclusions sont valides pour les valeurs de n utilisées et pour les deux niveaux nominaux, au moins pour les alternatives considérées dans notre étude.

Dans les tableaux 4.11 et 4.12 sont présentés les résultats des processus de modèles $AR(1)$ et $AR(2)$ avec les innovations iid et les comportements sont similaires aux tableaux 4.7 et 4.9, respectivement, sauf qu'il n'y a pas de situation où la puissance maximale des nouveaux tests est plus grande que celles des autres tests.

Comme mentionné précédemment, plus la valeur de α est grande, plus la volatilité est importante et donc, moins de régularité. Nos résultats sont donc en accord avec

Tableau 4.7. Puissances empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(1)×AR(4) avec des innovations GARCH(1,1), $(\phi_1; \phi_4) \in \{(0,2; 0,1); (0,3; 0,1); (0,3; 0,2)\}$, $\alpha \in \{0,3; 0,6\}$, $\beta = 0,1$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$.

		$(\phi_1=0,2; \phi_4 = 0,1)$				$(\phi_1=0,3; \phi_4 = 0,1)$				$(\phi_1=0,3; \phi_4 = 0,2)$			
		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
$\alpha=0,3$	Ljung-Box												
$\beta=0,1$	$m = 2$	7,48	14,14	7,16	13,52	7,94	13,74	7,04	12,96	11,08	17,58	9,70	16,84
	$m = 3$	6,84	12,68	6,76	12,08	6,72	12,30	6,86	11,94	9,12	15,38	8,64	14,32
	$m = 4$	37,86	51,40	67,94	78,58	37,38	50,48	67,56	77,78	95,86	97,86	99,98	100,00
	$m = 6$	32,72	44,46	60,96	72,80	32,20	43,56	60,64	72,24	93,12	96,12	99,88	99,98
	$m = 12$	23,76	35,28	49,14	62,28	24,30	35,18	48,62	61,46	85,80	92,24	99,38	99,76
	$m = 24$	18,24	27,86	37,46	49,72	18,08	27,34	37,08	49,46	72,90	82,88	98,24	99,02
	$m = 36$	15,98	23,74	29,06	41,84	16,32	23,26	29,32	41,76	65,74	75,34	95,62	97,76
	Francq-Roy-Zakoian												
	$m = 2$	7,82	13,74	7,44	13,34	7,52	13,84	7,26	13,20	11,08	18,00	9,92	17,18
	$m = 3$	6,76	12,90	6,88	12,32	6,90	12,82	6,66	11,96	9,56	15,44	8,50	14,72
	$m = 4$	38,58	52,36	67,94	78,56	38,56	51,60	67,68	78,22	95,50	97,74	99,86	99,98
	$m = 6$	31,60	44,96	60,68	72,98	31,28	44,48	60,16	72,28	92,20	96,02	99,76	99,92
	$m = 12$	24,02	35,32	48,36	62,46	23,98	34,94	47,98	62,22	85,40	91,78	99,30	99,74
	$m = 24$	17,82	27,40	37,26	49,66	17,90	27,36	36,50	49,64	72,18	83,14	97,96	99,04
	$m = 36$	15,94	24,38	29,68	41,90	15,40	24,58	29,58	41,76	63,28	75,60	95,56	97,76
	Nouveau test												
	$J = 1$	9,00	15,24	9,10	15,26	15,28	24,52	24,60	34,08	40,18	51,64	63,18	71,12
	$J = 2$	33,92	45,44	64,84	75,74	32,36	44,56	62,94	74,90	93,00	96,16	99,92	99,98
	$J = 3$	7,82	24,98	38,06	57,56	6,48	22,54	33,78	55,00	33,50	80,24	98,20	99,60
$\alpha=0,6$	Ljung-Box												
$\beta=0,1$	$m = 2$	6,86	13,08	6,64	13,64	6,68	13,70	6,62	12,46	8,60	17,22	8,80	15,64
	$m = 3$	6,32	11,92	6,02	11,86	6,30	11,74	5,96	11,64	7,64	13,84	7,08	13,32
	$m = 4$	14,42	27,86	21,26	43,24	13,56	27,36	19,48	42,52	64,50	83,80	90,94	97,60
	$m = 6$	12,66	24,14	18,48	38,46	12,36	24,62	17,54	36,64	57,80	79,52	88,24	96,52
	$m = 12$	12,96	23,22	16,54	34,66	12,10	22,82	16,70	33,64	53,54	73,72	85,56	95,52
	$m = 24$	12,50	23,22	17,70	32,24	12,30	22,50	16,46	31,40	50,60	68,06	81,64	93,20
	$m = 36$	12,36	21,00	15,56	30,78	12,12	20,94	15,80	31,22	46,52	62,48	77,86	90,88
	Francq-Roy-Zakoian												
	$m = 2$	7,28	13,14	7,20	12,92	7,40	13,42	7,36	13,04	10,24	17,52	10,28	16,66
	$m = 3$	6,86	12,18	6,68	11,98	6,60	12,28	7,04	11,96	8,46	15,42	9,12	14,52
	$m = 4$	26,00	37,88	42,76	55,38	25,16	37,00	42,24	54,64	77,66	85,48	91,38	94,34
	$m = 6$	22,08	34,72	40,16	52,76	22,28	33,54	39,80	51,66	73,34	82,68	90,06	93,78
	$m = 12$	17,90	28,64	33,64	48,08	17,42	28,90	33,34	47,86	66,44	77,80	88,22	92,92
	$m = 24$	14,78	24,96	28,28	41,80	14,76	24,60	27,98	41,34	58,26	70,90	85,76	91,08
	$m = 36$	14,56	23,88	24,62	37,88	14,36	23,38	24,34	36,80	52,86	65,96	83,00	89,44
	Nouveau test												
	$J = 1$	9,58	16,40	9,78	16,38	13,12	20,88	19,40	27,72	33,40	44,02	50,52	58,58
	$J = 2$	25,98	38,42	49,24	60,30	24,94	37,76	47,92	59,46	81,22	87,64	94,84	96,82
	$J = 3$	6,46	21,02	25,80	44,88	5,92	18,78	22,42	42,40	21,92	66,32	85,86	93,18

Tableau 4.8. Puissances empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(1)×AR(4) avec des innovations GARCH(1,1), $(\phi_1; \phi_4) \in \{(0,2; 0,1); (0,3; 0,1); (0,3; 0,2)\}$, $\alpha = 0,1$, $\beta \in \{0,3; 0,6\}$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$.

		$(\phi_1=0,2; \phi_4 = 0,1)$				$(\phi_1=0,3; \phi_4 = 0,1)$				$(\phi_1=0,3; \phi_4 = 0,2)$			
		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
$\alpha=0,1$	Ljung-Box												
$\beta=0,3$	$m = 2$	7,36	14,08	7,08	13,50	7,68	14,02	7,12	13,32	10,94	18,04	9,78	17,44
	$m = 3$	6,70	12,42	6,34	12,48	6,52	12,26	6,44	12,12	8,56	15,46	8,48	14,50
	$m = 4$	42,38	54,20	72,16	82,44	41,68	53,76	72,20	82,32	96,96	98,52	100,00	100,00
	$m = 6$	34,72	47,74	65,18	75,16	34,76	46,70	64,68	75,08	94,52	97,30	99,96	100,00
	$m = 12$	24,04	36,72	50,70	63,94	24,24	36,36	50,14	63,62	87,38	92,76	99,56	99,90
	$m = 24$	17,78	28,26	37,22	49,80	17,76	27,92	37,06	49,48	74,24	83,88	98,32	99,12
	$m = 36$	15,86	24,40	30,42	42,32	15,90	24,12	30,06	41,88	66,56	76,86	96,24	97,90
	Francq-Roy-Zakoian												
	$m = 2$	7,32	14,28	7,04	13,58	7,50	13,90	7,44	13,74	11,12	17,64	10,28	18,04
	$m = 3$	6,54	12,58	6,46	12,16	6,30	12,48	6,66	11,86	8,50	15,58	8,28	14,64
	$m = 4$	41,98	54,80	71,96	82,12	41,40	53,88	71,68	81,96	96,70	98,52	100,00	100,00
	$m = 6$	35,36	47,14	64,82	75,60	35,26	46,68	64,22	75,24	94,26	97,18	99,96	100,00
	$m = 12$	24,42	35,86	49,96	63,64	24,52	35,82	49,56	63,18	87,00	92,92	99,58	99,88
	$m = 24$	17,92	27,44	37,00	49,98	17,76	27,34	36,84	49,76	73,16	83,90	98,16	99,24
	$m = 36$	15,32	25,14	29,72	42,26	15,30	25,30	29,92	42,28	65,10	76,36	95,94	97,84
	Nouveau test												
	$J = 1$	8,28	14,40	8,00	15,32	15,82	25,08	24,30	36,46	41,52	53,56	64,86	74,50
	$J = 2$	34,68	47,08	65,30	76,86	33,30	45,32	64,62	76,16	93,74	96,54	99,98	100,00
	$J = 3$	9,16	25,42	41,10	58,70	6,98	23,92	36,70	55,84	41,28	81,74	98,92	99,76
$\alpha=0,1$	Ljung-Box												
$\beta=0,6$	$m = 2$	7,36	14,26	7,30	13,08	7,40	13,98	7,18	12,92	10,78	17,86	9,88	16,80
	$m = 3$	6,64	12,30	6,54	12,22	6,54	11,88	6,48	12,08	8,80	15,30	8,42	14,74
	$m = 4$	39,66	51,52	68,98	80,06	39,50	51,42	68,78	79,86	96,06	97,82	99,98	100,00
	$m = 6$	32,24	44,84	61,84	72,96	31,88	44,40	61,08	72,48	92,84	96,16	99,92	100,00
	$m = 12$	23,68	35,32	49,18	62,64	23,76	35,30	49,26	61,90	85,86	91,68	99,46	99,76
	$m = 24$	17,90	28,34	36,60	49,20	17,54	28,16	36,32	48,74	72,58	83,08	97,92	98,96
	$m = 36$	15,60	24,48	29,26	41,08	15,70	24,38	28,68	41,12	64,64	75,96	95,36	97,66
	Francq-Roy-Zakoian												
	$m = 2$	7,16	14,14	7,38	13,38	7,24	14,04	7,32	13,80	10,82	18,04	10,26	17,74
	$m = 3$	6,86	12,48	6,60	11,98	6,70	12,40	6,48	11,68	8,60	15,62	8,08	14,50
	$m = 4$	38,92	51,70	68,78	79,92	39,06	51,58	68,76	79,66	95,80	97,72	100,00	100,00
	$m = 6$	32,98	44,86	61,52	72,90	32,60	44,54	60,88	72,66	92,74	96,14	99,90	100,00
	$m = 12$	23,20	35,68	47,78	62,34	23,12	35,08	47,88	61,72	84,90	91,50	99,46	99,76
	$m = 24$	17,16	27,04	36,30	48,96	17,10	27,42	36,28	48,92	71,48	83,02	97,82	99,04
	$m = 36$	14,94	24,72	29,10	41,34	14,90	24,86	28,76	40,84	63,88	75,58	95,22	97,46
	Nouveau test												
	$J = 1$	8,32	14,38	8,02	15,12	15,24	24,42	23,80	35,88	40,40	52,34	63,78	73,50
	$J = 2$	32,38	44,82	62,66	74,36	31,02	43,42	60,98	73,52	92,42	95,78	99,94	100,00
	$J = 3$	7,84	24,04	37,80	55,74	6,64	21,88	32,92	53,28	35,84	77,82	98,24	99,62

Tableau 4.9. Puissances empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(2)×AR(4) avec des innovations GARCH(1,1), $(\phi_2; \phi_4) \in \{(0,2; 0,1); (0,3; 0,1); (0,3; 0,2)\}$, $\alpha \in \{0,3; 0,6\}$, $\beta = 0,1$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$.

		$(\phi_2=0,2; \phi_4 = 0,1)$				$(\phi_2=0,3; \phi_4 = 0,1)$				$(\phi_2=0,3; \phi_4 = 0,2)$			
		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
$\alpha=0,3$	Ljung-Box												
$\beta=0,1$	$m = 2$	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	$m = 3$	4,90	10,38	5,32	11,90	5,34	11,44	6,68	15,82	16,50	37,00	50,82	80,76
	$m = 4$	41,34	53,36	72,88	84,18	38,68	50,80	70,78	81,70	95,24	97,42	99,96	100,00
	$m = 6$	32,24	44,86	62,56	74,58	30,96	42,88	60,70	72,14	91,46	95,04	99,72	99,94
	$m = 12$	22,78	33,66	47,74	61,40	21,82	32,66	46,30	58,74	82,80	89,32	99,02	99,56
	$m = 24$	16,72	25,86	34,06	48,28	16,42	25,46	32,76	47,02	69,50	79,18	96,78	98,52
	$m = 36$	13,90	22,30	28,20	39,90	13,52	22,40	27,48	38,58	59,68	71,68	93,94	96,64
	Francq-Roy-Zakoian												
	$m = 2$	34,94	47,24	68,92	79,42	30,54	43,60	62,88	75,36	90,38	94,72	99,68	99,92
	$m = 3$	4,54	9,84	5,04	11,30	4,82	10,64	5,94	14,06	12,62	31,06	42,92	73,82
	$m = 4$	39,58	52,80	72,36	83,94	37,46	49,48	69,32	80,80	94,42	96,88	99,90	99,98
	$m = 6$	32,56	44,40	62,30	74,74	30,48	42,48	59,62	72,32	90,78	94,80	99,70	99,92
	$m = 12$	22,98	33,22	47,18	61,80	21,80	32,02	45,92	59,40	81,90	89,22	98,98	99,50
	$m = 24$	16,38	25,42	35,48	47,50	15,96	24,98	34,16	46,22	68,18	78,64	96,78	98,38
	$m = 36$	14,52	21,82	28,12	39,04	13,74	21,66	26,84	38,62	59,62	70,50	93,74	96,52
	Nouveau test												
	$J = 1$	34,80	48,16	67,88	78,14	38,08	51,06	70,10	81,46	94,04	97,00	99,92	99,96
	$J = 2$	31,18	42,44	61,14	73,50	24,50	38,18	56,44	69,62	86,00	92,42	99,54	99,80
	$J = 3$	15,00	29,58	45,36	59,32	13,88	27,10	40,78	56,50	67,24	83,08	98,16	99,20
$\alpha=0,6$	Ljung-Box												
$\beta=0,1$	$m = 2$	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	$m = 3$	4,66	9,34	4,70	9,74	4,40	9,34	4,92	10,60	5,82	15,48	7,86	27,52
	$m = 4$	19,90	35,92	33,02	54,94	18,66	33,16	29,80	50,18	77,00	88,36	94,76	97,90
	$m = 6$	14,78	28,84	25,24	47,10	14,56	27,40	23,24	43,70	68,66	82,46	92,38	97,26
	$m = 12$	14,00	25,06	21,04	40,18	12,90	23,50	19,70	37,68	60,34	76,28	88,50	95,84
	$m = 24$	11,86	21,16	18,24	35,24	11,40	21,14	17,28	33,22	53,12	67,82	83,76	92,72
	$m = 36$	11,54	19,54	18,50	30,72	10,90	19,66	16,76	29,30	47,16	61,70	79,76	89,66
	Francq-Roy-Zakoian												
	$m = 2$	20,70	31,34	42,86	55,76	18,04	29,38	40,12	52,44	69,10	78,76	89,54	93,08
	$m = 3$	4,40	9,24	4,84	9,94	4,52	9,70	4,90	11,30	6,86	17,40	17,10	38,24
	$m = 4$	28,94	39,86	50,36	63,58	25,36	37,32	46,54	60,88	78,36	85,88	92,90	95,90
	$m = 6$	23,36	35,52	43,90	58,42	21,14	33,10	41,10	55,76	73,86	82,54	91,10	94,78
	$m = 12$	17,14	28,60	37,04	49,70	16,34	27,26	35,20	47,38	65,22	76,76	89,40	93,48
	$m = 24$	13,16	22,84	28,64	42,10	12,66	22,64	27,38	40,28	55,06	68,48	86,06	91,44
	$m = 36$	12,92	21,32	23,62	36,10	12,20	20,42	22,96	34,54	50,80	62,30	81,82	88,72
	Nouveau test												
	$J = 1$	25,00	37,08	47,32	59,42	29,28	40,54	52,68	64,56	82,66	88,32	95,22	96,94
	$J = 2$	22,68	34,92	43,44	57,72	18,28	31,54	38,64	52,86	69,78	81,80	91,76	95,24
	$J = 3$	10,52	22,30	27,10	45,02	9,56	21,60	25,80	40,76	44,92	67,90	85,62	92,06

Tableau 4.10. Puissances empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(2)×AR(4) avec des innovations GARCH(1,1), $(\phi_2; \phi_4) \in \{(0,2; 0,1); (0,3; 0,1); (0,3; 0,2)\}$, $\alpha = 0,1$, $\beta \in \{0,3; 0,6\}$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$.

		$(\phi_2=0,2; \phi_4 = 0,1)$				$(\phi_2=0,3; \phi_4 = 0,1)$				$(\phi_2=0,3; \phi_4 = 0,2)$			
		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
$\alpha=0,1$	Ljung-Box												
$\beta=0,3$	$m = 2$	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	$m = 3$	5,00	10,40	5,92	12,84	5,46	12,02	7,66	17,74	19,56	42,02	61,26	86,46
	$m = 4$	43,72	56,30	76,20	85,30	41,78	54,28	74,58	83,72	96,46	98,40	99,96	100,00
	$m = 6$	34,14	45,78	65,08	76,92	32,62	44,40	62,98	74,50	92,52	96,10	99,86	99,94
	$m = 12$	22,92	32,94	48,96	61,58	22,52	32,88	47,60	60,30	84,16	90,10	99,36	99,66
	$m = 24$	17,52	25,40	35,00	47,36	17,52	24,94	33,18	45,86	71,10	79,26	96,92	98,52
	$m = 36$	14,72	22,56	28,34	39,88	14,32	22,14	27,70	38,72	62,16	72,38	94,46	96,78
	Francq-Roy-Zakoian												
	$m = 2$	39,48	51,48	73,30	82,66	35,44	47,26	67,32	78,70	93,36	96,38	99,92	99,98
	$m = 3$	5,00	9,92	5,32	12,02	5,28	11,10	6,84	16,30	16,40	37,14	54,60	83,36
	$m = 4$	43,00	54,32	76,12	84,90	40,24	51,98	73,54	83,20	95,92	98,02	100,00	100,00
	$m = 6$	33,38	45,04	64,84	77,02	32,12	43,70	63,30	74,56	92,08	95,70	99,86	99,96
	$m = 12$	23,16	33,30	48,76	61,56	22,32	32,02	47,02	60,10	83,48	89,88	99,26	99,68
	$m = 24$	16,94	25,20	34,92	47,20	16,82	24,80	33,54	45,82	69,82	78,92	96,90	98,58
	$m = 36$	14,28	22,90	28,06	39,38	14,02	22,16	26,58	38,44	60,90	71,72	94,08	96,74
	Nouveau test												
	$J = 1$	37,52	50,28	69,44	80,24	39,42	52,60	72,52	81,52	95,06	97,64	99,98	100,00
	$J = 2$	31,88	43,54	63,58	74,64	26,18	39,30	58,28	71,72	88,28	93,48	99,82	99,86
	$J = 3$	17,42	30,96	47,42	61,34	15,52	29,30	45,16	58,86	71,70	86,12	98,98	99,56
$\alpha=0,1$	Ljung-Box												
$\beta=0,6$	$m = 2$	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	$m = 3$	4,96	10,16	5,48	12,32	5,36	12,18	7,42	16,52	17,86	39,28	56,10	82,20
	$m = 4$	41,28	53,70	73,06	82,72	38,68	51,58	71,42	81,60	95,34	97,64	99,94	100,00
	$m = 6$	31,54	43,60	61,38	74,20	30,38	42,30	59,78	72,22	90,74	94,96	99,82	99,90
	$m = 12$	22,70	33,02	47,22	59,76	22,00	32,16	46,36	58,66	82,64	88,80	99,14	99,50
	$m = 24$	17,14	25,62	34,42	46,06	16,80	25,14	32,88	45,06	69,32	78,38	96,74	98,14
	$m = 36$	14,32	22,00	27,56	39,36	14,12	22,34	26,80	38,14	60,38	71,66	93,30	96,44
	Francq-Roy-Zakoian												
	$m = 2$	37,12	49,12	70,02	80,36	32,88	44,80	64,78	76,02	91,80	95,18	99,84	99,98
	$m = 3$	4,84	9,82	5,18	11,90	5,06	10,88	6,78	15,36	15,06	34,20	49,94	79,14
	$m = 4$	40,44	52,32	72,66	82,76	37,78	49,56	70,68	80,28	94,80	97,20	99,94	100,00
	$m = 6$	31,82	43,36	61,48	75,08	30,36	41,74	59,14	72,70	90,36	94,64	99,84	99,94
	$m = 12$	22,42	33,02	46,78	60,24	21,62	31,16	45,54	58,80	81,54	88,60	99,04	99,48
	$m = 24$	16,42	25,12	34,20	46,56	16,90	24,46	32,94	45,34	69,22	78,04	96,50	98,06
	$m = 36$	13,62	22,40	27,78	38,60	13,68	21,92	26,52	37,86	59,32	71,00	93,36	96,34
	Nouveau test												
	$J = 1$	35,92	48,26	66,02	78,26	37,74	51,06	68,66	79,82	93,78	97,04	99,92	99,98
	$J = 2$	29,98	41,74	60,88	72,02	24,96	37,54	55,56	69,26	86,64	92,26	99,60	99,84
	$J = 3$	15,88	29,96	44,30	58,80	14,52	28,24	41,24	56,20	68,68	84,44	98,48	99,30

Tableau 4.11. Puissances empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(1)×AR(4) avec des innovations iid, $(\phi_1, \phi_4) \in \{(0,2; 0,1); (0,3; 0,1); (0,3; 0,2)\}$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$.

	$(\phi_1=0,2; \phi_4 = 0,1)$				$(\phi_1=0,3; \phi_4 = 0,1)$				$(\phi_1=0,3; \phi_4 = 0,2)$			
	$n = 500$		$n = 1000$		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
Ljung-Box												
$m = 2$	7,26	13,94	6,92	13,88	7,42	13,68	7,16	13,60	10,74	17,90	10,20	17,70
$m = 3$	6,50	12,40	6,18	12,16	6,38	11,98	6,22	11,96	8,88	14,90	8,24	14,54
$m = 4$	43,52	56,10	74,44	83,96	43,10	55,82	74,62	84,10	97,58	98,90	100,00	100,00
$m = 6$	36,14	48,86	65,72	76,72	35,56	48,04	65,42	76,98	94,98	97,58	99,98	100,00
$m = 12$	24,12	37,04	50,94	63,76	24,22	37,02	50,80	63,50	87,70	93,14	99,70	99,94
$m = 24$	18,48	28,18	36,66	50,14	18,22	28,02	36,56	49,84	74,88	84,40	98,16	99,18
$m = 36$	15,76	24,42	30,18	42,00	15,68	24,30	29,94	41,68	66,64	77,64	96,16	97,96
Francq-Roy-Zakoian												
$m = 2$	7,34	14,26	7,04	13,92	7,22	13,96	7,22	13,70	11,02	18,04	10,18	18,14
$m = 3$	6,28	12,38	6,24	12,08	6,24	12,16	6,14	11,84	8,34	15,08	7,90	14,44
$m = 4$	43,26	55,50	73,80	84,02	42,96	55,32	73,72	83,82	97,22	98,68	100,00	100,00
$m = 6$	36,22	48,00	65,36	76,66	36,12	47,84	65,32	76,48	94,94	97,40	99,98	100,00
$m = 12$	24,26	37,02	50,38	63,68	24,32	36,60	50,14	63,16	87,46	92,98	99,66	99,90
$m = 24$	18,04	27,44	36,94	49,74	18,02	27,48	36,68	49,82	74,22	83,80	98,06	99,24
$m = 36$	15,12	25,44	29,44	41,76	14,88	25,00	29,36	41,52	65,44	76,88	95,88	97,86
Nouveau test												
$J = 1$	7,78	14,72	7,98	14,72	15,42	25,30	25,24	36,34	41,88	54,60	66,02	75,12
$J = 2$	34,18	48,28	66,40	77,56	33,58	46,56	65,72	76,64	94,12	96,90	99,98	100,00
$J = 3$	10,26	25,04	42,10	59,86	8,36	22,92	37,80	56,20	49,48	81,18	99,04	99,80

l'intuition puisque les nouveaux tests semblent posséder un potentiel accru en terme de puissance comparativement aux autres tests quand α est plus substantiel. De plus, l'effet du coefficient ϕ_4 augmente la puissance des nouveaux tests plus rapidement que pour les autres tests. Ceci suggère qu'une composante saisonnière augmente le potentiel des nouveaux tests par rapport aux autres.

Dans la prochaine section, nous illustrons les tests avec une application utilisant des données réelles provenant de l'économétrie.

Tableau 4.12. Puissances empiriques (en pourcentage) avec niveaux nominaux de 5% et 10% pour des modèles AR(2) \times AR(4) avec des innovations iid, $(\phi_1, \phi_4) \in \{(0,2; 0,1); (0,3; 0,1); (0,3; 0,2)\}$, tailles d'échantillons $n \in \{500, 1000\}$ et nombre d'itérations $N = 5000$.

	$(\phi_2=0,2; \phi_4 = 0,1)$				$(\phi_2=0,3; \phi_4 = 0,1)$				$(\phi_2=0,3; \phi_4 = 0,2)$			
	$n = 500$		$n = 1000$		$n = 500$		$n = 1000$		$n = 500$		$n = 1000$	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
Ljung-Box												
$m = 2$	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
$m = 3$	4,98	10,40	5,66	12,76	5,54	12,28	8,34	19,52	21,40	43,24	65,20	89,26
$m = 4$	45,22	57,30	77,56	85,72	43,64	55,18	75,18	84,56	96,98	98,68	99,98	100,00
$m = 6$	34,80	46,18	65,96	77,70	32,82	44,72	64,34	75,36	92,94	96,30	99,86	99,94
$m = 12$	23,42	34,26	49,62	62,20	22,76	33,16	48,12	60,52	84,76	90,70	99,36	99,70
$m = 24$	17,44	26,10	34,26	47,72	17,00	25,28	33,06	46,18	70,56	79,74	96,92	98,68
$m = 36$	14,36	22,54	27,72	40,28	13,76	22,14	27,10	38,90	62,08	72,74	94,30	96,86
Francq-Roy-Zakoian												
$m = 2$	40,96	53,10	75,12	83,62	35,94	49,14	69,22	80,02	93,86	96,98	99,94	99,98
$m = 3$	5,04	9,78	5,50	12,46	5,30	11,78	7,76	18,24	17,96	39,56	61,58	87,24
$m = 4$	44,18	55,52	76,44	85,64	42,22	53,60	74,44	84,00	96,56	98,42	100,00	100,00
$m = 6$	34,70	45,70	65,82	77,60	32,50	44,22	64,34	75,24	92,80	96,14	99,86	99,96
$m = 12$	22,86	34,30	49,70	62,20	22,38	33,20	47,66	60,38	83,86	90,66	99,36	99,70
$m = 24$	16,88	25,62	34,02	47,56	16,60	24,96	32,74	46,04	70,34	79,32	96,78	98,64
$m = 36$	14,58	22,48	27,64	39,92	14,24	22,22	26,72	38,70	62,24	72,16	93,90	96,70
Nouveau test												
$J = 1$	38,62	51,18	70,16	80,30	41,08	53,30	72,64	82,24	95,58	98,12	99,96	100,00
$J = 2$	32,10	44,40	65,06	75,68	27,42	40,30	59,92	72,74	89,34	93,98	99,84	99,86
$J = 3$	17,72	32,50	48,46	62,02	16,94	30,82	46,44	59,84	74,78	87,68	99,18	99,68

Chapitre 5

Exemple avec données réelles

L'objectif de ce chapitre est d'utiliser les tests avec des données réelles. Pour ce faire, considérons les valeurs quotidiennes de l'indice *Standard & Poor 500* (noté S&P500). Cet indice est bien connu dans la littérature économétrique et financière. Nous considérons comme période temporelle le 3 janvier 1980 jusqu'au 31 août 2010. Ceci donne une taille échantillonnale $n = 7737$. Comme mentionné dans les chapitres précédents, nos applications impliquent souvent de grandes tailles échantillonnales et de ce point de vue les valeurs de n du chapitre 4 peuvent même être considérées modérées.

Les données proviennent du site web de Tsay (2013), et ont été téléchargées à l'adresse:

<http://faculty.chicagobooth.edu/ruey.tsay/teaching/introTS/>.

On note que le site de Tsay est très utile car il contient plusieurs exemples et données associées à l'ouvrage Tsay (2013).

Le modèle ARMA(4,0) avec l'innovation d'un modèle GARCH(1,1) a été ajusté dans Tsay (2013, p. 238) et les résultats obtenus sont:

$$\begin{aligned}r_t &= 0,00055 + 0,0145r_{t-1} - 0,0111r_{t-2} - 0,0221r_{t-3} - 0,034r_{t-4} + \varepsilon_t, \\ \varepsilon_t &= \sigma_t \eta_t, \quad \eta \sim \mathcal{N}(0,1), \\ \sigma_t^2 &= 1,248 \times 10^{-6} + 0,0756\varepsilon_{t-1}^2 + 0,9158\sigma_{t-1}^2,\end{aligned}$$

où $\{r_t\}$ est le rendement et $\{\varepsilon_t\}$ est l'innovation. Seul le coefficient de délai quatre des coefficients AR est significatif au niveau de 5% et les coefficients du GARCH sont tous significatifs. La validité du modèle a été vérifiée par les tests de LB sur les résidus standardisés et les résidus standardisés au carré.

Dans cette section, nous ajustons les rendements selon des modèles AR(0), AR(1), AR(2) et AR(4) et l'innovation selon des modèles GARCH(0,0) et GARCH(1,1). Pour l'innovation de modèle GARCH(1,1), nous analysons les résidus non-standardisés ainsi que les résidus standardisés. Ainsi, nous nous proposons de ré-analyser les données considérées dans Tsay (2013), et nous complétons ses analyses en utilisant les nouveaux tests.

Nous utilisons la fonction `arima()` pour ajuster les modèles d'innovation GARCH(0,0) et la fonction `garchFit()` de la librairie `fGARCH` pour ajuster les modèles d'innovation GARCH(1,1). Nous obtenons les modèles suivants:

Pour GARCH(0,0):

$$AR(0) : r_t = 0,0003 + \varepsilon_t,$$

$$AR(1) : r_t = 0,0003 - 0,0232r_{t-1} + \varepsilon_t,$$

$$AR(2) : r_t = 0,0003 - 0,0245r_{t-1} - 0,0530r_{t-2} + \varepsilon_t,$$

$$AR(4) : r_t = 0,0003 - 0,0242r_{t-1} - 0,0543r_{t-2} - 0,0029^*r_{t-3} - 0,0260r_{t-4} + \varepsilon_t.$$

Pour GARCH(1,1):

$$AR(0) : r_t = 0,00053 + \varepsilon_t,$$

$$\sigma_t^2 = 1,238 \times 10^{-6} + 0,0749\varepsilon_{t-1}^2 + 0,9165\sigma_{t-1}^2,$$

$$AR(1) : r_t = 0,00052 + 0,0156^*r_{t-1} + \varepsilon_t,$$

$$\sigma_t^2 = 1,240 \times 10^{-6} + 0,0750\varepsilon_{t-1}^2 + 0,9164\sigma_{t-1}^2,$$

$$AR(2) : r_t = 0,00053 + 0,0157^*r_{t-1} - 0,0111^*r_{t-2} + \varepsilon_t,$$

$$\sigma_t^2 = 1,240 \times 10^{-6} + 0,0750\varepsilon_{t-1}^2 + 0,9165\sigma_{t-1}^2,$$

$$AR(4) : r_t = 0,00055 + 0,0145^*r_{t-1} - 0,0111^*r_{t-2} - 0,0221^*r_{t-3} - 0,0340r_{t-4} + \varepsilon_t,$$

$$\sigma_t^2 = 1,248 \times 10^{-6} + 0,0756\varepsilon_{t-1}^2 + 0,9158\sigma_{t-1}^2.$$

Les estimateurs avec un astérisque (*) ne sont pas significatifs au niveau 5%.

Les résultats des tests de LB, FRZ et des nouveaux tests sont présentés dans le tableau 5.1. Pour un seuil de signification de 5%, les tests de LB rejettent pour les modèles AR(0) et AR(1), mais seulement à partir de $m = 12$ pour les modèles AR(2) et AR(4). Ces résultats suggèrent que l'hypothèse d'innovation iid pour les modèles AR(0) et AR(1)

est rejetée. Pour les modèles AR(2) et AR(4), on remarque que pour des valeurs de m supérieures ou égales à 12 l'hypothèse nulle d'adéquation est rejetée.

Les nouveaux tests et les tests de FRZ ne rejettent pas l'hypothèse nulle, sauf pour le modèle AR(2), où les tests de FRZ rejettent pour $m = 1$ et $m = 2$. Ces résultats suggèrent que la modélisation d'un modèle AR(0) est possible, mais sous l'hypothèse que l'innovation soit une différence de martingales conditionnellement hétéroscédastique. Nous pouvons dire avec plus de certitude que le bruit est semi-fort, car les nouveaux tests ne rejettent pas l'hypothèse de bruit semi-fort. Si nous ne mettons en oeuvre que les tests FRZ, alors nous ne sommes pas en mesure de conclure que le bruit est semi-fort, dans la mesure où certaines valeurs- p sont encore petites.

Les valeurs des autocorrélations résiduelles sont présentées dans le tableau 5.2. Chaque colonne du tableau est également présentée graphiquement dans les figures 5.1, 5.2 et 5.3 selon les modèles d'innovation GARCH(0,0), GARCH(1,1) avec les résidus non-standardisés ou GARCH(1,1) avec les résidus standardisés, respectivement.

En observant les autocorrélations des résidus non-standardisés des graphiques de la figure 5.2, nous constatons que les ajustements d'un AR(0), AR(1) ou AR(2) ne semblent pas aider à retirer les corrélations. Cependant, l'ajustement d'un AR(4) diminue la valeur de l'autocorrélation de délai 4. Ceci explique en partie la raison pour laquelle seul le coefficient de délai 4 semble être significatif lorsqu'on ajuste des modèles AR(p) avec $p = 0,1,2,4$.

En observant les autocorrélations des résidus standardisés des graphiques de la figure 5.3, nous constatons qu'il reste de la dépendance de délai un quand on ajuste un modèle AR(0). Cependant, le coefficient de délai un n'est pas significatif lorsqu'on ajuste des modèles AR(p) avec $p = 1,2,4$.

En conclusion, nos analyses complètent le modèle AR(4) proposé par Tsay (2013). Nous concluons également que les erreurs sont hétéroscédastiques. Un modèle AR(4) avec erreurs GARCH(1,1) devrait convenir. Les conclusions des tests FRZ et de nos tests sont les mêmes. Compte tenu que nos tests étaient puissants avec des modèles saisonniers, ceci donne une évidence additionnelle en faveur de ce modèle.

Tableau 5.1. Valeurs-p des différents tests pour les résidus des rendements du S&P500 avec l'innovation suivant un modèle GARCH(0,0) et les rendements suivant des modèles (a) AR(0), (b) AR(1), (c) AR(2) et (d) AR(4).

(a)										
m	1	2	3	4	6	12	18	24	30	36
Statistique de LB	4,1781	25,4210	25,6970	29,8945	30,3409	51,9861	75,8314	85,2207	97,3950	135,8615
LB	0,0410	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
FRZ	0,3623	0,1245	0,1911	0,2006	0,4010	0,4258	0,3524	0,4545	0,5111	0,2836
J	1	2	3	4						
Nouvelle statistique	3,5812	4,7495	5,5719	18,6715						
Nouveau test	0,1669	0,3140	0,6951	0,2861						
(b)										
m	1	2	3	4	6	12	18	24	30	36
Statistique de LB	0,0117	21,6277	21,7654	25,9702	26,4068	47,2207	70,5618	79,4698	91,6600	128,9990
LB	NA	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000
FRZ	0,0392	0,0373	0,0812	0,0957	0,2892	0,3815	0,3371	0,4034	0,4220	0,2817
J	1	2	3	4						
Nouvelle statistique	3,0502	5,1374	6,0498	19,8922						
Nouveau test	0,2176	0,2735	0,6417	0,2251						
(c)										
m	1	2	3	4	6	12	18	24	30	36
Statistique de LB	0,0003	0,0133	0,0627	5,2821	5,7687	27,3420	48,2963	56,9238	70,0995	104,9305
LB	NA	NA	0,8022	0,0713	0,2171	0,0023	0,0000	0,0001	0,0000	0,0000
FRZ	0,8585	0,4541	0,8915	0,4211	0,8171	0,6192	0,5148	0,5880	0,5766	0,3750
J	1	2	3	4						
Nouvelle statistique	0,1359	1,0976	7,5882	16,8993						
Nouveau test	0,9343	0,8947	0,4747	0,3921						
(d)										
m	1	2	3	4	6	12	18	24	30	36
Statistique de LB	0,0002	0,0003	0,0078	0,0080	0,5085	23,5531	45,3978	54,2502	67,3543	102,2939
LB	NA	NA	NA	NA	0,7755	0,0027	0,0000	0,0001	0,0000	0,0000
FRZ	0,8035	0,9562	0,6397	0,8216	0,9459	0,5596	0,4508	0,5326	0,5271	0,3282
J	1	2	3	4						
Nouvelle statistique	0,4378	5,4791	8,4183	16,1120						
Nouveau test	0,8034	0,2416	0,3937	0,4452						

Tableau 5.2. Valeurs d'autocorrélation des résidus des rendements du S&P500 suivant des modèles AR avec l'innovation suivant des modèles GARCH.

Lag	GARCH(0,0)				GARCH(1,1) non-standardisé				GARCH(1,1) standardisé			
	AR(0)	AR(1)	AR(2)	AR(4)	AR(0)	AR(1)	AR(2)	AR(4)	AR(0)	AR(1)	AR(2)	AR(4)
0	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
1	-0,0232	-0,0012	0,0002	-0,0002	-0,0232	-0,0380	-0,0385	-0,0381	0,0256	0,0117	0,0114	0,0115
2	-0,0524	-0,0528	-0,0013	0,0001	-0,0524	-0,0523	-0,0416	-0,0440	-0,0076	-0,0078	0,0020	0,0014
3	0,0060	0,0042	0,0025	-0,0010	0,0060	0,0071	0,0068	0,0267	-0,0179	-0,0173	-0,0173	0,0021
4	-0,0233	-0,0233	-0,0260	0,0001	-0,0233	-0,0234	-0,0239	0,0091	-0,0196	-0,0192	-0,0194	0,0110
5	-0,0058	-0,0062	-0,0072	-0,0078	-0,0058	-0,0055	-0,0057	-0,0076	-0,0181	-0,0177	-0,0180	-0,0183
6	0,0050	0,0043	0,0032	0,0021	0,0050	0,0054	0,0052	0,0038	-0,0100	-0,0096	-0,0099	-0,0107
7	-0,0245	-0,0243	-0,0247	-0,0252	-0,0245	-0,0247	-0,0248	-0,0256	-0,0092	-0,0090	-0,0092	-0,0101
8	0,0048	0,0042	0,0050	0,0056	0,0048	0,0052	0,0054	0,0051	-0,0066	-0,0063	-0,0062	-0,0068
9	0,0007	0,0010	-0,0015	-0,0018	0,0007	0,0005	-0,0001	0,0007	-0,0059	-0,0059	-0,0061	-0,0059
10	0,0106	0,0100	0,0123	0,0125	0,0106	0,0109	0,0114	0,0110	0,0153	0,0154	0,0157	0,0155
11	-0,0239	-0,0228	-0,0225	-0,0238	-0,0239	-0,0248	-0,0247	-0,0267	-0,0056	-0,0064	-0,0062	-0,0077
12	0,0386	0,0382	0,0386	0,0398	0,0386	0,0389	0,0390	0,0398	0,0303	0,0302	0,0303	0,0296

Figure 5.1. Graphique d'autocorrélation des résidus pour les rendements du S&P500 avec l'innovation suivant un modèle GARCH(0,0) et les rendements suivant des modèles (a) AR(0), (b) AR(1), (c) AR(2) et (d) AR(4).

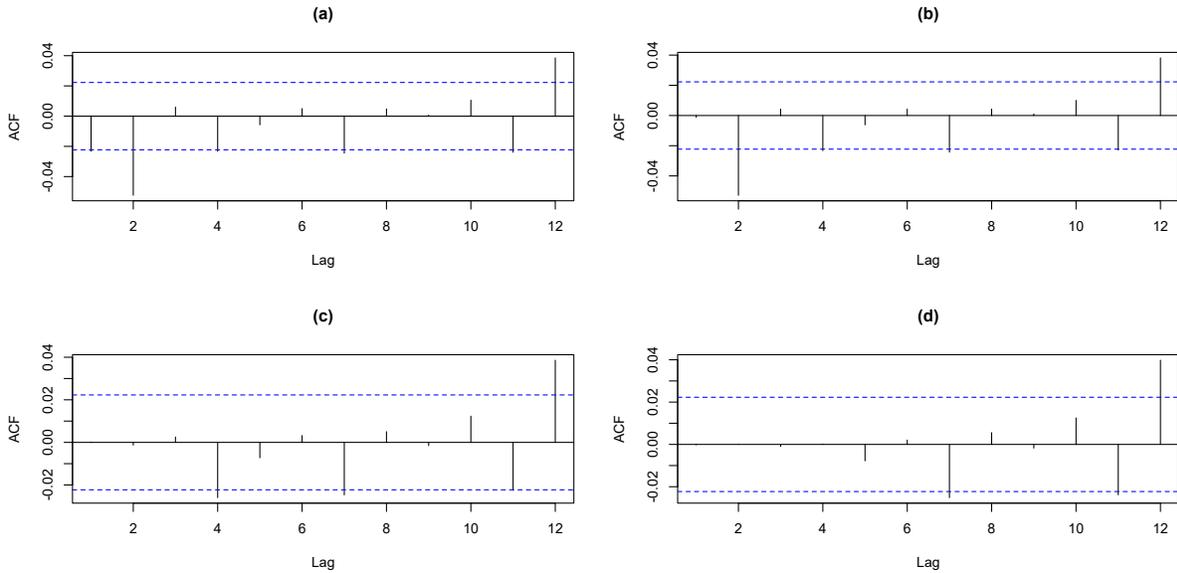


Figure 5.2. Graphique d'autocorrélation des résidus non-standardisés pour les rendements du S&P500 avec l'innovation suivant un modèle GARCH(1,1) et les rendements suivant des modèles (a) AR(0), (b) AR(1), (c) AR(2) et (d) AR(4).

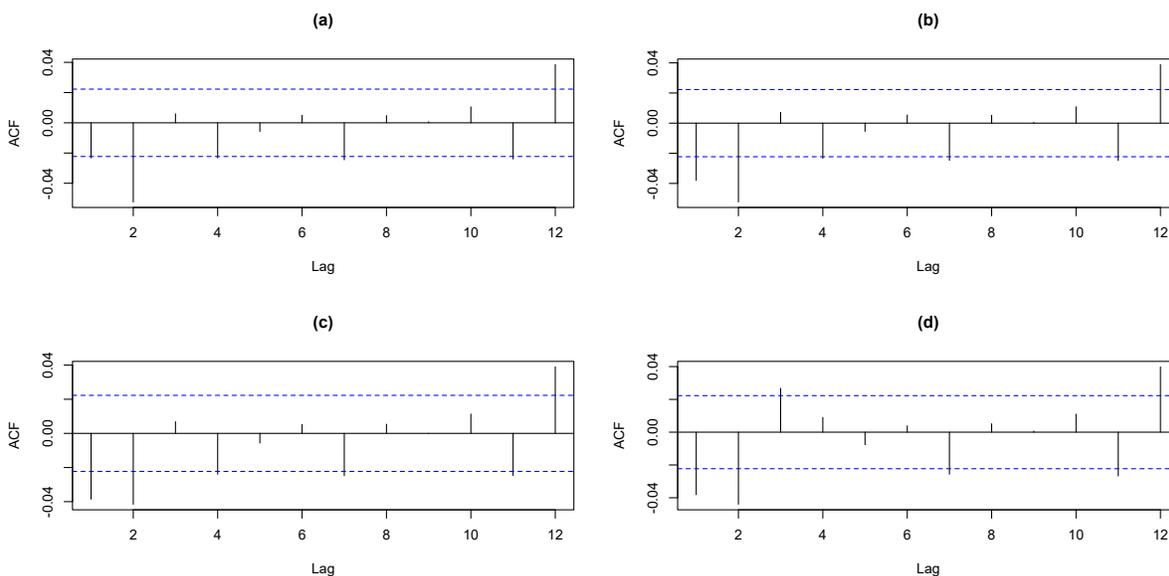
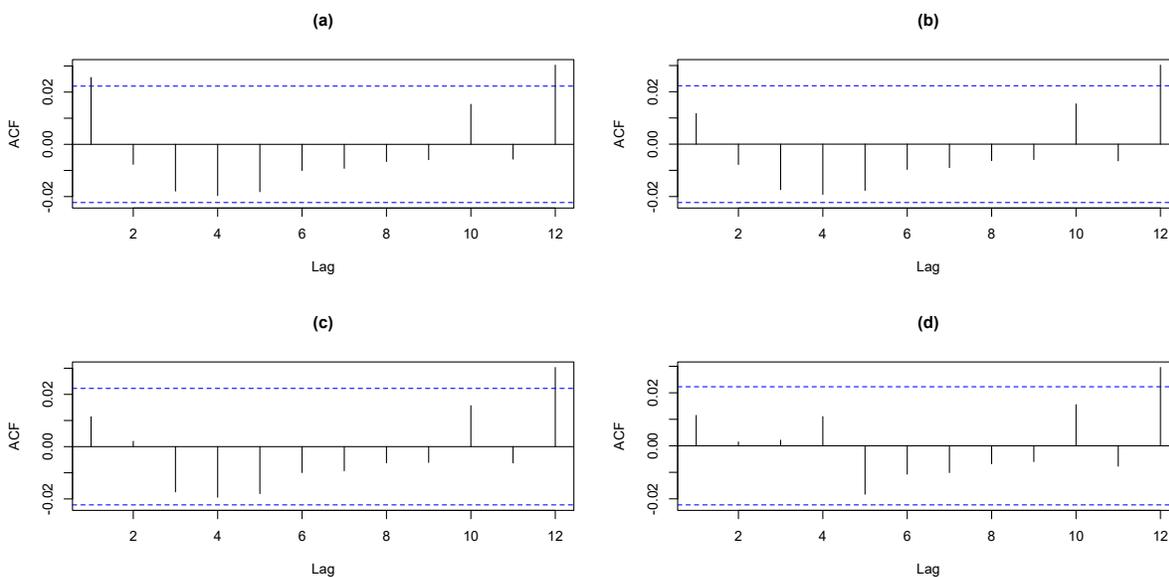


Figure 5.3. Graphique d'autocorrélation des résidus standardisés pour les rendements du S&P500 avec l'innovation suivant un modèle GARCH(1,1) et les rendements suivant des modèles (a) AR(0), (b) AR(1), (c) AR(2) et (d) AR(4).



Conclusion

Dans le mémoire, nous avons développé des nouveaux tests à base d'ondelettes afin de tester la qualité d'ajustement d'un modèle ARMA où l'innovation est une différence de martingales avec hétéroscédasticité conditionnelle. Les nouveaux tests reposent sur les résidus du modèle ajusté. Pour un modèle bien ajusté, les autocovariances empiriques des résidus suivent conjointement une distribution normale. De ceux-ci, nous déduisons que les distributions asymptotiques des coefficients d'ondelettes de la densité spectrale de l'expansion d'ondelettes de l'innovation suivent aussi une distribution normale, et leur variance et covariance asymptotiques dépendent des modèles ARMA ainsi que de la structure de covariance du moment d'ordre deux des innovations. Les nouveaux tests reposent sur une fonction d'ondelette et nous avons énoncé les conditions précises que l'ondelette doit satisfaire dans la mise en oeuvre du test: dans notre cas, l'ondelette de Franklin nous permet d'obtenir les résultats techniques. Les nouvelles statistiques de test suivent une distribution de type chi-carré.

De nos simulations de Monte Carlo, nous avons montré que les tests de Box-Pierce-Ljung classiques ne sont pas valides pour les ARMA semi-fort, tandis que les niveaux empiriques des nouveaux tests convergent rapidement vers les niveaux nominaux pour les modèles ARMA semi-fort ou iid. Cependant, elles ne convergent pas aussi vite que les tests de Francq-Roy-Zakoïan qui en plus sont valides sur les processus avec bruits blancs faibles. L'avantage des nouveaux tests repose sur le gain en puissance pour les séries avec beaucoup d'irrégularités et les séries saisonnières. Ceci a été illustré dans nos simulations quand nous augmentons le coefficient α des modèles GARCH et le coefficient ϕ_4 des modèles multiplicatifs, respectivement. De plus, nous n'avons besoin de calculer la

statistique que pour quelques valeurs de J contrairement aux autres tests qui nécessitent le calcul des statistiques pour plusieurs valeurs de m .

Nous avons utilisé les différents tests afin de modéliser les rendements de la série chronologique financière S&P500. Les résultats des nouveaux tests et des tests de Francq-Roy-Zakoïan sont plutôt similaires. Cependant, sans les nouveaux tests, il est plus difficile de déduire que l'innovation est semi-fort. Donc, avec simplement les tests de Francq-Roy-Zakoïan, des démarches supplémentaires sont nécessaires afin d'avoir une évidence que l'innovation est un bruit blanc semi-fort.

Pour conclure, pour les processus de modèle ARMA avec des innovations qui ne sont pas indépendantes, les tests de Box-Pierce-Ljung ne sont pas adéquats. L'utilisation des nouveaux tests portemanteaux développés dans le mémoire sont recommandés quand les innovations sont des différences de martingale avec hétéroscédasticité conditionnelle et ils offrent de bonnes propriétés de puissance pour les séries saisonnières et quand les densités spectrales possèdent moins de régularité ou plus d'inhomogénéité.

Annexe

Preuve du lemme 2.4.2. Puisque la preuve est très similaire à la preuve du lemme 2 dans McLeod (1978, p. 298) et la preuve du théorème 1 dans Francq et al. (2005, p. 542), nous ne fournissons que les grandes lignes de l'argument. Soit $\partial S_n/\partial \boldsymbol{\theta}_0$ et $\partial S_n/\partial \hat{\boldsymbol{\theta}}$ les vecteurs de dérivées partielles de $S_n(\boldsymbol{\theta})$ (i.e., $\partial S_n(\boldsymbol{\theta})/\partial \boldsymbol{\theta}$) évalués à $\boldsymbol{\theta}_0$ et $\hat{\boldsymbol{\theta}}$, respectivement. Sous l'hypothèse nulle, il peut être montré que $\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 = \mathbf{O}_p(1/\sqrt{n})$ et la dérivée partielle de $S_n(\boldsymbol{\theta})$ de tout ordre est $\mathbf{O}_p(1/\sqrt{n})$. Alors, une expansion de Taylor de $\partial S_n(\boldsymbol{\theta})/\partial \boldsymbol{\theta}$ au point $\boldsymbol{\theta}_0$ et évaluée à $\hat{\boldsymbol{\theta}}$, donne le développement suivant:

$$\mathbf{0} = \frac{\partial S_n}{\partial \boldsymbol{\theta}_0} + \frac{\partial^2 S_n}{\partial \boldsymbol{\theta}_0 \partial \boldsymbol{\theta}_0^\top} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) + \mathbf{O}_p(1/n). \quad (5.0.1)$$

Comme dans Francq et al. (2005, p. 543), la dérivée des bruits peut être représentée par:

$$\frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = (v_{t-1}(\boldsymbol{\theta}), \dots, v_{t-p}(\boldsymbol{\theta}), u_{t-1}(\boldsymbol{\theta}), \dots, u_{t-q}(\boldsymbol{\theta}))^\top, \quad (5.0.2)$$

où $v_t(\boldsymbol{\theta}) = -\phi_{\boldsymbol{\theta}}^{-1}(B)\varepsilon_t(\boldsymbol{\theta})$, $u_t(\boldsymbol{\theta}) = \psi_{\boldsymbol{\theta}}^{-1}(B)\varepsilon_t(\boldsymbol{\theta})$ et $\varepsilon_t(\boldsymbol{\theta}) = \psi_{\boldsymbol{\theta}}^{-1}(B)\phi_{\boldsymbol{\theta}}(B)X_t$. Quand $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, nous obtenons $\partial \varepsilon_t/\partial \boldsymbol{\theta}_0 = \sum_{i \geq 1} \varepsilon_{t-i} \boldsymbol{\lambda}_i$. Utilisant le résultat $\|\varepsilon_t(\boldsymbol{\theta}) - e_t(\boldsymbol{\theta})\| \leq K\rho^t$ pour des constantes $K > 0$ and $0 < \rho < 1$ et $\|e_t(\boldsymbol{\theta})\| \leq k$ pour une constante $k > 0$, nous avons, à part le terme $\mathbf{O}_p(1/n)$ qui est négligé:

$$\frac{1}{2} \frac{\partial S_n}{\partial \boldsymbol{\theta}_0} = \frac{1}{n} \sum_{t=1}^n \varepsilon_t \frac{\partial \varepsilon_t}{\partial \boldsymbol{\theta}_0} + \mathbf{O}_p(1/n) = \frac{1}{n} \sum_{t=1}^n \sum_{i \geq 1} \varepsilon_t \varepsilon_{t-i} \boldsymbol{\lambda}_i + \mathbf{O}_p(1/n). \quad (5.0.3)$$

Écrivons $(1/2)\partial S_n(\boldsymbol{\theta})/\partial \boldsymbol{\theta} = n^{-1} \sum_{t=1}^n \varepsilon_t(\boldsymbol{\theta}) \partial \varepsilon_t(\boldsymbol{\theta})/\partial \boldsymbol{\theta} + \mathbf{O}_p(1/n)$. Utilisant l'expression de $\partial \varepsilon_t(\boldsymbol{\theta})/\partial \boldsymbol{\theta}$ dans (5.0.2), avec un peu d'algèbre et en appliquant le théorème 1 dans Hannan et Heyde (1972, p. 2060), nous pouvons obtenir le résultat suivant:

$$\frac{1}{2} \frac{\partial^2 S_n}{\partial \boldsymbol{\theta}_0 \partial \boldsymbol{\theta}_0^\top} = \mathbf{J} + \mathbf{O}_p(1/\sqrt{n}),$$

où la matrice \mathbf{J} de dimensions $(p+q) \times (p+q)$ est:

$$\mathbf{J} = \begin{pmatrix} \gamma_{vv}(i-j) & \gamma_{vu}(i-j) \\ \gamma_{uv}(i-j) & \gamma_{uu}(i-j) \end{pmatrix} = \sigma^2 \sum_{i=1}^{\infty} \boldsymbol{\lambda}_i \boldsymbol{\lambda}_i^\top = \sigma^2 \boldsymbol{\Lambda}_\infty \boldsymbol{\Lambda}_\infty^\top.$$

Les éléments (i, j) dans chaque partition de la matrice est indiquée et $\gamma_{vv}(k) = \mathbb{E}(v_t v_{t+k})$, $\gamma_{uu}(k) = \mathbb{E}(u_t u_{t+k})$ et $\gamma_{vu}(k) = \mathbb{E}(v_t u_{t+k})$. Les $v_t = v_t(\boldsymbol{\theta}_0)$ et $u_t = u_t(\boldsymbol{\theta}_0)$ sont fournis dans (5.0.2). Ensuite, de (5.0.1) et (5.0.3), ceci montre le lemme.

Preuve du lemme 2.4.3. La preuve est similaire à celle du lemme 3 dans McLeod (1978, p. 299). Puisque toutes les combinaisons linéaires des éléments de $\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0$ et \mathbf{r}_m sont, à part du terme $\mathbf{O}_p(1/n)$, une moyenne d'une série de différence de martingales, il peut être montré que \sqrt{n} fois cette combinaison linéaire est asymptotiquement normale par le théorème centrale limite pour les martingales. Du lemme 2.4.2, la covariance asymptotique de $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ et $\sqrt{nr}(l), l \geq 1$ est:

$$\text{Cov}(\sqrt{n}\mathbf{J}^{-1}\mathbf{Y}_n, \sqrt{nr}(l)) = -\mathbf{J}^{-1} \frac{1}{n} \sum_{t=1}^n \sum_{t'=1}^{n-l} \mathbb{E}(\varepsilon_t \frac{\partial \varepsilon_t}{\partial \boldsymbol{\theta}_0} \varepsilon_{t'} \varepsilon_{t'+l}).$$

Utilisant $\partial \varepsilon_t / \partial \boldsymbol{\theta}_0 = \sum_{i \geq 1} \varepsilon_{t-i} \boldsymbol{\lambda}_i$ dans le lemme 2.4.2, il peut être montré que:

$$\text{Cov}(\sqrt{n}\mathbf{J}^{-1}\mathbf{Y}_n, \sqrt{nr}(l)) \longrightarrow -\mathbf{J}^{-1} \sum_{i \geq 1} \Gamma(i, l) \boldsymbol{\lambda}_i,$$

quand $n \rightarrow \infty$, sous l'hypothèse A. Donc, nous avons:

$$\text{Cov}(\sqrt{n}\mathbf{J}^{-1}\mathbf{Y}_n, \sqrt{nr}_m) = -\mathbf{J}^{-1} \boldsymbol{\Lambda}_\infty \boldsymbol{\Gamma}_{\infty, m},$$

ce qui donne le lemme.

Preuve du lemme 2.4.4. Similaire à la preuve du théorème 2 dans Francq et al. (2005, p. 543), pour $l = 1, 2, \dots, m$, nous obtenons:

$$\begin{aligned} \hat{r}(l) &= r(l) + E\left(\varepsilon_t \frac{\partial \varepsilon_{t+l}}{\partial \boldsymbol{\theta}_0}\right) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) + R_n + O_p(1/n), \\ &= r(l) + \sigma^2 \boldsymbol{\lambda}_l^\top (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) + R_n + O_p(1/n), \end{aligned}$$

où $R_n = n^{-1} \sum_{t=1}^{n-l} \{e_t(\hat{\boldsymbol{\theta}}) e_{t+l}(\hat{\boldsymbol{\theta}}) - \varepsilon_t(\hat{\boldsymbol{\theta}}) \varepsilon_{t+l}(\hat{\boldsymbol{\theta}})\}$. De $\|\varepsilon_t(\boldsymbol{\theta}) - e_t(\boldsymbol{\theta})\| \leq K\rho^t$ et $\|e_t(\boldsymbol{\theta})\| \leq k$, pour des constantes $K, k > 0$ et $0 < \rho < 1$, il peut être montré que $R_n = O_p(1/n)$. Donc,

nous pouvons écrire:

$$\hat{\mathbf{r}}_m = \mathbf{r}_m + \sigma^2 \mathbf{\Lambda}_m^\top (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) + \mathbf{O}_p(1/n).$$

Du lemme 2.4.3 et du *Continuous Mapping Theorem*, ceci permet de montrer le résultat.

Preuve des lemmes 2.6.1 et 2.6.2. Compte tenu de l'expression pour $\hat{\beta}_{jk}$ dans (2.6.13), il est facile de montrer que $\hat{\beta}_{jk_1} = \hat{\beta}_{jk_2}$ tant que $k_1 + k_2 = 2^j + 1$. La preuve pour les β_{jk} est similaire et est donc omise.

Preuve du théorème 2.7.1. En vue de (2.6.11), nous écrivons $\sqrt{n} \hat{\beta}_{jk}$ tel que:

$$\sqrt{n} \hat{\beta}_{jk} = \sum_{h=1}^m \sqrt{n} \hat{r}_\varepsilon(h) b_{jk}(h) + \sum_{h=m+1}^{n-1} \sqrt{n} \hat{r}_\varepsilon(h) b_{jk}(h) = \mathbf{B}_{jk,m}^\top \sqrt{n} \hat{\mathbf{r}}_m + Q_{n,m}.$$

Du lemme 2.4.4, nous avons, pour $m \geq 1$ fixé, $\mathbf{B}_{jk,m}^\top \sqrt{n} \hat{\mathbf{r}}_m \rightarrow_d \mathcal{N}(\mathbf{0}, \mathbf{B}_{jk,m}^\top \boldsymbol{\Sigma}_{\hat{\mathbf{r}}_m} \mathbf{B}_{jk,m}^\top)$, quand $n \rightarrow \infty$. Nous avons également $\mathbf{B}_{jk,m}^\top \boldsymbol{\Sigma}_{\hat{\mathbf{r}}_m} \mathbf{B}_{jk,m}^\top \rightarrow \sigma_{jk}^2$, quand $m \rightarrow \infty$. Pour le terme restant $Q_{n,m}$, il peut être montré que $\sup_{1 \leq h < n} \text{Var}\{\hat{r}_\varepsilon(h)\} = O(n^{-1})$ sous l'hypothèse A. Donc, nous avons:

$$E(Q_{n,m}^2) \leq C \sum_{h_1=m+1}^{n-1} \sum_{h_2=m+1}^{n-1} |b_{jk}(h_1)| |b_{jk}(h_2)| = C \left(\sum_{h=m+1}^{n-1} |b_{jk}(h)| \right)^2.$$

Puisque $|\hat{\eta}(w)| \leq C(1 + |w|)^{-2}$ avec l'ondelette de Franklin η dans (2.5.7) et utilisant $\hat{\eta}_{jk}(2\pi h) = e^{-i2\pi h k / 2^j} 2^{-j/2} \hat{\eta}(2\pi h / 2^j)$, nous avons $|\hat{\eta}_{jk}(2\pi h)| \leq C 2^{3j/2} (2^j + 2\pi h)^{-2}$. Ce qui implique que nous avons $|b_{jk}(h)| \leq C 2^{3j/2} (2^j + 2\pi h)^{-2}$ aussi. Donc nous avons $\sum_h |b_{jk}(h)| < \infty$, pour j et k fixés et nous avons $\lim_{m \rightarrow \infty} \limsup_n E(Q_{n,m}^2) = 0$. De l'inégalité de Markov, nous obtenons que $\lim_{m \rightarrow \infty} \limsup_n Q_{n,m} = 0$ en probabilité. En appliquant la proposition 6.3.9 de Brockwell et Davis (1991, p. 207), nous prouvons que $\sqrt{n} \hat{\beta}_{jk} \rightarrow \mathcal{N}(0, \sigma_{jk}^2)$. De même, nous pouvons montrer le terme de covariance.

Preuve du lemme 3.3.2. De la définition de $\hat{\varepsilon}_t = e_t(\hat{\boldsymbol{\theta}})$ et $\varepsilon_t = \varepsilon_t(\boldsymbol{\theta}_0)$, nous écrivons

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^n \hat{\varepsilon}_t^2 = \frac{1}{n} \sum_{t=1}^n (e_t^2(\hat{\boldsymbol{\theta}}) - \varepsilon_t^2(\hat{\boldsymbol{\theta}})) + \frac{1}{n} \sum_{t=1}^n (\varepsilon_t^2(\hat{\boldsymbol{\theta}}) - \varepsilon_t^2(\boldsymbol{\theta}_0)) + \frac{1}{n} \sum_{t=1}^n \varepsilon_t^2(\boldsymbol{\theta}_0). \quad (5.0.4)$$

De $\|\varepsilon_t(\boldsymbol{\theta}) - e_t(\boldsymbol{\theta})\| \leq K\rho^t$ et $\|e_t(\boldsymbol{\theta})\| \leq k$, pour des constantes $K, k > 0$ et $0 < \rho < 1$, le premier terme du côté droit de (5.0.4) est $O_p(1/n)$. De $\hat{\boldsymbol{\theta}} \rightarrow_p \boldsymbol{\theta}_0$, le deuxième terme

du côté droit de (5.0.4) est $O_p(1)$. Le dernier terme converge vers σ^2 en probabilité sous l'hypothèse A. Du théorème de Slutsky, nous prouvons que $\hat{\sigma}^2 \rightarrow_p \sigma^2$. La preuve que $\hat{\Gamma}(l, l') = n^{-1} \sum_{t=1}^n \hat{\varepsilon}_t^2 \hat{\varepsilon}_{t-l} \hat{\varepsilon}_{t-l'} \rightarrow_p \sigma^4 \tau_{l, l'}$ est similaire sous l'hypothèse A.

Bibliographie

- Anderson, T. W. (1993). ‘Goodness of fit tests for spectral distributions’. *The Annals of Statistics* **21**, 830–847.
- Andrews, D. et Ploberger, W. (1996). ‘Testing for serial correlation against an ARMA(1,1) process’. *Journal of the American Statistical Association* **91**, 1331–1342.
- Berk, K. N. (1974). ‘Consistent autoregressive spectral estimates’. *The Annals of Statistics* **2**, 489–502.
- Box, G. E. P. et Pierce, D. A. (1970). ‘Distribution of residual autocorrelations in autoregressive integrated moving average time series models’. *Journal of the American Statistical Association* **65**, 1509–1526.
- Brockwell, P. J. et Davis, R. A. (1991). *Time Series: Theory and Methods*. 2^{ième} Édition. Springer-Verlag, New York.
- Brockwell, P. J. et Davis, R. A. (2016). *Introduction to Time Series and Forecasting*. 3^{ième} Édition. Springer, New York.
- Daubechies, I. (1992). *Ten Lectures on Wavelets*. SIAM, Philadelphia.
- Davies, N., Triggs, C. M. et Newbold, P. (1977). ‘Significance levels of the Box-Pierce portmanteau statistic in finite samples’. *Biometrika* **64**, 517–522.
- Deo, R. S. (2000). ‘Spectral tests of the martingale hypothesis under conditional heteroscedasticity’. *Journal of Econometrics* **99**, 291–315.
- Duchesne, P. et Lafaye De Micheaux, P. (2010). ‘Computing the distribution of quadratic forms: further comparison between the Liu-Tang-Zhang approximation and exact methods’. *Computational Statistics & Data Analysis* **54**, 858–862.
- Duchesne, P., Li, L. et Vandermeerschen, J. (2010). ‘On testing for serial correlation of unknown form using wavelet thresholding’. *Computational Statistics & Data Analysis*

54, 2512–2531.

- Durbin, J. (1960). ‘The fitting of time-series models’. *Review of the International Statistical Institute* **28**, 233–244.
- Durlauf, S. N. (1991). ‘Spectral based testing of the martingale hypothesis’. *Journal of Econometrics* **50**, 355–376.
- Escanciano, J. C. et Lobato, I. N. (2009). ‘An automatic portmanteau test for serial correlation’. *Journal of Econometrics* **151**, 140–149.
- Fan, J. (1996). ‘Test of significance based on wavelet thresholding and Neyman’s truncation’. *Journal of the American Statistical Association* **91**, 674–688.
- Francq, C. et Zakoïan, J. (1998). ‘Estimating linear representations of nonlinear processes’. *Journal of Statistical Planning and Inference* **68**, 145–165.
- Francq, C., Roy, R. et Zakoïan, J. (2005). ‘Diagnostic checking in ARMA models with uncorrelated errors’. *Journal of the American Statistical Association* **100**, 532–544.
- Francq, C. et Zakoïan, J. (2009). *Modèles GARCH : structure, inférence statistique et applications financières*. Economica, Paris.
- Francq, C. et Zakoïan, J. (2010). *GARCH Models: Structure, Statistical Inference and Financial Applications*. John Wiley & Sons Ltd., Hoboken, New Jersey.
- Fuller, W. A. (1996). *Introduction to Statistical Time Series*. 2^{ième} Édition. Wiley Series in Probability and Statistics, John Wiley & Sons Inc., New York.
- Gençay, R. et Signori, D. (2015). ‘Multi-scale tests for serial correlation’. *Journal of Econometrics* **184**, 62–80.
- Hannan, E. J. et Heyde, C. C. (1972). ‘On limit theorems for quadratic functions of discrete time series’. *The Annals of Mathematical Statistics* **43**, 2058–2066.
- Hernández, E. et Weiss, G. (1996). *A First Course on Wavelets*. CRC Press, Boca Raton, FL.
- Herrndorf, N. (1984). ‘A functional central limit theorem for weakly dependent sequences of random variables’. *The Annals of Probability* **12**, 141–153.
- Hong, Y. (1996). ‘Consistent testing for serial correlation of unknown form’. *Econometrica* **64**, 837–864.

- Hong, Y. (1999). ‘Hypothesis testing in time series via the empirical characteristic function: a generalized spectral density approach’. *Journal of the American Statistical Association* **94**, 1201–1220.
- Hong, Y. et Kao, C. (2004). ‘Wavelet-based testing for serial correlation of unknown form in panel models’. *Econometrica* **72**, 1519–1563.
- Horowitz, J. L., Lobato, I. N., Nankervis, J. C. et Savin, N. E. (2006). ‘Bootstrapping the Box-Pierce Q-test: A robust test of uncorrelatedness’. *Journal of Econometrics* **133**, 841–862.
- Imhof, J. P. (1961). ‘Computing the distribution of quadratic forms in normal variables’. *Biometrika* **48**, 419–426.
- Lee, J. et Hong, Y. (2001). ‘Testing for serial correlation of unknown form using wavelet methods’. *Econometric Theory* **17**, 386–423.
- Li, L., Yao, S. et Duchesne, P. (2014). ‘On wavelet-based testing for serial correlation of unknown form using Fan’s adaptive Neyman method’. *Computational Statistics & Data Analysis* **70**, 308–327.
- Li, W. K. et McLeod, A. I. (1981). ‘Distribution of the residual autocorrelations in multivariate ARMA time series models’. *Journal of the Royal Statistical Society, Series B* **43**, 231–239.
- Li, W. K. (2004). *Diagnostic Checks in Time Series*. Chapman & Hall/CRC, New York.
- Ljung, G. M. et Box, G. E. P. (1978). ‘On a measure of lack of fit in time series models’. *Biometrika* **65**, 297–303.
- Ljung, G. M. (1986). ‘Diagnostic testing of univariate time series models’. *Biometrika* **73**, 725–730.
- Lobato, I. N. et Robinson, P. M. (1998). ‘A nonparametric test for I(0)’. *The Review of Economic Studies* **65**, 475–495.
- Lobato, I. N., Nankervis, J. C. et Savin, N. E. (2001). ‘Testing for autocorrelation using a modified Box-Pierce Q test’. *International Economic Review* **42**, 187–205.
- Lobato, I. N., Nankervis, J. C. et Savin, N. E. (2002). ‘Testing for zero autocorrelation in the presence of statistical dependence’. *Econometric Theory* **18**, 730–743.

- Mallat, S. (1999). *A Wavelet Tour of Signal Processing*. Academic Press Inc., San Diego, CA.
- McLeod, A. I. (1978). ‘On the distribution of residual autocorrelations in Box-Jenkins method’. *Journal of the Royal Statistical Society, Series B* **40**, 296–302.
- Paparoditis, E. (2000). ‘Spectral density based goodness-of-fit tests for time series models’. *Scandinavian Journal of Statistics* **27**, 143–176.
- Priestley, M. B. (1996). ‘Wavelets and time-dependent spectral analysis’. *Journal of Time Series Analysis* **17**, 85–103.
- Robinson, P. M. (1991). ‘Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression’. *Journal of Econometrics* **47**, 67–84.
- Rosenblatt, M. (1956). ‘A central limit theorem and a strong mixing condition’. *Proceedings of the National Academy of Sciences, U.S.A.* **42**, 43–47.
- Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons Ltd., New York.
- Shao, X. (2011). ‘A bootstrap-assisted spectral test of white noise under unknown dependence’. *Journal of Econometrics* **162**, 213–224.
- Shumway, R. H. et Stoffer, D. S. (2011). *Time Series Analysis and Its Applications: With R Examples*. 3^{ième} Édition. Springer, New York.
- Tsay, R. S. (2013). *An Introduction to Analysis of Financial Data with R*. John Wiley & Sons Inc., Hoboken, New Jersey.
- Vidakovic, B. (1999). *Statistical Modeling by Wavelets*. Wiley Series in Probability and Statistics, John Wiley & Sons Inc., New York.
- Whittle, P. (1963). ‘On the fitting of multivariate autoregressions and the approximate canonical factorization of a spectral density matrix’. *Biometrika* **40**, 129–134.