



**HAL**  
open science

# Validation des modèles ARMA saisonniers et/ou périodiques avec innovations dépendantes

Abdoulkarim Ilmi Amir

► **To cite this version:**

Abdoulkarim Ilmi Amir. Validation des modèles ARMA saisonniers et/ou périodiques avec innovations dépendantes. Probabilités [math.PR]. Université Bourgogne Franche-Comté, 2021. Français. NNT : 2021UBFCD015 . tel-03346451

**HAL Id: tel-03346451**

**<https://theses.hal.science/tel-03346451>**

Submitted on 16 Sep 2021

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

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

# Validation des modèles ARMA saisonniers et/ou périodiques avec innovations dépendantes

Thèse de doctorat de  
l'Université Bourgogne Franche-Comté

École doctorale CARNOT-PASTEUR

présentée et soutenue publiquement le 20 mai 2021  
en vue de l'obtention du grade de

Docteur de l'Université Bourgogne Franche-Comté

discipline : Mathématiques

par

**Abdoulkarim ILMI AMIR**

## *Composition du jury :*

<i>Rapporteurs :</i>	Mme Sophie DABO-NIANG M. ABDELOUAHAB Bibi	Université de Lille Université Larbi Ben M'Hidi, Oum EL Bouaghi
<i>Présidente :</i>	Mme Camélia GOGA	Université Bourgogne Franche-Comté
<i>Examineurs :</i>	M. Eugen URSU M. Ahmed EL GHINI	Université de Bordeaux Université de Mohammed V de Rabat
<i>Directeur de thèse :</i>	M. Yacouba BOUBACAR MAÏNASSARA	Université Bourgogne Franche-Comté

# Remerciements

Je voudrais tout d'abord remercier grandement mon directeur de thèse, Monsieur Yacouba BOUBACAR MAÏNASSARA, pour toute son aide. Je suis ravi d'avoir travaillé en sa compagnie car outre son appui scientifique, il a toujours été là pour me soutenir et me conseiller au cours de l'élaboration de cette thèse.

J'adresse tous mes remerciements à Monsieur le Professeur ABDELOUAHAB Bibi et à Madame la Professeure Sophie DABO-NIANG de l'honneur qu'ils me font d'avoir accepté d'être les rapporteurs de cette thèse et de l'intérêt porté à l'égard de mon travail.

Je remercie également Monsieur le Professeur Ahmed EL GHINI, Madame la Professeure Camélia GOGA et Monsieur Eugen URSU qui ont bien voulu être examinateurs.

Un grand merci à Madame la Professeure Sophie DABO-NIANG et Monsieur Ulrich RAZAFISON d'avoir constitué mon comité de suivi de thèse.

J'adresse mes remerciements à l'ensemble des membres du Laboratoire de Mathématiques de Besançon et en particulier à Youssef ESSTAFI.

Je tiens aussi à remercier Docteur Hassan ALI BARKAD et Docteur Hassan ELMI ROBLEH pour leurs soutiens.

Mes remerciements vont à ma chère épouse qui a tout fait pour m'aider, qui m'a soutenu et surtout supporté dans tout ce que j'ai entrepris. Enfin, je remercie mes parents, mes frères et sœurs, l'ensemble de ma famille et mes amis pour leur soutien constant et leur présence sans faille.

---

*Je dédie ce travail à ma fille Zahra, à mon fils Mouhoumed, à mon épouse Ilham et à la personne sans laquelle je n'en serais pas là aujourd'hui qui est à la fois une mère et sœur pour moi.*

# Résumé

Dans cette thèse nous nous intéressons principalement à la validation des modèles ARMA saisonniers et/ou périodiques (SARMA pour Seasonal AutoRegressive Moving-Average, PARMA pour Periodic ARMA et SPARMA Seasonal PARMA) en considérant des termes d'erreur non corrélés mais qui peuvent contenir des dépendances non linéaires. Ces modèles sont appelés des ARMA saisonniers et/ou périodiques faibles et permettent de traiter des processus qui peuvent avoir des dynamiques non linéaires très générales. Par opposition nous appelons modèles ARMA saisonniers et/ou périodiques forts quand le terme d'erreur est indépendant et identiquement distribué (iid). Relâcher l'hypothèse d'indépendance sur le terme d'erreur (une hypothèse habituellement imposée dans la littérature) permet aux modèles SARMA, PARMA et SPARMA faibles d'élargir considérablement leurs champs d'application. Nous étudions les propriétés asymptotiques de l'estimateur du quasi-maximum de vraisemblance et de l'estimateur des moindres carrés généralisés et quasi-généralisés des modèles SARMA faibles et SPARMA faibles. Ensuite nous nous intéressons aux tests fondés sur les résidus, qui ont pour objet de vérifier que les résidus des modèles estimés sont bien des estimations de bruits blancs. Plus particulièrement, nous nous intéressons aux tests portmanteau, aussi appelés tests d'autocorrélation. Nous montrons que les autocorrélations résiduelles de ces modèles ont une distribution asymptotique normal de matrice de covariance différente de celle obtenue dans le cas standard (c'est-à-dire sous les hypothèses iid sur le bruit). Cela nous permet d'avoir le comportement asymptotique des statistiques de tests portmanteau et de proposer ainsi des versions modifiées de ces tests. La distribution asymptotique des tests portmanteau est approximée par un chi-deux lorsque le terme d'erreur est supposé iid. Dans le cas général, nous montrons que cette distribution asymptotique est celle d'une somme pondérée de chi-deux, elle peut être très différente de l'approximation chi-deux usuelle du cas fort (iid). Nous proposons également une approche qui nous permet d'éviter l'estimation de la matrice de covariance asymptotique des autocorrélations empiriques dite d'auto-normalisation.

**Mots-clés :** Test d'ajustement, estimation du quasi-maximum de vraisemblance, tests de Box-Pierce et de Ljung-Box portmanteau, autocorrélation résiduelle, saisonnalité, auto-normalisation, modèles SARMA faibles, modèles SVARMA faibles, modèles PARMA faibles, moindres carrés pondérés, moindres carrés généralisés, moindres carrés quasi-généralisés, modèles SPARMA faibles.

# Abstract

In this thesis, we are mainly interested in the validation of seasonal and/or periodic ARMA models (SARMA for Seasonal AutoRegressive Moving-Average, PARMA for Periodic ARMA and SPARMA for Seasonal PARMA) with uncorrelated but non-independent error terms. These models are called weak seasonal and/or periodic ARMA and allow to deal with processes that may have very general non-linear dynamics. By opposition we call strong seasonal and/or periodic ARMA models when the error term is independent and identically distributed (iid). Relaxing the independence assumption on the error term (a assumption usually imposed in the literature) allows the weak SARMA, PARMA and SPARMA models to considerably broaden their fields of application. We study the asymptotic properties of the quasi-maximum likelihood estimator and the generalized and quasi-generalized least squares estimator of weak SARMA and weak SPARMA models. We are interested in portmanteau tests, also known as autocorrelation tests. We show that the residual autocorrelations of these models have a normal asymptotic covariance matrix distribution different from that obtained in the standard case (i.e. under the iid assumptions on noise). This allows us to have the asymptotic behaviour of portmanteau test statistics and thus to offer modified versions of these tests. The asymptotic distribution of portmanteau tests is approximated by a chi-squared when the error term is assumed to be iid. In the general case, we show that this asymptotic distribution is that of a weighted sum of chi-squared. It can be very different from the usual chi-square approximation of the strong case (iid). We also propose an approach that allows us to avoid the estimation of the asymptotic covariance matrix of empirical autocorrelations called self-normalization.

**Keywords :** Goodness-of-fit test, quasi-maximum likelihood estimation, Box-Pierce and Ljung-Box portmanteau tests, residual autocorrelation, seasonality, self-normalization, weak SARMA models, weak SVARMA models, weak PARMA models, weighted least squares, generalized least squares, quasi-generalized least squares, weak SPARMA models.

# Travaux réalisés

## Articles publiés :

[1] Boubacar Maïnassara, Yacouba and Ilmi Amir, Abdoukarim (2019), Distribution of residual autocorrelations for multiplicative seasonal ARMA models with uncorrelated but nonindependent error terms, *Statistica Neerlandica*, 454–474.

[2] Boubacar Maïnassara, Yacouba and Ilmi Amir, Abdoukarim (2020), Multivariate portmanteau tests for weak multiplicative seasonal VARMA models, *Statistical Papers*, 2529–2560.

## Articles soumis ou en cours de finalisation :

[3] Boubacar Maïnassara, Yacouba and Ilmi Amir, Abdoukarim (2021), Portmanteau tests for Periodic ARMA models with uncorrelated but non-independent error terms, soumis.

[4] Boubacar Maïnassara, Yacouba and Ilmi Amir, Abdoukarim (2021), Estimating and validation SPARMA models with uncorrelated but non-independent error terms. En cours de finalisation

# Table des matières

<b>Remerciements</b>	<b>i</b>
<b>Résumé</b>	<b>iii</b>
<b>Abstract</b>	<b>iv</b>
<b>Travaux réalisés</b>	<b>i</b>
<b>1 Introduction</b>	
	<b>5</b>
1.1 Concepts de base et contexte bibliographique . . . . .	5
1.2 Contribution de la thèse . . . . .	16
1.3 Résultats du chapitre 2 . . . . .	17
1.4 Résultats du chapitre 3 . . . . .	24
1.5 Résultats du chapitre 4 . . . . .	31
1.6 Résultats du chapitre 5 . . . . .	40
Bibliographie . . . . .	52
<b>2 Distribution of residual autocorrelations for multiplicative seasonal ARMA models with uncorrelated but non-independent error terms</b>	
	<b>53</b>
2.1 Introduction . . . . .	53
2.2 Estimating weak SARMA models . . . . .	55
2.3 Diagnostic checking in weak SARMA models . . . . .	57
2.4 Numerical illustrations . . . . .	60
2.5 Conclusion . . . . .	68
2.6 Proofs . . . . .	69
Bibliographie . . . . .	74
<b>3 Multivariate portmanteau tests for weak multiplicative seasonal VARMA models</b>	
	<b>75</b>
3.1 Introduction . . . . .	75
3.2 Estimating weak SVARMA models . . . . .	77
3.3 Diagnostic checking in SVARMA models . . . . .	81
3.4 Numerical illustrations . . . . .	85
3.5 Proofs . . . . .	93



<b>4</b>	<b>Portmanteau tests for Periodic ARMA models with uncorrelated but non-independent error terms</b>	<b>98</b>
4.1	Introduction . . . . .	98
4.2	Preliminaries . . . . .	100
4.3	Diagnostic checking in weak PARMA models . . . . .	103
4.4	Numerical illustration . . . . .	110
4.5	Adequacy of weak PARMA models for real datasets . . . . .	124
4.6	Proofs . . . . .	137
	Bibliographie . . . . .	148
<b>5</b>	<b>Estimation and validation of SPARMA models with uncorrelated but non-independent error terms</b>	<b>149</b>
5.1	Introduction . . . . .	149
5.2	Causality, stationarity and invertibility of SPARMA models . . . . .	151
5.3	Estimating weak SPARMA models . . . . .	153
5.4	Estimating the asymptotic variance matrix . . . . .	155
5.5	Diagnostic checking in weak SPARMA models . . . . .	157
5.6	Numerical illustration . . . . .	163
5.7	Appendix: Proofs . . . . .	168
	Bibliographie . . . . .	193
	<b>Conclusion générale et perspectives</b>	<b>x</b>

# Liste des figures

2.1	Autocorrelation of a realization of size $n = 2,000$ for weak SARMA(0,1)(0,1) <sub>12</sub> model (2.11)-(2.12) with $\alpha_1 = 1.3$ . The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong SARMA assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak SARMA assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Proposition 2.3.1. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 2.3.4. . . . . .	62
2.2	Autocorrelation of the particular SARMA(3,1)(0,1) <sub>12</sub> model residuals for the mean-logarithm-corrected of monthly mean total sunspot number. The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong SARMA assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak SARMA assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Proposition 2.3.1. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 2.3.4. . . . . .	67
3.1	Autocorrelations for Weak SVARMA(1,0)(0,1) <sub>12</sub> model (3.12)-(3.14). Approximately 95% of the sample autocorrelations of a realization of size $n = 2,000$ should lie between the bands shown as solid lines (red color). The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 3.3.3. The bands shown as dotted lines (blue color) correspond to the standard (Strong) SVARMA formula. The Lag printed is the real lag (in standard case) divided by the season $s = 12$ (the real standard lags used in the simulation are the lags printed multiplied by the season $s = 12$ ). . . . . .	88
3.2	The monthly U.S. housing starts (series 1 corresponds to HOUST) and housing sold (series 2 corresponds to HSN1F), in thousands. The observations covered the period from January 01, 1963 to July 31, 2018 which correspond to $n = 667$ observations. . . . . .	89
3.3	Autocorrelations of the SVARMA(1,0)(1,1) <sub>12</sub> residuals for the mean-logarithm-corrected of monthly U.S. housing starts (serie 1 corresponds to HOUST) and housing sold (serie 2 corresponds to HSN1F). The real standard lags used are the lags printed multiplied by the season $s = 12$ ). . . . . .	89
4.1	Autocorrelation of the PARMA <sub>5</sub> (1,1) residuals for the squares of the CAC 40 returns. The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong PARMA <sub>5</sub> (1,1) assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak PARMA <sub>5</sub> (1,1) assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 4.3.1. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 4.3.5.	133

- 4.2 Autocorrelation of the  $\text{PARMA}_5(1, 1)$  residuals for the squares of the DAX returns. The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong  $\text{PARMA}_5(1, 1)$  assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak  $\text{PARMA}_5(1, 1)$  assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 4.3.1. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 4.3.5. . . . . 134
- 4.3 Autocorrelation of the  $\text{PARMA}_5(1, 1)$  residuals for the squares of the SP500 returns. The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong  $\text{PARMA}_5(1, 1)$  assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak  $\text{PARMA}_5(1, 1)$  assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 4.3.1. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 4.3.5. . . . . 135
- 4.4 Autocorrelation of the  $\text{PARMA}_5(1, 1)$  residuals for the squares of the Nikkei returns. The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong  $\text{PARMA}_5(1, 1)$  assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak  $\text{PARMA}_5(1, 1)$  assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 4.3.1. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 4.3.5. . . . . 136

# Liste des tableaux

2.1	Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of SARMA(0,1)(0,1) <sub>s</sub> model (2.11)-(2.12) with $\alpha_1 = 0$ . The nominal asymptotic level of the tests is $\alpha = 5\%$ . The number of replications is $N = 1,000$ .	63
2.2	Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of SARMA(0,1)(0,1) <sub>s</sub> model (2.11)-(2.12) with $\alpha_1 = 0.45$ . The nominal asymptotic level of the tests is $\alpha = 5\%$ . The number of replications is $N = 1,000$ .	64
2.3	Empirical power (in %) of the modified and standard versions of the LB and BP tests in the case of SARMA(0,1)(0,1) <sub>s</sub> model (2.13)-(2.12) with $\alpha_1 = 0$ . The number of replications is $N = 1,000$ .	65
2.4	Empirical power (in %) of the modified and standard versions of the LB and BP tests in the case of ARMA(0,1)(0,1) <sub>s</sub> model (2.13)-(2.12) with $\alpha_1 = 0.45$ . The number of replications is $N = 1000$ .	66
3.1	Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of SVARMA(1,0)(0,1) <sub>s</sub> model. The nominal asymptotic level of the tests is $\alpha = 1\%$ .	90
3.2	Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of SVARMA(1,0)(0,1) <sub>s</sub> model. The nominal asymptotic level of the tests is $\alpha = 5\%$ .	91
3.3	Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of SVARMA(1,0)(0,1) <sub>s</sub> model. The nominal asymptotic level of the tests is $\alpha = 10\%$ .	92
4.1	Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a strong PARMA <sub>2</sub> (1,1) given by (4.33) with the parameter $\alpha_0 = (0,0,0,0)'$ and where the innovation $\epsilon_t \sim \mathcal{N}(0,1)$ .	113
4.2	Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a weak PARMA <sub>2</sub> (1,1) given by (4.33) with the parameter $\alpha_0 = (0,0,0,0)'$ and where the innovation $\epsilon_t \sim (4.34)$ .	114
4.3	Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a weak PARMA <sub>2</sub> (1,1) given by (4.33) with the parameter $\alpha_0 = (0,0,0,0)'$ and where the innovation $\epsilon_t \sim (4.35)$ .	115
4.4	Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a strong PARMA <sub>2</sub> (1,1) given by (4.33) with the parameter $\alpha_0 = (0.8,0.9,0,0)'$ and where the innovation $\epsilon_t \sim \mathcal{N}(0,1)$ .	116
4.5	Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a weak PARMA <sub>2</sub> (1,1) given by (4.33) with the parameter $\alpha_0 = (0.8,0.9,0,0)'$ and where the innovation $\epsilon_t \sim (4.34)$ .	117
4.6	Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a weak PARMA <sub>2</sub> (1,1) given by (4.33) with the parameter $\alpha_0 = (0.8,0.9,0,0)'$ and where the innovation $\epsilon_t \sim (4.35)$ .	118

4.7	Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a strong PARMA <sub>2</sub> (1, 1) given by (4.33) with the parameter $\alpha_0 = (0.8, 0.9, -0.5, -0.45)'$ and where the innovation $\epsilon_t \sim \mathcal{N}(0, 1)$ . . . . .	119
4.8	Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a weak PARMA <sub>2</sub> (1, 1) given by (4.33) with the parameter $\alpha_0 = (0.8, 0.9, -0.5, -0.45)'$ and where the innovation $\epsilon_t \sim (4.34)$ . . . . .	120
4.9	Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a weak PARMA <sub>2</sub> (1, 1) given by (4.33) with the parameter $\alpha_0 = (0.8, 0.9, -0.5, -0.45)'$ and where the innovation $\epsilon_t \sim (4.35)$ . . . . .	121
4.10	Empirical power (in %) of the modified and standard versions of the LBM and BP tests for the null hypothesis of a periodic noise against a PARMA <sub>2</sub> (1, 1) given by (4.33) with the parameter $\alpha_0 = (0.8, 0.9, 0.0, 0.0)'$ . Model I with innovation $\epsilon_t \sim \mathcal{N}(0, 1)$ ; Model II with innovation $\epsilon_t \sim (4.34)$ and Model III with innovation $\epsilon_t \sim (4.35)$ . The nominal asymptotic level of the tests is $\alpha = 5\%$ . . . . .	122
4.11	Empirical power (in %) of the modified and standard versions of the LBM and BP tests for the null hypothesis of a PARMA <sub>2</sub> (1, 0) against a PARMA <sub>2</sub> (1, 1) given by (4.33) with the parameter $\alpha_0 = (0.8, 0.9, 0.5, 0.45)'$ . Model I with innovation $\epsilon_t \sim \mathcal{N}(0, 1)$ ; Model II with innovation $\epsilon_t \sim (4.34)$ and Model III with innovation $\epsilon_t \sim (4.35)$ . The nominal asymptotic level of the tests is $\alpha = 5\%$ . . . . .	123
4.12	Sample correlations between $r_{5n+1}$ (Monday) and $r_{5n+v-h}$ (other days); between $r_{5n+1}^2$ and $r_{5n+v-h}^2$ for the four indices considered. . . . .	127
4.13	Modified and standard versions of portmanteau tests to check the null hypothesis that the returns follow a periodic white noise, based on $m$ squared residuals autocovariances. . . . .	128
4.14	Modified and standard versions of portmanteau tests to check the null hypothesis that the returns follow a periodic white noise, based on $m$ squared residuals autocovariances. . . . .	129
4.15	QLS estimates, their $p$ -values (in parentheses) and their estimated standard errors (in brackets) of a weak PARMA <sub>5</sub> (1, 1) model fitted to the mean corrected series of the squared returns of the CAC 40 and DAX indices. . . . .	130
4.16	QLS estimates, their $p$ -values (in parentheses) and their estimated standard errors (in brackets) of a weak PARMA <sub>5</sub> (1, 1) model fitted to the mean corrected series of the squared returns of the SP500 and Nikkei indices. . . . .	130
4.17	Modified and standard versions of portmanteau tests to check the null hypothesis that the squared returns follow a weak PARMA <sub>5</sub> (1, 1) models, based on $m$ squared residuals autocovariances. . . . .	131
4.18	Modified and standard versions of portmanteau tests to check the null hypothesis that the squared returns follow a weak PARMA <sub>5</sub> (1, 1) models, based on $m$ squared residuals autocovariances. . . . .	132
5.1	Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a strong SPARMA <sub>2</sub> (1, 1)(1, 1) given by (5.37) where the innovation $\epsilon_{nT+v} \sim \mathcal{N}(0, \sigma_v^2)$ . . . . .	166
5.2	Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a weak SPARMA <sub>2</sub> (1, 1)(1, 1) given by (5.37) where the innovation $\epsilon_{nT+v} \sim (5.38)$ . . . . .	167
5.3	Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a weak SPARMA <sub>2</sub> (1, 1)(1, 1) given by (5.37) where the innovation $\epsilon_{nT+v} \sim (5.39)$ . . . . .	168

# Introduction

## 1.1 Concepts de base et contexte bibliographique

L'étude des séries temporelles correspond à l'analyse statistique d'observations régulièrement espacées dans le temps. Nous rencontrons de nombreuses séries temporelles provenant de divers domaines d'études, tel que la météorologie, l'hydrologie et l'économie entre autres.

Notre recherche portera essentiellement sur la validation de modèles saisonniers et/ou périodiques (que nous définirons dans la suite) en imposant une hypothèse moins restrictive (hypothèse qu'on définira plus tard) sur le terme d'erreur contenue dans ces modèles. Avant de présenter les différentes classes de modèles qu'on développera dans notre thèse, commençons par rappeler la classe des modèles autorégressifs moyennes mobiles (ARMA pour AutoRegressive Moving Average), qui constituent la classe la plus utilisée de modèles de séries temporelles.

### 1.1.1 Modèle autorégressif moyenne mobile

En 1927, [86] introduit les modèles autoregressifs (AR pour AutoRegressive) dans la littérature, il propose le modèle

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t \quad (1.1)$$

et trouve qu'un modèle AR d'ordre 2 représente assez bien les comportements des données de Wolfer<sup>1</sup> sur les taches solaires (1749 – 1924) où  $\phi_1, \phi_2$  sont des paramètres inconnus et  $\epsilon_t$  sera défini par la suite. Indépendamment de [86], [76] présente les modèles moyennes mobiles (MA pour Moving Average).

$$Y_t = \mu_0 \epsilon_t + \mu_1 \epsilon_{t-1} + \dots + \mu_q \epsilon_{t-q} \quad (1.2)$$

où les  $(\mu_i)_{0 \leq i \leq q}$  sont des paramètres inconnus. Ces processus deviendront respectivement les AR( $p$ ) et MA( $q$ ) où  $p$  et  $q$  sont les ordres des modèles.

L'apparition de ces deux modèles a ouvert une voie à l'étude des séries temporelles. En 1954 Wold a mis en application les modèles linéaires ARMA (combinaisons des modèles (1.1) et (1.2)).

Un processus  $X = (X_t)_{t \in \mathbb{Z}}$  stationnaire au second ordre<sup>2</sup> à valeurs réelles, admet une représentation ARMA( $p, q$ ) si, pour tout  $t \in \mathbb{Z}$

1. Alfred Wolfer (1854-1931) était un astronome suisse

2. Le processus  $(X_t)$  à valeurs réelles est dit stationnaire au second ordre ou faible si :

$$a(L)X_t = b(L)\epsilon_t, \quad \forall t \in \mathbb{Z}, \quad (1.3)$$

où les polynômes AR et MA sont définis respectivement par  $a(L) = 1 - \sum_{i=1}^p a_{0i}L^i$  et  $b(L) = 1 - \sum_{i=1}^q b_{0i}L^i$  et  $L$  est l'opérateur retard usuel (i.e  $X_{t-1} = LX_t$ ,  $\epsilon_{t-k} = L^k\epsilon_t$  pour tout  $t$ ,  $k \in \mathbb{Z}$ ). On suppose que  $a_{00}^2 + b_{00}^2 \neq 0$  (par convention  $a_{00} = b_{00} = 1$ ). Le terme d'erreur  $\epsilon_t$  est un bruit blanc faible qui vérifie l'hypothèse suivante :

**(H0) :** (i)  $\mathbb{E}(\epsilon_t) = 0$ , (ii)  $\text{Var}(\epsilon_t) = \sigma_0^2$  et (iii)  $\text{Cov}(\epsilon_t, \epsilon_{t-h}) = 0 \quad \forall h, t \in \mathbb{Z}, h \neq 0$ .

**Remarque 1.1.1** *Il est parfois nécessaire de remplacer l'hypothèse (iii) par l'hypothèse plus forte -(iii') les variables  $\epsilon_t$  et  $\epsilon_{t-h}$  sont indépendantes. On parle alors de bruit blanc fort. Il est important de noter qu'aucune hypothèse d'indépendance n'est faite dans la définition du bruit blanc faible. Les variables aux différentes dates sont seulement non corrélées.*

Parmi les processus pouvant être ajustés à des bruits blancs faibles, nous pouvons citer les processus autorégressifs conditionnellement hétéroscédastique (ARCH pour AutoRegressive Conditional Heteroskedasticity) introduit par [31] ainsi que les extensions GARCH (Generalized ARCH) introduit par [12]. Ces processus sont non corrélés mais peuvent contenir néanmoins des dépendances temporelles. On peut aussi citer les modèles all-pass [4] qui génèrent également des séries non corrélées. D'autres types de processus pouvant être ajustés à des bruits blancs faibles ont été introduit par [68].

Ainsi, on parlera d'un modèle ARMA( $p, q$ ) fort, si dans (1.3),  $\epsilon = (\epsilon_t)_{t \in \mathbb{Z}}$  est un bruit fort, à savoir une suite indépendante et identiquement distribuée (i.e iid). Et si on suppose simplement que  $\epsilon$  est un bruit blanc faible mais qu'il n'est pas forcément iid, on dit que  $X$  admet une représentation ARMA faible. La distinction entre modèle ARMA fort ou faible n'est donc qu'une question d'hypothèse sur le bruit. Comme la contrainte sur  $\epsilon$  est moins forte pour un modèle ARMA faible que pour un modèle ARMA fort, nous en déduisons la relation suivante

$$\{\text{ARMA forts}\} \subset \{\text{ARMA faibles}\}.$$

Ces modèles sont généralement utilisés avec des hypothèses fortes sur le bruit qui en limitent la généralité. Ainsi relâcher l'hypothèse d'indépendance permet d'élargir le champs d'applications des modèles ARMA faibles. Par exemple [34] et [35] ont montré qu'il existe des processus très variés admettant, à la fois, des représentations non linéaires<sup>3</sup> et linéaires de type ARMA, pourvu que les hypothèses sur le bruit du modèle ARMA soient suffisamment peu restrictives. Pour l'étude des séries temporelles, la notion de stationnarité faible conduit à imposer des contraintes sur les racines des polynômes AR et MA. Si les racines de  $a(z) = 0$  sont en dehors du disque unité, alors  $X_t$  est une combinaison linéaire de  $\epsilon_t, \epsilon_{t-1}, \dots$  (cette combinaison est une solution non anticipative ou causale de (1.3)). Si les racines de  $b(z) = 0$  sont en dehors du disque

---

(i)  $\mathbb{E}(X_t^2) < \infty \quad \forall t \in \mathbb{Z}$

(ii)  $\mathbb{E}(X_t) = m \quad \forall t \in \mathbb{Z}$

(iii)  $\text{Cov}(X_t, X_{t+h}) = \gamma(h) \quad \forall h, t \in \mathbb{Z}$ .

La fonction  $\gamma(\cdot)$ , est appelée fonction d'autocovariance de  $(X_t)$ .

3. Pour définir les notions de linéarité et de non linéarité d'un modèle, rappelons la décomposition de [85] qui affirme que si  $(X_t)_{t \in \mathbb{Z}}$  est un processus stationnaire, purement non déterministe, il peut être exprimé pour tout  $t \in \mathbb{Z}$  sous la forme :

$$X_t = \sum_{i=0}^{+\infty} \psi_i \epsilon_{t-i},$$

où  $\psi_0 = 1$ ,  $\sum_{i=0}^{+\infty} \psi_i < \infty$  (voir [25]). Le processus  $(\epsilon_t)_{t \in \mathbb{Z}}$  est l'innovation linéaire du processus  $X_t$ . Nous parlons de modèles linéaires lorsque le processus d'innovation  $(\epsilon_t)_{t \in \mathbb{Z}}$  est iid et de modèles non linéaires dans le cas contraire.

unité, alors  $\epsilon_t$  est une combinaison linéaire de  $X_t, X_{t-1}, \dots$  (l'équation (1.3) est dite inversible). Les deux hypothèses précédentes sur les polynômes AR et MA seront regroupées dans l'hypothèse suivante

$$(H1) : a(z)b(z) = 0 \Rightarrow |z| > 1.$$

De nombreux outils statistiques ont été développés pour l'analyse, l'identification, l'estimation et la validation des modèles ARMA forts (voir [44], [69], [39], [41], [79], [24], [37], [57], [61]) pour ne citer que ceux-ci.

Dans le cas des ARMA faibles, plusieurs travaux ont été réalisés. Par exemple, [34] ont montré la consistance et la normalité asymptotique de l'estimateur des moindres carrés ordinaires sous certaines hypothèses d'ergodicité<sup>4</sup> et de mélange<sup>5</sup>. Il se trouve que la précision de cet estimateur dépend de manière cruciale des hypothèses faites sur le bruit. [68] ont étudié les comportements asymptotiques des autocorrélations empiriques. [33] ont étudié des tests de validation des modèles ARMA faible et ont montré que dans ce cas, les tests standards de Box-Pierce et Ljung-Box peuvent mal fonctionner.

Il arrive que dans de nombreux cas, on tombe sur des séries présentant des phénomènes qui se répètent régulièrement sur une période assez courte dont il faudrait en tenir compte. Le nombre d'observation sur cette période est appelée saison et sera notée  $s$  dans la suite. Un mouvement saisonnier annuel ou trimestriel par exemple, correspond respectivement à  $s = 12$  et  $s = 4$ . On obtient la saison en comptant le nombre d'observation séparant les pics et/ou les creux. La capacité de la classe de modèles ARMA à capter cette dynamique est limitée. D'où l'intérêt d'étendre cette classe de modèle.

### 1.1.2 Modèle saisonnier autorégressif moyenne mobile

Pour améliorer les modèles ARMA existant, nous présentons une classe de modèles, qui est conçue pour modéliser des séries chronologiques présentant des phénomènes saisonniers, introduite à l'origine par [23].

Un processus  $X = (X_t)_{t \in \mathbb{Z}}$  stationnaire au second ordre à valeurs réelles, admet une représentation saisonnière autorégressive moyenne mobile multiplicative (SARMA pour Seasonal ARMA),

4. Nous allons rappeler la notion de stationnarité stricte, avant de définir l'ergodicité.

Un processus  $(X_t)_{t \in \mathbb{Z}}$  est dit strictement stationnaire si  $(X_1, \dots, X_t)$  et  $(X_{1+h}, \dots, X_{t+h})$  ont la même loi jointe pour tout  $h \in \mathbb{Z}$  et tout  $t \in \mathbb{Z}$ .

On dit qu'une suite stationnaire est ergodique si elle satisfait la loi forte des grands nombres.

Un processus strictement stationnaire  $Z = (Z_t)_{t \in \mathbb{Z}}$ , à valeurs réelles, est dit ergodique si et seulement si, pour tout borélien  $B$  de  $\mathbb{R}^{h-1}$  et tout entier  $h$

$$\frac{1}{n} \sum_{t=1}^n I_B(Z_t, Z_{t+1}, \dots, Z_{t+h}) \rightarrow \mathbb{P}[(Z_1, \dots, Z_{t+h}) \in B]$$

avec probabilité 1.

5. Dans un espace probabilisé  $(\Omega, \mathcal{A}_0, \mathbb{P})$ , considérons  $\mathcal{A}$  et  $\mathcal{B}$  deux sous-tribus de  $\mathcal{A}_0$ . Nous définissons le coefficient de  $\alpha$ -mélange (ou de mélange fort) entre deux tribus par

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

Il est évident que si  $\mathcal{A}$  et  $\mathcal{B}$  sont indépendantes alors  $\alpha(\mathcal{A}, \mathcal{B}) = 0$ . Si par contre  $\mathcal{A} = \mathcal{B} = \{A, A^c, \emptyset, \Omega\}$  avec  $\mathbb{P} = \frac{1}{2}$  alors  $\alpha(\mathcal{A}, \mathcal{B}) = \frac{1}{4}$ . Les coefficients de mélange fort d'un processus  $X = (X_t)$  sont définis par

$$\alpha_X(h) = \sup_t \alpha[\sigma(X_u \leq t), \sigma(X_u \leq t+h)].$$

Si  $X$  est stationnaire, on peut enlever le terme  $\sup_t$ . On dit que  $X$  est fortement mélangé si  $\alpha_X(h) \rightarrow 0$  quand  $h \rightarrow \infty$ .



notée  $SARMA(p, q)(P, Q)_s$  si, pour tout  $t \in \mathbb{Z}$

$$a(L)\mathbf{a}(L)X_t = b(L)\mathbf{b}(L)\epsilon_t, \quad \forall t \in \mathbb{Z}, \quad (1.4)$$

où  $a(L)$  et  $b(L)$  sont respectivement les polynômes non saisonniers AR et MA, tandis que les polynômes saisonniers AR et MA sont donnés respectivement par  $\mathbf{a}(L) = 1 - \sum_{i=1}^P \mathbf{a}_{0i}L^{si}$  et  $\mathbf{b}(L) = 1 - \sum_{i=1}^Q \mathbf{b}_{0i}L^{si}$ . On suppose que  $a_{0p}^2 + b_{0q}^2 + \mathbf{a}_{0p}^2 + \mathbf{b}_{0q}^2 \neq 0$  (par convention  $a_{00} = b_{00} = \mathbf{a}_{00} = \mathbf{b}_{00} = 1$ ). Le terme d'erreur  $\epsilon_t$  est un bruit blanc faible ou fort.

Comme pour les modèles ARMA, nous parlons de modèle SARMA fort quand le bruit blanc dans (1.4) est fort et de modèle SARMA faible dans le cas contraire. Nous en déduisons également la relation suivante

$$\{\text{SARMA forts}\} \subset \{\text{SARMA faibles}\}.$$

Ces modèles sont généralement utilisés avec des hypothèses fortes sur le bruit qui en limitent la généralité. Ainsi, relâcher l'hypothèse d'indépendance permet d'élargir le champ d'applications des modèles SARMA faibles. Contrairement aux modèles ARMA, le modèle SARMA est dit causal, si les racines de  $a(z) = 0$  et  $\mathbf{a}(z) = 0$  sont en dehors du disque unité et il est dit inversible, si les racines de  $b(z) = 0$  et  $\mathbf{b}(z) = 0$  sont en dehors du disque unité. Nous pouvons énoncer ces contraintes sous l'hypothèse suivante

$$(\mathbf{H}'1) : a(z)\mathbf{a}(z)b(z)\mathbf{b}(z) = 0 \Rightarrow |z| > 1.$$

Cette classe des modèles tient compte de ces caractéristiques répétitives qui indiquent la présence d'une saisonnalité et constitue une extension de modèles ARMA. Le modèle (1.4) peut être vu comme un modèle ARMA d'ordre  $(p + Ps, q + Qs)$  dans lequel certains coefficients sont nuls et les autres sont fonctions du vecteurs des paramètres des polynômes saisonniers et non saisonniers.

Nous illustrons sur l'exemple suivant la structure de la fonction d'autocorrélations des modèles ARMA et SARMA.

**Exemple 1.1.1** Soit un processus  $X_t$  satisfaisant un modèle ARMA  $(0, 1)$

$$X_t = \epsilon_t - b\epsilon_{t-1}, \quad \text{avec } \epsilon_t \text{ un bb faible}$$

$$\begin{aligned} \text{Cov}(X_t, X_{t-h}) &= \text{Cov}(\epsilon_t - b\epsilon_{t-1}, \epsilon_{t-h} - b\epsilon_{t-1-h}) \\ &= \text{Cov}(\epsilon_t, \epsilon_{t-h}) - \text{Cov}(\epsilon_t, b\epsilon_{t-1-h}) - \text{Cov}(b\epsilon_{t-1}, \epsilon_{t-h}) + \text{Cov}(b\epsilon_{t-1}, b\epsilon_{t-1-h}) \end{aligned}$$

La fonction d'autocovariance du processus s'écrit

$$\gamma(h) = \begin{cases} \sigma^2(1 + b^2) & \text{si } h = 0 \\ -b\sigma^2 & \text{si } h = 1 \\ 0 & \forall h > 1 \end{cases}$$

La fonction d'autocorrélation du processus s'écrit

$$\rho(h) = \begin{cases} \frac{-b}{(1+b^2)} & \text{si } h = 1 \\ 0 & \forall h > 1 \end{cases}$$

Soit un processus  $X_t$  satisfaisant un modèle SARMA  $(0, 1)(0, 1)_s$

$$X_t = \epsilon_t - a\epsilon_{t-s} - b\epsilon_{t-1} + ba\epsilon_{t-1-s}, \quad \text{avec } \epsilon_t \text{ un bb faible}$$

$$\begin{aligned}
 \text{Cov}(X_t, X_{t-h}) &= \text{Cov}(\epsilon_t - a\epsilon_{t-s} - b\epsilon_{t-1} + ba\epsilon_{t-1-s}, \epsilon_{t-h} - a\epsilon_{t-s-h} - b\epsilon_{t-1-h} + ba\epsilon_{t-1-s-h}) \\
 &= \text{Cov}(\epsilon_t, \epsilon_{t-h}) - \text{Cov}(\epsilon_t, a\epsilon_{t-s-h}) - \text{Cov}(\epsilon_t, b\epsilon_{t-1-h}) + \text{Cov}(\epsilon_t, ba\epsilon_{t-1-s-h}) \\
 &\quad - \text{Cov}(a\epsilon_{t-s}, \epsilon_{t-h}) + \text{Cov}(a\epsilon_{t-s}, a\epsilon_{t-s-h}) + \text{Cov}(a\epsilon_{t-s}, b\epsilon_{t-1-h}) - \text{Cov}(a\epsilon_{t-s}, ba\epsilon_{t-1-s-h}) \\
 &\quad - \text{Cov}(b\epsilon_{t-1}, \epsilon_{t-h}) + \text{Cov}(b\epsilon_{t-1}, a\epsilon_{t-s-h}) + \text{Cov}(b\epsilon_{t-1}, b\epsilon_{t-1-h}) - \text{Cov}(b\epsilon_{t-1}, ba\epsilon_{t-1-s-h}) \\
 &\quad + \text{Cov}(ba\epsilon_{t-1-s}, \epsilon_{t-h}) - \text{Cov}(ba\epsilon_{t-1-s}, a\epsilon_{t-s-h}) - \text{Cov}(ba\epsilon_{t-1-s}, b\epsilon_{t-1-h}) \\
 &\quad + \text{Cov}(ba\epsilon_{t-1-s}, ba\epsilon_{t-1-s-h})
 \end{aligned}$$

La fonction d'autocovariance du processus s'écrit

$$\gamma(h) = \begin{cases} \sigma_\epsilon^2(1 + a^2 + b^2 + a^2b^2) & \text{si } h = 0 \text{ et } s \neq 1 \\ \sigma_\epsilon^2(1 + 2ab + a^2 + b^2 + a^2b^2) & \text{si } h = 0 \text{ et } s = 1 \text{ ou } s = -1 \\ -\sigma_\epsilon^2(a + b + a^2b + ab^2) & \text{si } h = 1 \text{ et } s = 1 \text{ ou } s = -1 \\ -\sigma_\epsilon^2b(1 + a^2) & \text{si } h = 1 \text{ et } s \neq 1 \\ 0 & \text{si } h > 1 \text{ et } s \neq |h| \\ -\sigma_\epsilon^2a(1 + b^2) & \text{si } h > 1 \text{ et } s = |h| \end{cases}$$

La fonction d'autocorrélation du processus s'écrit

$$\rho(h) = \begin{cases} -\frac{b(1+a^2)}{(1+a^2+b^2+a^2b^2)} & \text{si } h = 1 \text{ et } s \neq 1 \\ -\frac{(a+b+a^2b+ab^2)}{(1+2ab+a^2+b^2+a^2b^2)} & \text{si } h = 1 \text{ et } s = 1 \text{ ou } s = -1 \\ 0 & \text{si } h > 1 \text{ et } s \neq |h| \\ -\frac{a(1+b^2)}{(1+a^2+b^2+a^2b^2)} & \text{si } h > 1 \text{ et } s = |h| \end{cases}$$

Nous observons que les fonctions d'autocorrélations des modèles ARMA dépendent seulement du retard  $h$  tandis que pour les modèles SARMA, en plus du retard  $h$  la saisonnalité intervient dans le calcul.

Le problème qui nous préoccupera premièrement dans cette thèse est la validation des modèles SARMA faibles.

Les travaux consacrés à la validation des modèles SARMA faibles n'ont pas encore reçu autant d'attention et se limitent essentiellement dans le cadre fort. [29] et [57], ont examiné les tests de corrélations sur le SARMA. [29] a proposée une statistique de test basée sur un estimateur de la densité spectrale. Nous pouvons citer par exemple [81] qui ont étudié les propriétés asymptotiques des estimateurs du maximum de vraisemblance conditionnelle des modèles vectoriels saisonniers ARMA (SVARMA pour Seasonal Vector ARMA) et ont proposé un test portmanteau pour ces modèles. [56] a conçu des statistiques de test portmanteau dans lesquelles les fonctions d'autocorrélations sont multiples de la saisonnalité ( $1s, 2s, 3s, \dots$ ).

Dans de nombreuses situations, ces tests sont mis en œuvre pour vérifier la validité des modèles SARMA forts. Cette méthodologie standard doit être adapté pour tenir compte de possible dépendance des termes d'erreurs. D'où l'intérêt de notre thèse à s'orienter sur cet axe de travail.

De nombreux travaux ont montré (voir par exemple [34] et [35]) que les modèles ARMA forts sont mal adaptés pour l'analyse de certaines séries économiques, hydrologiques, environnementales etc. Nous pourrions supposer qu'une conclusion similaire reste aussi vraie pour les modèles SARMA forts.

Dans le cadre des SARMA faibles, les travaux notables sont : [19] qui ont proposé des statistiques de test portmanteau modifiées afin de tester l'adéquation des modèles SARMA faibles et [20] qui ont également étendu le travail de [56] aux modèles SVARMA faibles.

On peut être amené à travailler ou à modéliser des séries chronologiques dont la structure de corrélation dépend de la saison. Par conséquent, les modèles SARMA ne seront pas en mesure de bien les modéliser. En effet, les méthodes d'ajustement saisonnier traitent les observations dans tous les saisons de la même manière.

A titre illustratif, considérons l'exemple donné dans [78], sur les relevés journaliers de la concentration d'ozone à Los Angeles.

**Exemple 1.1.1** *La concentration d'ozone à Los Angeles s'accumule généralement pendant les heures du matin, elle atteint un pic en début d'après-midi et elle se dissipe, presque à zéro, la nuit.*

Notons  $X = (X_1, X_2, \dots, X_r)$  la série d'observation de cette concentration d'ozone. Ainsi nous distinguons 3 périodes de relevés journaliers différents. Dans la suite nous noterons  $v$  une saison ou une période quelconques (dans cet exemple  $v \in \{1, 2, 3\}$ ). Ce qui nous donne :

- $v = 1$  va correspondre aux relevés "des heures du matin",
- $v = 2$  va correspondre aux relevés "les heures d'après midi",
- $v = 3$  va correspondre aux relevés "les heures de nuit".

Notons  $T$  le nombre de périodes considérées (ici  $T = 3$ ) et notons également  $n \in \{0, \dots, N-1\}$  un jour quelconque qui est en général appelé cycle.  $N$  représente le nombre des jours (dans l'exemple), si la série d'observation est trimestrielle  $N$  représentera le nombre total des trimestres contenus dans la série et si la série est annuelle  $N$ , sera le nombre d'année. Pour un  $v$  fixé et une valeur  $T$  donnée, nous désignons  $(X_{nT+v})$  la réalisation de la  $v$ -ème saison avec  $v \in \{1, \dots, T\}$  du  $n$ -ème cycle.

On pourrait donc s'attendre à ce que la corrélation entre les relevés d'ozone du matin ( $X_{nT+1}$ ) et du début de l'après-midi ( $X_{nT+2}$ ) d'une journée donnée soit considérablement plus forte que la corrélation entre le relevé du matin ( $X_{nT+1}$ ) et celui de minuit ( $X_{nT+3}$ ). Nous remarquons que  $r = NT$  désigne le nombre total d'observations.

Par conséquent les modèles SARMA ne sont pas adaptés à ces types des données.

D'autres exemples peuvent être trouvés dans [43] qui illustrent ces types de données.

**Exemple 1.1.2** *Considérons le cas du ruissellement provenant de la fonte de neige à la fin de l'hiver ou de début de printemps dans une région nordique. Si la fonte des neiges est un facteur important du ruissellement, qui peut se produire soit en mars soit en avril, la corrélation entre les débits observés pour ces mois (c'est-à-dire entre  $X_{nT+3}$  et  $X_{nT+4}$ , ici  $T = 12$  et le cycle  $n$  représente une année quelconque) peut être négative alors qu'à d'autres moment de l'année, elle est généralement positive.*

Si on ajuste un modèle SARMA à ces types des séries chronologiques mentionnées plus haut, le modèle SARMA peut ne pas être en mesure de décrire la corrélation périodique ou saisonnière contenue dans la série.

### 1.1.3 Modèle périodique autorégressif moyenne mobile

Les modèles périodiques sont utilisés depuis longtemps dans les disciplines de l'environnement, de l'hydrologie et de la météorologie voir par exemple [50], [66], [11], [84]. Et dans les années 1980, des auteurs comme [67] et [65], entre autres les introduisent dans l'économie. Nous pouvons aussi nous référer à [36] pour plus de détails sur les motivations de choix d'un modèle périodique dans la modélisation des séries chronologiques.

Un processus  $X = (X_t)_{t \in \mathbb{Z}}$  à valeurs réelles, admet une représentation périodique autorégressif moyenne mobile (PARMA pour Periodic ARMA), s'il satisfait l'équation suivante

$$X_t - \sum_{i=1}^p \phi_i(t)X_{t-i} = \epsilon_t - \sum_{j=1}^q \theta_j(t)\epsilon_{t-j}, \quad (1.5)$$

où les coefficients vérifient  $\phi_i(t+T) = \phi_i(t)$  pour  $1 \leq i \leq p$  et  $\theta_j(t+T) = \theta_j(t)$  pour  $1 \leq j \leq q$ .  $\{\epsilon_t\}$  est un bruit blanc périodique de moyenne nulle avec  $\epsilon_t$  et  $\epsilon_s$  non corrélés quand  $t \neq s$  et  $\text{Var}(\epsilon_t) = \sigma^2(t) > 0$  avec  $\sigma^2(t+T) = \sigma^2(t)$  (voir par exemple [45] pour plus de détails). Dans la suite nous utiliserons les notations introduites dans l'exemple 1.1.1. Le modèle (1.5) peut s'écrire de manière équivalente

$$X_{nT+v} - \sum_{k=1}^p \phi_k(v)X_{nT+v-k} = \epsilon_{nT+v} - \sum_{l=1}^q \theta_l(v)\epsilon_{nT+v-l}. \quad (1.6)$$

où  $\phi_k(v)_{k=1, \dots, p(v)}$  et  $\theta_l(v)_{l=1, \dots, q(v)}$  sont les coefficients respectifs AR et MA du modèle pour la saison  $v \in \{1, \dots, T\}$ .

De manière générale, on peut permettre aussi aux ordres  $p$  et  $q$  de varier périodiquement si nécessaire à savoir  $p(v)$  et  $q(v)$  (voir par exemple [55] et [32]). Ainsi l'équation (1.6) devient :

$$X_{nT+v} - \sum_{k=1}^{p(v)} \phi_k(v)X_{nT+v-k} = \epsilon_{nT+v} - \sum_{l=1}^{q(v)} \theta_l(v)\epsilon_{nT+v-l} \quad (1.7)$$

où les ordres des polynômes AR et MA sont respectivement donnés par  $p(v)$  et  $q(v)$ . Si le processus  $(X_{nT+v})$  suit le modèle (1.7), nous le notons  $\text{PARMA}_T(p_1, \dots, p_v, \dots, p_T; q_1, \dots, q_v, \dots, q_T)$ . Il n'y a pas de perte de généralité à prendre  $p_1 = \dots = p_T = p$  et  $q_1 = \dots = q_T = q$  comme dans (1.6) qui sera noté comme un  $\text{PARMA}_T(p, q)$ .

Les modèles PARMA sont une extension des modèles ARMA dans le sens où ils permettent aux paramètres des modèles ARMA de varier avec le temps (voir l'équation (1.5)).

**Remarque 1.1.2** Si  $T = 1$ , le modèle PARMA se réduit au modèle ARMA.

Nous utiliserons de manière interchangeable la notation non périodique  $X_t$  avec la notation périodique  $X_{nT+v}$ , qui nous permet de faire ressortir la saison  $v \in \{1, \dots, T\}$  dans le cycle  $n$ . Le terme d'erreur  $(\epsilon_t) = (\epsilon_{nT+v})$  est un bruit blanc périodique faible qui vérifie l'hypothèse suivante :

$$(H0') : \quad (i) \mathbb{E}(\epsilon_{nT+v}) = 0, \quad (ii) \text{Var}(\epsilon_{nT+v}) = \sigma_v^2 \text{ et } (iii) \text{Cov}(\epsilon_{nT+v}, \epsilon_{nT+v-h}) = 0 \quad \forall h, t \in \mathbb{Z}, h \neq 0.$$

**Remarque 1.1.3** Il est parfois nécessaire de remplacer l'hypothèse (iii) par l'hypothèse plus forte -(iii') les variables  $\epsilon_{nT+v}$  et  $\epsilon_{nT+v-h}$  sont indépendantes. On parle alors de bruit blanc périodique fort. Il est important de noter qu'aucune hypothèse d'indépendance n'est faite dans la définition du bruit blanc périodique faible. Les variables aux différentes dates sont seulement non corrélées.

Un exemple de bruit blanc périodique faible est le processus PGARCH pour Periodic GARCH (voir par exemple [13], [38] et [2]). De nombreux d'autres exemples peuvent être trouvés également dans [32].

Nous parlons de modèle PARMA fort quand le bruit blanc périodique dans (1.7) est fort et de modèle PARMA faible dans le cas contraire. Comme pour les autres modèles présentés dans les parties précédentes, la distinction entre un PARMA fort ou faible n'est donc qu'une question d'hypothèse sur le bruit également. Nous en déduisons la relation suivante

$$\{\text{PARMA forts}\} \subset \{\text{PARMA faibles}\}.$$

Comme pour les autres modèles présentés précédemment, les modèles PARMA sont généralement utilisés avec des hypothèses fortes sur le bruit qui en limitent leur généralité. Ainsi, relâcher l'hypothèse d'indépendance permet d'élargir également le champs d'applications des

modèles PARMA (voir par exemple [32]).

Nous présentons sur cette exemple la structure de la fonction d'autocorrélation du modèle PARMA.

**Exemple 1.1.2** Soit un processus  $(X_{nT+v})_{n \in \mathbb{Z}}$  satisfaisant un modèle  $PARMA_2(0, 1)$

$$X_{2n+v} = \epsilon_{2n+v} - b(v)\epsilon_{2n+v-1}, \text{ avec } \epsilon_{2n+v} \text{ un bb périodique faible}$$

$$\begin{aligned} Cov(X_t, X_{t-h}) &= Cov(\epsilon_{2n+v} - b(v)\epsilon_{2n+v-1}, \epsilon_{2n+v-h} - b(v)\epsilon_{2n+v-1-h}) \\ &= Cov(\epsilon_{2n+v}, \epsilon_{2n+v-h}) - Cov(\epsilon_{2n+v}, b(v)\epsilon_{2n+v-1-h}) \\ &\quad - Cov(b(v)\epsilon_{2n+v-1}, \epsilon_{2n+v-h}) + Cov(b(v)\epsilon_{2n+v-1}, b(v)\epsilon_{2n+v-1-h}) \end{aligned}$$

La fonction d'autocovariance du processus s'écrit

$$\gamma_v(h) = \begin{cases} \sigma_v^2(1 + b^2(v)) & \text{si } h = 0, \text{ pour } v = \{1, 2\} \\ -b(v)\sigma_v^2 & \text{si } h = 1, \text{ pour } v = \{1, 2\} \\ 0 & \forall h > 1 \end{cases}$$

La fonction d'autocorrélation du processus s'écrit

$$\rho_v(h) = \begin{cases} -\frac{\sigma_v b(v)}{\sigma_{v-1}(1+b^2(v))} & \text{si } h = 1, \text{ pour } v = \{1, 2\} \\ 0 & \forall h > 1 \end{cases}$$

Les modèles de séries chronologiques saisonnières et périodiques sont très différents, car la nature de la structure de corrélation diffère sur plusieurs aspect importants : purement stochastique dans les modèles saisonniers (voir l'exemple 1.1.1) et purement déterministe dans les modèles périodiques (voir [80] pour plus de détails).

Contrairement aux modèles précédents, l'étude des conditions d'inversibilité et de stationnarité des modèles périodiques, nécessite une réécriture des modèles sous forme vectoriels. Pour cela, nous posons  $\mathbf{X}_n = (X_{nT+1}, \dots, X_{nT+T})$  et  $\mathbf{\epsilon}_n = (\epsilon_{nT+1}, \dots, \epsilon_{nT+T})$ . Définissons également  $p^* = [(p-1)/T] + 1$  et  $q^* = [(q-1)/T] + 1$  qui seront les nouveaux ordres du modèle vectoriel.

Ainsi, le modèle (1.6) peut s'écrire sous forme d'un modèle VARMA( $p^*, q^*$ ) (voir par exemple [32] pour plus de détails) suivant :

$$\Phi_0 \mathbf{X}_n - \sum_{k=1}^{p^*} \Phi_k \mathbf{X}_{n-k} = \Theta_0 \mathbf{\epsilon}_n - \sum_{l=1}^{q^*} \Theta_l \mathbf{\epsilon}_{n-l} \quad (1.8)$$

où les matrices des coefficients de  $\Phi_k, k = 0, \dots, p^*$  et  $\Theta_l, l = 0, \dots, q^*$  sont définies par

$$(\Phi_0)_{i,j} = \begin{cases} 1 & \text{Si } i = j \\ 0 & \text{Si } i < j \\ -\phi_{i-j}(i) & \text{Si } i > j \end{cases} \quad (\Theta_0)_{i,j} = \begin{cases} 1 & \text{Si } i = j \\ 0 & \text{Si } i < j \\ -\theta_{i-j}(i) & \text{Si } i > j \end{cases} \quad (1.9)$$

avec

$$(\Phi_k)_{i,j} = \phi_{kT+i-j}(i, j), \quad k = 1, \dots, p^* \text{ et } (\Theta_l)_{i,j} = \theta_{lT+i-j}(i) \quad l = 1, \dots, q^*. \quad (1.10)$$

Nous supposons que  $\phi_h(v) = 0, h \notin \{1, \dots, p^*\}$  et  $\theta_h(v) = 0, h \notin \{1, \dots, q^*\}$ . La matrice de variance covariance du bruit blanc  $\mathbf{\epsilon}_n$  de dimension  $T$  sera notée  $\Sigma_\epsilon = (\sigma_1^2, \dots, \sigma_T^2) > 0$ . Définissons

les polynômes matriciels suivants  $\Phi(z) = \Phi_0 - \sum_{i=1}^{p^*} \Phi_i z^i$  et  $\Theta(z) = \Theta_0 - \sum_{i=1}^{q^*} \Theta_i z^i$ . Ainsi, nous réécrivons l'équation (1.8) sous sa forme compacte suivante :

$$\Phi(L)\mathbf{X}_n = \Theta(L)\boldsymbol{\epsilon}_n. \quad (1.11)$$

Il est important de noter que l'opérateur retard  $L$  opère sur l'indice du cycle  $n$ . Lorsqu'il agit sur l'indice de temps  $t = nT + v$  du processus périodique  $(Z_t)$ , nous avons  $L^k Z_{nT+v} = Z_{(n-1)T+v}$ . A partir du modèle (1.8), nous pouvons déduire en principe les propriétés de l'estimation, de l'identification et de la validation des modèles PARMA faibles à partir des résultats existants sur l'estimation, l'identification et la validation des modèles VARMA faibles (voir par exemples [30], [18], [16], [21], [15], [22]). Toutefois, cette approche n'est pas recommandée pour diverses raisons (voir par exemple [32] et [43] pour plus de détails). Néanmoins, nous allons citer certaines de ces raisons ici. Premièrement, les modèles PARMA saisissent bien les caractéristiques physiques des phénomènes naturels décrites dans les séries chronologiques. Ce qui n'est pas le cas des modèles VARMA qui sont une extension naturelle des modèles ARMA au cas multidimensionnel. Il y a moins de paramètres (ce qui inclu une interprétation plus simple) et leur estimation est plus facile. Par exemple, dans un modèle PARMA<sub>T</sub>(1, 1), il y a  $2T$  paramètres AR et MA à estimer. La représentation VARMA(1, 1) correspondante contient  $4T^2$  paramètres parmi lesquels  $2T(2T - 1)$  sont contraints à 1 ou 0 selon (1.9) et (1.10). La différence dans le nombre de paramètres est énorme, en particulier avec les données mensuelles ( $T = 12$ ) ou hebdomadaires ( $T = 52$ ). Deuxièmement, la représentation VARMA (1.8) n'est pas standard, les matrices  $\Phi_0$  et  $\Theta_0$  ne sont pas en générale les matrices identités. La modification du bruit vectoriel dans (1.8) via  $\boldsymbol{\epsilon}_n^* = \Theta_0 \boldsymbol{\epsilon}_n$  puis la multiplication des deux côtés par  $\Theta_0^{-1}$  conduit à une représentation VARMA standard. Et pour les matrices  $\Phi_0$  et  $\Theta_0$ , il faudrait vérifier certaines conditions d'inversibilité, d'identifiabilité contraignantes (voir par exemple [14]).

En utilisant les propriétés du modèle VARMA (voir par exemple [25]), il en résulte que le modèle PARMA est causal et inversible si,

$$(\mathbf{H1}''') : \det \Phi(z) \det \Theta(z) = 0 \Rightarrow |z| > 1.$$

Les modèles des séries chronologiques périodiques sont conçus pour traiter les données avec des structures périodiques. Ces modèles saisissent bien la corrélation variante selon les saisons qui est contenue dans la série. [78] montrent notamment comment des modèles peuvent être mal spécifiés avec des modèles SARMA pour des séries présentant des propriétés périodiques. Voir la partie VI de [43] pour une vaste étude des modèles périodiques et leurs spécificités pratiques dans l'ajustement des séries chronologiques.

Le problème qui nous occupera deuxièmement dans cette thèse est la validation des modèles PARMA faibles.

Nous trouvons dans la littérature de nombreuses recherches sur l'identification, l'estimation et la validation des modèles PARMA forts. [64] proposent l'utilisation de fonction d'autocorrélation (ACF) et de fonction d'autocorrélation partielle (PACF) périodique pour identifier les modèles PARMA en s'appuyant sur les résultats d'une vaste étude de cas. Voir [77], [70] et [62] pour d'autres méthodes d'identification. [83] a développé une technique pour obtenir des estimations de maximum de vraisemblance des paramètres d'un modèle PARMA. [71] ont étudié l'estimation des moments des modèles PARMA d'ordre peu élevé. [6] ont étudié les propriétés asymptotiques des estimations de paramètres du modèle PARMA. [55] ont développé un algorithme pour l'estimation de maximum de vraisemblance des modèles PARMA. Récemment, [82] ont développé une méthode de modélisation robuste pour identifier et estimer le modèle PAR fort afin de bien saisir les variations périodiques du débit des rivières. [8], [9] ont proposé une procédure de regroupement sur les paramètres PAR fort afin d'obtenir un modèle plus parcimonieux et concis. [43] proposent des tests portmanteau pour le modèle PAR fort et soulignent



que ces tests peuvent être adaptés avec les modifications appropriées aux modèles PARMA. [58] sur la base des études de cas, met au point un test de validation PARMA.

Dans le cadre de modèles PARMA faibles très peu des travaux sont disponibles. L'exception notable reste [32] qui ont étudié les propriétés asymptotiques de l'estimation de moindres carrés ordinaires, généralisés et quasi-généralisés pour les modèles PARMA faibles. [2] ont étudié des résultats sur l'estimation du quasi-maximum de vraisemblance des processus PGARCH ainsi que pour de modèles PARMA avec un bruit PGARCH. À notre connaissance, il n'y a pas de test de validation pour les modèles PARMA faibles d'ou l'intérêt de compléter ces travaux comme mentionné précédemment.

Dans la pratique, on peut aussi être amené à modéliser des données plus complexes. Par exemple, considérons  $X$  comme la consommation électrique mensuelle. Notons  $X_{nT+v}$  l'observation du  $v$ -ème mois de la  $n$ -ème année. On pourrait se pencher sur la corrélation entre  $X_{nT+v}$  et  $X_{nT+v+1}$  (mois successif), de même que la corrélation entre  $X_{nT+v}$  et  $X_{(n+1)T+v}$  (même mois pour les années successives). Ces données peuvent présenter des dépendances d'observations consécutives au cours d'une période et des dépendances d'observations entre périodes. Pour pouvoir modéliser ces types des séries chronologiques, nous présentons une classe de modèle qui permet aux paramètres de modèles SARMA de varier périodiquement comme pour les modèles ARMA avec l'extension PARMA.

#### 1.1.4 Modèles saisonniers périodiques autorégressifs moyennes mobiles

Un processus  $(X_{nT+v})$  à valeurs réelles, admet une représentation saisonnière périodique autorégressif moyenne mobile (SPARMA pour Seasonal PARMA), s'il satisfait l'équation suivante

$$a(L, v)\mathbf{a}(L, v)X_{nT+v} = b(L, v)\mathbf{b}(L, v)\epsilon_{nT+v}. \quad (1.12)$$

Nous réécrivons l'équation (1.12) sous la forme développée suivante :

$$\begin{aligned} X_{nT+v} - \sum_{i=1}^{p(v)} a_i(v)X_{nT+v-i} - \sum_{j=1}^{P(v)} \mathbf{a}_j(v)X_{(n-j)T+v} + \sum_{i=1}^{p(v)} \sum_{j=1}^{P(v)} \mathbf{a}_j(v)a_i(v)X_{(n-j)T+v-i} = - \sum_{i=1}^{q(v)} b_i(v)\epsilon_{nT+v-i} \\ - \sum_{j=1}^{Q(v)} \mathbf{b}_j(v)\epsilon_{(n-j)T+v} + \sum_{i=1}^{q(v)} \sum_{j=1}^{Q(v)} \mathbf{b}_j(v)b_i\epsilon_{(n-j)T+v-i} + \epsilon_{nT+v} \end{aligned} \quad (1.13)$$

Le processus  $(X_{nT+v})$ , sera noté  $\text{SPARMA}_T(p_1, \dots, p_T; q_1, \dots, q_T)(P_1, \dots, P_T; Q_1, \dots, Q_T)$

Les ordres des polynômes saisonniers AR et MA sont donnés respectivement par  $P(v)$  et  $Q(v)$ . Les coefficients saisonniers AR et MA du modèle pour la saison  $v \in \{1, \dots, T\}$  sont respectivement  $\mathbf{a}_j(v)_{j=1, \dots, P(v)}$  et  $\mathbf{b}_j(v)_{j=1, \dots, Q(v)}$ . Le terme d'erreur  $(\epsilon_t) = (\epsilon_{nT+v})$  est un bruit blanc périodique déjà défini précédemment.

Nous parlons de modèle SPARMA fort quand le bruit blanc périodique dans (1.13) est fort et de modèle SPARMA faible dans le cas contraire. La distinction entre un PARMA fort ou faible n'est donc qu'une question d'hypothèse sur le bruit. Nous en déduisons la relation suivante

$$\{\text{SPARMA forts}\} \subset \{\text{SPARMA faibles}\}.$$

Comme pour les autres modèles présentés précédemment, nous relâcherons l'hypothèse d'indépendance afin d'élargir également le champs d'applications des modèles SPARMA.

Il n'y a pas de perte de généralité à prendre  $p_1 = \dots = p_T = p$ ,  $q_1 = \dots = q_T = q$ ,  $P_1 = \dots = P_T = P$  et  $Q_1 = \dots = Q_T = Q$  dans (1.13). Ainsi, le processus  $(X_{nT+v})$  noté SPARMA $_T(p,q)(P,Q)$  vérifie l'équation suivante :

$$\begin{aligned} X_{nT+v} - \sum_{i=1}^p a_i(v)X_{nT+v-i} - \sum_{j=1}^P \mathbf{a}_j(v)X_{(n-j)T+v} + \sum_{i=1}^p \sum_{j=1}^P \mathbf{a}_j(v)a_i(v)X_{(n-j)T+v-i} = - \sum_{i=1}^q b_i(v)\epsilon_{nT+v-i} \\ - \sum_{j=1}^Q \mathbf{b}_j(v)\epsilon_{(n-j)T+v} + \sum_{i=1}^q \sum_{j=1}^Q \mathbf{b}_j(v)b_i\epsilon_{(n-j)T+v-i} + \epsilon_{nT+v} \end{aligned} \quad (1.14)$$

Nous illustrons sur cet exemple la structure de la fonction d'autocorrélations des modèles SPARMA.

**Exemple 1.1.3** Soit un processus  $(X_{nT+v})_{n \in \mathbb{Z}}$  satisfaisant un modèle SPARMA $_2(0,1)(0,1)$

$$X_{2n+v} = \epsilon_{2n+v} - a\epsilon_{2n+v-2} - b\epsilon_{2n+v-1} + ba\epsilon_{2n+v-1-2}, \text{ avec } \epsilon_{2n+v} \text{ un bb faible}$$

$$\begin{aligned} \text{Cov}(X_{2n+v}, X_{2n+v-h}) = \\ \text{Cov}(\epsilon_{2n+v} - a\epsilon_{2n+v-2} - b\epsilon_{2n+v-1} + ba\epsilon_{2n+v-1-2}, \epsilon_{2n+v-h} - a\epsilon_{2n+v-2-h} - b\epsilon_{2n+v-1-h} + ba\epsilon_{2n+v-1-2-h}) \\ = \text{Cov}(\epsilon_{2n+v}, \epsilon_{2n+v-h}) - \text{Cov}(\epsilon_{2n+v}, a\epsilon_{2n+v-2-h}) - \text{Cov}(\epsilon_{2n+v}, b\epsilon_{2n+v-1-h}) + \text{Cov}(\epsilon_{2n+v}, ba\epsilon_{2n+v-1-2-h}) \\ - \text{Cov}(a\epsilon_{2n+v-2}, \epsilon_{2n+v-h}) + \text{Cov}(a\epsilon_{2n+v-2}, a\epsilon_{2n+v-2-h}) + \text{Cov}(a\epsilon_{2n+v-2}, b\epsilon_{2n+v-1-h}) - \text{Cov}(a\epsilon_{2n+v-2}, ba\epsilon_{2n+v-1-2-h}) \\ - \text{Cov}(b\epsilon_{2n+v-1}, \epsilon_{2n+v-h}) + \text{Cov}(b\epsilon_{2n+v-1}, a\epsilon_{2n+v-2-h}) + \text{Cov}(b\epsilon_{2n+v-1}, b\epsilon_{2n+v-1-h}) - \text{Cov}(b\epsilon_{2n+v-1}, ba\epsilon_{2n+v-1-2-h}) \\ + \text{Cov}(ba\epsilon_{2n+v-1-2}, \epsilon_{2n+v-h}) - \text{Cov}(ba\epsilon_{2n+v-1-2}, a\epsilon_{2n+v-2-h}) - \text{Cov}(ba\epsilon_{2n+v-1-2}, b\epsilon_{2n+v-1-h}) \\ + \text{Cov}(ba\epsilon_{2n+v-1-2}, ba\epsilon_{2n+v-1-2-h}) \end{aligned}$$

La fonction d'autocovariance du processus s'écrit pour un  $T = 2$  fixé et pour  $v = \{1, 2\}$

$$\gamma_v(h) = \begin{cases} \sigma_v^2(1 + a^2(v) + b^2(v) + a^2(v)b^2(v)) & \text{si } h = 0 \\ -\sigma_{(v)}^2 b(v)(1 + a^2(v)) & \text{si } h = 1 \\ 0 & \text{si } h > 1 \text{ et } T \neq |h| \\ -\sigma_{(v)}^2 a(v)(1 + b^2(v)) & \text{si } h > 1 \text{ et } T = |h| \end{cases}$$

La fonction d'autocorrélation du processus s'écrit pour un  $T$  (on prend  $T = 2$  dans cet exemple) fixé et pour  $v = \{1, 2\}$

$$\rho(h) = \begin{cases} -\frac{\sigma_v b(v)(1+a^2(v))}{\sigma_{v-1}(1+a^2(v)+b^2(v)+a^2(v)b^2(v))} & \text{si } h = 1 \text{ et } T \neq |h| \\ 0 & \text{si } h > 1 \text{ et } T \neq |h| \\ -\frac{\sigma_v a(v)(1+b^2(v))}{\sigma_{v-1}(1+a^2(v)+b^2(v)+a^2(v)b^2(v))} & \text{si } h > 1 \text{ et } T = |h| \end{cases}$$

Nous constatons également que les calculs des fonctions autocovariances et d'autocorrélations du modèles SPARMA, ne dépendent pas que du retard  $h$  mais aussi du valeur prise par la saisonnalité comme pour les SARMA vue dans l'exemple 1.1.1.

Cette classe de modèles contient les modèles SARMA et PARMA comme cas particuliers. Comme pour les modèles PARMA, l'étude des conditions d'inversibilités et de stationnarités, des modèles SPARMA nécessite autant une réécriture des modèles sous forme vectorielle. Les modèles SARMA sont considérés comme des modèles ARMA( $p+Ps, q+Qs$ ), mais avec des trous et une paramétrisation parcimonieuse. De même, les modèles SPARMA( $p, q$ )( $P, Q$ ) peuvent être formulés comme un processus PARMA( $p+PT, q+QT$ ).

$$X_{nT+v} - \sum_{i=1}^{p+PT} a_i^*(v)X_{nT+v-i} = - \sum_{i=1}^{q+QT} b_i^*(v)\epsilon_{nT+v-i} + \epsilon_{nT+v} \quad (1.15)$$



où

$$a_i^*(v) = \begin{cases} a_i(v) & \text{Si } 1 \leq i \leq p \\ 0 & \text{Si } p < i < T \end{cases} \quad a_{Tj}^*(v) = \mathbf{a}_j(v) \text{ pour } 1 \leq j \leq P \text{ et } a_{Tj+i}^*(v) = -a_i(v)\mathbf{a}_j(v)$$

pour  $1 \leq i \leq p, 1 \leq j \leq P$

et

$$b_i^*(v) = \begin{cases} b_i(v) & \text{Si } 1 \leq i \leq q \\ 0 & \text{Si } q < i < T \end{cases} \quad b_{Tj}^*(v) = \mathbf{b}_j(v) \text{ pour } 1 \leq j \leq Q \text{ et } b_{0Tj+i}^*(v) = -b_{0i}(v)\mathbf{b}_j(v)$$

pour  $1 \leq i \leq q, 1 \leq j \leq Q$

Nous réécrivons l'équation (1.15) sous la forme vectorielle VARMA( $p^*, q^*$ ) suivante :

$$A_0 \mathbf{X}_n - \sum_{k=1}^{p^*} A_k \mathbf{X}_{n-k} = B_0 \boldsymbol{\epsilon}_n - \sum_{l=1}^{q^*} B_l \boldsymbol{\epsilon}_{n-l} \quad (1.16)$$

où

$$\mathbf{X}_n = (X_{nT+1}, \dots, X_{nT+T}), \quad \boldsymbol{\epsilon}_n = (\epsilon_{nT+1}, \dots, \epsilon_{nT+T}) \quad (1.17)$$

$p^* = [(p + PT - 1)/T] + 1$ ,  $q^* = [(q + QT - 1)/T] + 1$ , les matrices ( $T \times T$ ) des coefficients de  $A_k, k = 0, \dots, p^*$  et  $B_l, l = 0, \dots, q^*$  sont définies par

$$(A_0)_{i,j} = \begin{cases} 1 & \text{Si } i = j \\ 0 & \text{Si } i < j \\ -a_{i-j}^*(i) & \text{Si } i > j \end{cases}$$

$$(B_0)_{i,j} = \begin{cases} 1 & \text{Si } i = j \\ 0 & \text{Si } i < j \\ -b_{i-j}^*(i) & \text{Si } i > j \end{cases}$$

$$\text{et } (A_k)_{i,j} = a_{kT+i-j}^*(i), \quad k = 1, \dots, p^* \quad (B_l)_{i,j} = b_{lT+i-j}^*(i), \quad l = 1, \dots, q^*$$

Nous avons réécrit le modèle SPARMA sous la forme d'un modèle VARMA. Ainsi l'hypothèse (H1''), nous assure les conditions de stationnarités, d'inversibilités et de causalités du modèle SPARMA.

Il existe peu d'études sur les modèles saisonniers périodiques auto-régressifs moyennes mobiles. Nous avons dans le cas fort, [7], qui ont étudié les conditions de stationnarité et la distribution asymptotique des estimateurs des moindres carrés. Ils ont effectué ce travail dans un cas particulier d'un SPARMA(1,0)(1,0). [80] ont généralisé le travail de [7] au cas de modèles SPVARMA( $p, 0$ )( $P, 0$ ) (pour Seasonal Periodic VARMA). Nous allons les étendre au cadre SPARMA( $p, P$ )( $P, Q$ ) avec un bruit blanc périodique non corrélé mais non nécessairement indépendant.

## 1.2 Contribution de la thèse

L'objectif principal de la thèse est d'étudier de quelle manière il convient d'adapter les outils statistiques standards pour la validation de modèles saisonniers et/ou périodiques.

Le deuxième chapitre est consacré à la validation de modèles SARMA faibles. Dans un premier temps, nous rappelons les propriétés asymptotiques des estimateurs du quasi-maximum de vraisemblance (en anglais QMLE). Ensuite nous étudions la distribution asymptotique jointe

de l'estimateur QMLE et des autocovariances empiriques. Ceci nous permet d'obtenir les distributions asymptotiques des autocovariances et autocorrélations résiduelles. Enfin, nous déduisons le comportement asymptotique des statistiques portmanteau. Dans le cadre standard (c'est-à-dire sous les hypothèses iid sur le bruit), il est connu que la distribution asymptotique des tests portmanteau est approximée par chi-deux. Dans le cas général, nous montrons que cette distribution asymptotique est celle d'une somme pondérée de chi-deux. Cette distribution peut être donc très différente de l'approximation chi-deux usuelle du cas fort. Nous en déduisons des tests portmanteau modifiés pour tester l'adéquation des modèles SARMA faibles. Une deuxième méthode alternative basée sur une approche d'auto-normalisation est proposée dans ce chapitre pour tester l'adéquation des modèles SARMA faibles. L'utilisation de cette technique a pour principal objectif d'éviter l'estimation de la matrice de covariance asymptotique des autocorrélations empiriques. Les valeurs critiques de la statistique de test introduite sont tabulées dans [54]. Nous établissons dans cette partie une extension de résultat de [22] au cas de modèles SARMA.

Le troisième chapitre est consacré aussi à la validation de modèles SVARMA faibles. Nous étendons dans cette partie le travail de [56] aux cas de modèles SVARMA faibles. Nous avons suivi la même méthodologie que le chapitre précédent mais en regardant cette fois ci, que les fonctions autocorrélations qui sont multiples de la saisonnalité ( $2s, 3s \dots$ ).

Dans le quatrième chapitre, nous rappelons dans un premier temps les résultats sur les estimateurs de moindres carrés ordinaires, généralisés et quasi-généralisés des paramètres des modèles PARMA obtenus par [32]. Ensuite nous étudions la distribution asymptotique jointe de l'estimateur moindres carrés quasi-généralisés et des autocovariances empiriques. Ceci nous permet d'obtenir les distributions asymptotiques des autocovariances et autocorrélations résiduelles. Enfin, nous déduisons le comportement asymptotique des statistiques portmanteau. Dans le cadre standard (c'est-à-dire sous les hypothèses iid sur le bruit), il est connu que la distribution asymptotique des tests portmanteau est approximée par chi-deux. Dans le cas général nous montrons que cette distribution asymptotique est celle d'une somme pondérée de chi-deux. Cette distribution peut être très différente de l'approximation chi-deux usuelle du cas fort. Nous en déduisons des tests portmanteau modifiés pour tester l'adéquation des modèles PARMA faibles. De la même manière que le chapitre 2, une deuxième méthode alternative basée sur une approche d'auto-normalisation est proposée dans ce chapitre pour tester l'adéquation des modèles PARMA faibles.

Dans le dernier chapitre nous présentons un modèle saisonnier autorégressif moyenne mobile avec des paramètres périodiques (SPARMA). Dans un premier temps, nous allons étudier les propriétés asymptotiques de l'estimateur de moindres carrés quasi-généralisés des modèles SPARMA induits par un bruit non corrélé mais qui peut contenir des dépendances non linéaires générales. Dans un deuxième, temps nous allons développer la distribution asymptotique des statistiques portmanteau des modèles SPARMA faibles. Nous en déduisons des tests portmanteau modifiés pour tester l'adéquation des modèles SPARMA faibles. Nous proposons également dans cette partie la méthode d'auto-normalisation pour les modèles SPARMA faibles.

## 1.3 Résultats du chapitre 2

Après l'étape d'estimation, une étape importante dans la modélisation des modèles SARMA faibles consiste à vérifier si le modèle estimé ajuste bien correctement les données. Cette étape d'adéquation permet de valider ou d'invalider le choix des ordres du modèle. Ce choix est important pour la précision des prévisions linéaires et pour une bonne interprétation du modèle. L'objectif de ce chapitre est de rappeler dans un premiers temps, les propriétés asymptotiques de l'estimateur du quasi-maximum de vraisemblance (QMLE en anglais) du modèle SARMA obtenus dans [20]. Dans un second temps, nous étudions les distributions asymptotiques des

autocovariances et autocorrélations résiduelles induits par un bruit non corrélé mais qui peut contenir des dépendances non linéaires très générales. Ceci nous permet de développer des versions modifiées de tests portmanteau pour la validation des modèles SARMA faibles. Nous proposons enfin, une deuxième méthode alternative basée sur une approche d'auto-normalisation pour tester l'adéquation des modèles SARMA faibles. Cette méthode à l'avantage de contourner le problème de l'estimation de la matrice de variance du vecteur joint de l'estimateur QML et des autocovariances empiriques.

### 1.3.1 Propriétés asymptotiques de l'estimateur QML

Pour l'estimation des paramètres, nous utiliserons la méthode du quaxi-maximum de vraisemblance, c'est une méthode couramment utilisée, qui peut également être considérée comme une estimation non linéaire.

Soit  $X = (X_t)_{t \in \mathbb{Z}}$  le processus défini dans (1.4). Il est d'usage courant pour le traitement statistique des modèles SARMA de considérer que

$$\theta_0 = (a_{01}, a_{02}, \dots, a_{0p}, b_{01}, b_{02}, \dots, b_{0q}, \mathbf{a}_{01}, \mathbf{a}_{02}, \dots, \mathbf{a}_{0P}, \mathbf{b}_{01}, \mathbf{b}_{02}, \dots, \mathbf{b}_{0Q})'$$

est l'unique paramètre vérifiant l'équation (1.4). Nous garantissons cette condition d'identifiabilité du paramètre par le fait de considérer que les polynômes saisonniers  $\mathbf{a}(\cdot)$  et non saisonniers  $a(\cdot)$  AR ne partagent respectivement aucun facteur en commun avec les polynômes saisonniers  $\mathbf{b}(\cdot)$  et non saisonniers  $b(\cdot)$  MA.

Le vrai paramètre (inconnu)  $\theta_0$  est supposé appartenir à l'espace des paramètres

$$\Theta = \left\{ \theta = (a_1, \dots, a_p, b_1, \dots, b_q, \mathbf{a}_1, \dots, \mathbf{a}_p, \mathbf{b}_1, \dots, \mathbf{b}_q)' \in \mathbb{R}^{k_0}, \text{ où } k_0 = p + q + P + Q, \right. \\ \left. a_\theta(z) = 1 - \sum_{i=1}^p a_i z^i, \mathbf{a}_\theta(z) = 1 - \sum_{j=1}^P \mathbf{a}_j z^{s_j}, b_\theta(z) = 1 - \sum_{i=1}^p b_i z^i \text{ et } \mathbf{b}_\theta(z) = 1 - \sum_{j=1}^Q \mathbf{b}_j z^{s_j} \right. \\ \left. \text{ont leurs zéros en dehors du disque unité} \right\}.$$

**(H2) :** Le processus  $\epsilon = (\epsilon_t)_{t \in \mathbb{Z}}$  est stationnaire et ergodique.

Pour la normalité asymptotique du QMLE, nous avons besoin des hypothèses supplémentaires. Il est d'abord nécessaire de supposer que le paramètre  $\theta_0$  n'est pas situé sur le bord de l'espace de paramètre  $\Theta$ .

**(H3) :** Nous avons  $\theta_0 \in \overset{\circ}{\Theta}$ , où  $\overset{\circ}{\Theta}$  est à l'intérieur de  $\Theta$ .

Pour contrôler la dépendance dans le temps du processus stationnaire  $(\epsilon_t)$ , nous introduisons les coefficients de mélange fort  $\alpha_\epsilon(h)$  définis par

$$\alpha_\epsilon(h) = \sup_{A \in \mathcal{F}_{-\infty}^t, B \in \mathcal{F}_{t+h}^{+\infty}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|,$$

où  $\mathcal{F}_{-\infty}^t = \sigma(\epsilon_u, u \leq t)$  et  $\mathcal{F}_{t+h}^{+\infty} = \sigma(\epsilon_u, u \geq t+h)$ . Nous ferons une hypothèse d'intégrabilité sur le moment du bruit et une condition de sommabilité sur les coefficients de mélange fort  $(\alpha_\epsilon(k))_{k \geq 0}$ .

**(H4) :** Nous avons  $\mathbb{E}|\epsilon_t|^{4+2\nu} < \infty$  et  $\sum_{k=0}^{\infty} \{\alpha_\epsilon(k)\}^{\frac{\nu}{2+\nu}} < \infty$  pour un certain  $\nu > 0$ .

L'hypothèse **(H4)** est une condition technique pour prouver les propriétés asymptotiques du QMLE. L'hypothèse d'intégrabilité sur le moment du bruit n'est pas restrictive dans ce cadre car le processus d'innovation  $(\epsilon_t)$  est directement observé (voir [68]).

Étant donné une observation de longueur  $n$ ,  $X_1, X_2, \dots, X_n$  satisfaisant (1.4), les variables aléatoires peuvent être approximées, pour  $0 < t \leq n$  par  $e_t(\theta)$  définies récursivement par

$$e_t(\theta) = X_t - \sum_{i=1}^p a_i X_{t-i} - \sum_{j=1}^P \mathbf{a}_j X_{t-sj} + \sum_{i=1}^p \sum_{j=1}^P a_i \mathbf{a}_j X_{t-sj-i} + \sum_{i=1}^q b_i e_{t-i}(\theta) + \sum_{j=1}^Q \mathbf{b}_j e_{t-sj}(\theta) - \sum_{i=1}^q \sum_{j=1}^Q b_i \mathbf{b}_j e_{t-sj-i}(\theta), \quad (1.18)$$

où les valeurs initiales inconnues sont mises à zéro :  $e_0(\theta) = \dots = e_{1-q-sQ}(\theta) = X_0 = \dots = X_{1-p-sP} = 0$ . La quasi-vraisemblance est donnée par

$$L_n(\theta, \sigma^2) = \prod_{t=1}^n \frac{1}{(2\pi)^{1/2} \sqrt{\sigma^2}} \exp \left\{ -\frac{e_t^2(\theta)}{2\sigma^2} \right\}.$$

Un QMLE de  $(\theta, \sigma^2)$  est une solution mesurable  $(\hat{\theta}_n, \hat{\sigma}^2)$  de

$$(\hat{\theta}_n, \hat{\sigma}^2) = \arg \min_{\theta, \sigma^2} \left\{ \log(\sigma^2) + \frac{1}{2\sigma^2} Q_n(\theta) \right\} \text{ où } Q_n(\theta) = \frac{1}{n} \sum_{t=1}^n e_t^2(\theta).$$

Dans toute la suite, nous notons par  $\xrightarrow{d}$ , la convergence en distribution. Le symbole  $o_{\mathbb{P}}(1)$  est utilisé pour une suite de variables aléatoires qui convergent vers zéro en probabilité. Sous les hypothèses **(H'1)**, **(H2)**, **(H3)** et **(H4)**, nous avons montré dans [20] presque sûrement que  $\hat{\theta}_n \rightarrow \theta_0$  quand  $n \rightarrow +\infty$  et que :

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = -J^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^n Y_t + o_{\mathbb{P}}(1) \xrightarrow[n \rightarrow +\infty]{d} \mathcal{N}(0, \Sigma = J^{-1} I J^{-1}), \quad (1.19)$$

où

$$J = J(\theta_0) = \frac{2}{\sigma_0^2} \mathbb{E} \left[ \frac{\partial \epsilon_t(\theta_0)}{\partial \theta} \frac{\partial \epsilon_t(\theta_0)}{\partial \theta'} \right], \quad I = I(\theta_0) = \sum_{h=-\infty}^{+\infty} \text{Cov}(Y_t, Y_{t-h}) \text{ et } Y_t = \frac{2}{\sigma_0^2} \epsilon_t(\theta_0) \frac{\partial \epsilon_t(\theta_0)}{\partial \theta}.$$

Notons que l'existence de la matrice  $I(\theta_0)$  est une conséquence de **(H4)** et de l'inégalité de Davydov (voir [27]).

### 1.3.2 La distribution asymptotique des autocovariances et autocorrélations résiduelles

Soit  $\hat{\epsilon}_t = \hat{e}_t = e_t(\hat{\theta}_n)$  les résidus du QMLE quand  $p > 0$ ,  $P > 0$ ,  $q > 0$  et  $Q > 0$ . En utilisant (1.18), nous avons  $\hat{\epsilon}_t = 0$  pour  $t \leq 0$  et  $t > 0$ . De (1.4), nous en déduisons que

$$\hat{\epsilon}_t = \hat{X}_t - \sum_{i=1}^p \hat{a}_i \hat{X}_{t-i} - \sum_{j=1}^P \hat{\mathbf{a}}_j \hat{X}_{t-sj} + \sum_{i=1}^p \sum_{j=1}^P \hat{a}_i \hat{\mathbf{a}}_j \hat{X}_{t-sj-i} + \sum_{i=1}^q \hat{b}_i \hat{\epsilon}_{t-i} + \sum_{j=1}^Q \hat{\mathbf{b}}_j \hat{\epsilon}_{t-sj} - \sum_{i=1}^q \sum_{j=1}^Q \hat{b}_i \hat{\mathbf{b}}_j \hat{\epsilon}_{t-sj-i},$$

pour  $t = 1, \dots, n$ , avec  $\hat{X}_t = 0$  pour  $t \leq 0$  et  $\hat{X}_t = X_t$  pour  $t \geq 1$ .

Pour un entier  $m \geq 1$  fixé, définissons les autocovariances résiduelles

$$\hat{\gamma}_m = (\hat{\gamma}(1), \dots, \hat{\gamma}(m))' \text{ où } \hat{\gamma}(h) = \frac{1}{n} \sum_{t=h+1}^n \hat{\epsilon}_t \hat{\epsilon}_{t-h} \text{ pour } 0 \leq h < n.$$

Dans la suite nous aurons également besoin du vecteur des  $m$  premières autocorrélations résiduelles

$$\hat{\rho}_m = (\hat{\rho}(1), \dots, \hat{\rho}(m))' \text{ où } \hat{\rho}(\ell) = \hat{\gamma}(\ell) / \hat{\gamma}(0)$$

Le test portmanteau a été introduit par [24] (noté **BP**) afin de mesurer la qualité d'ajustement d'un modèle ARMA fort. Ensuite [53] (noté **LB**) ont proposé une version modifiée du test portmanteau de **BP** qui est à nos jours l'un des outils de diagnostic le plus populaire dans la modélisation des séries chronologiques. En se basant sur les autocorrélations empiriques résiduelles du bruit, leurs statistiques de test sont définies respectivement par

$$Q_m^{\text{BP}} = n \sum_{h=1}^m \hat{\rho}^2(h) \text{ et } Q_m^{\text{LB}} = n(n+2) \sum_{h=1}^m \frac{\hat{\rho}^2(h)}{n-h}. \quad (1.20)$$

Ces statistiques sont souvent utilisées pour tester l'hypothèse nulle

( $\mathcal{H}0$ ) :  $(X_t)_{t \in \mathbb{Z}}$  satisfait une représentation SARMA( $p, q$ )( $P, Q$ ) $_s$ ;

contre l'alternative

( $\mathcal{H}1$ ) :  $(X_t)_{t \in \mathbb{Z}}$  n'admet pas de représentation SARMA( $p, q$ )( $P, Q$ ) $_s$  ou  $(X_t)_{t \in \mathbb{Z}}$  satisfait une représentation SARMA( $p', q'$ )( $P', Q'$ ) $_s$  avec  $p' > p$  ou  $q' > q$  ou  $P' > P$  ou  $Q' > Q$ .

Ces tests sont des outils très utiles pour vérifier la signification globale des autocorrélations résiduelles.

Nous définissons les autocovariances empiriques, avant d'énoncer les résultats qui donnent les distributions jointes de  $\hat{\theta}_n$  et des autocovariances empiriques, et les distributions des autocorrélations et autocovariances résiduelles.

Notons tout d'abord que l'hypothèse (**H4**) impliquera la normalité asymptotique des autocovariances "empiriques"

$$\gamma_m = (\gamma(1), \dots, \gamma(m))' \text{ où } \gamma(h) = \frac{1}{n} \sum_{t=h+1}^n \epsilon_t \epsilon_{t-h} \text{ pour } 0 \leq h < n.$$

Il est à noter que les  $\gamma(h)$  ne sont pas de statistiques (saut si  $p = P = q = Q = 0$ ) puisqu'elles dépendent des innovations théoriques  $\epsilon = \epsilon(\theta_0)$ .

Un développement de Taylor autour de  $\theta_0$  donne

$$\begin{aligned} \hat{\gamma}(h) &= \gamma(h) + \frac{1}{N} \sum_{t=h+1}^n \left\{ \epsilon_{t-h} \frac{\partial \epsilon_t(\theta)}{\partial \theta'} + \frac{\partial \epsilon_{t-h}(\theta)}{\partial \theta'} \epsilon_t(\theta) \right\}_{\theta=\theta_n^*} (\hat{\theta}_n - \theta_0) + o_{\mathbb{P}}(1/n) \\ &= \gamma(h) + \mathbb{E} \left( \epsilon_{t-h}(\theta_0) \frac{\partial \epsilon_t(\theta)}{\partial \theta'} \right) (\hat{\theta}_n - \theta_0) + o_{\mathbb{P}}(1/n), \end{aligned}$$

où  $\theta_n^*$  est compris entre  $\hat{\theta}_n$  et  $\theta_0$ . La dernière égalité découle de la consistance de  $\hat{\theta}_n$  et du fait que  $(\partial \epsilon_{t-h} / \partial \theta')(\theta_0)$  n'est pas corrélé avec  $\epsilon_t$  quand  $h \geq 0$ . Ainsi, pour  $h = 1, \dots, m$ , nous avons

$$\hat{\gamma}_m := (\hat{\gamma}(1), \dots, \hat{\gamma}(m))' = \gamma_m + \Phi_m (\hat{\theta}_n - \theta_0) + o_{\mathbb{P}}(1/n), \quad (1.21)$$

où

$$\Phi_m = \mathbb{E} \left\{ \left( \begin{array}{c} \epsilon_{t-1} \\ \vdots \\ \epsilon_{t-m} \end{array} \right) \frac{\partial \epsilon_t}{\partial \theta'} \right\}. \quad (1.22)$$

Nous en déduisons également que

$$\sqrt{n} \hat{\rho}_m = \sqrt{n} \frac{\hat{\gamma}_m}{\sigma_\epsilon^2} + o_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right). \quad (1.23)$$

Ainsi, de (1.23) on remarque que la distribution asymptotique des autocorrélations résiduelles  $\sqrt{n}\hat{\rho}_m$  dépend de la distribution de  $\hat{\gamma}_m$ . D'après l'équation (1.21), l'étude de la distribution jointe du vecteur  $\sqrt{n}(\hat{\theta}'_n - \theta'_0, \gamma'_m)'$  nous permettra de déduire la loi asymptotiques des autocovariances résiduelles  $\sqrt{n}\hat{\gamma}_m$ . Définissons la matrice

$$\Xi = \begin{pmatrix} \Sigma & \Sigma_{\hat{\theta}_n, \gamma_m} \\ \Sigma'_{\hat{\theta}_n, \gamma_m} & \Sigma_{\gamma_m} \end{pmatrix} = \sum_{h=-\infty}^{+\infty} \mathbb{E} w_t w'_{t-h}, \text{ avec } w_t = \begin{pmatrix} w_{1t} \\ w_{2t} \end{pmatrix} \in \mathbb{R}^{k_0+m} \quad (1.24)$$

où  $w_{1t} = -2\sigma_0^{-2}\epsilon_t J^{-1}(\partial\epsilon_t/\partial\theta)$  et  $w_{2t} = (\epsilon_{t-1}, \dots, \epsilon_{t-m})'\epsilon_t$ . Notons que l'existence de la matrice  $\Xi$  est une conséquence de (H4) et de l'inégalité de Davydov (voir [27]). En appliquant le théorème central limite pour le processus mélangeant (voir [42]), nous avons

$$\sqrt{n}(\hat{\theta}'_n - \theta'_0, \gamma'_m)' = \frac{1}{\sqrt{n}} \sum_{t=1}^n w_t \xrightarrow[n \rightarrow +\infty]{d} \mathcal{N}(0, \Xi).$$

La proposition suivante, qui est une généralisation du théorème 3 obtenu par [20], fournit la distribution asymptotique des autocovariances et autocorrélations résiduelles de modèles SARMA.

**Proposition 1.3.1** *Quand,  $p > 0$ ,  $P > 0$ ,  $q > 0$  et  $Q > 0$ , sous certaines hypothèses, nous avons*

$$\sqrt{n}\hat{\gamma}_m \xrightarrow[n \rightarrow +\infty]{d} \mathcal{N}(0, \Sigma_{\hat{\gamma}_m}) \quad \text{et} \quad \sqrt{n}\hat{\rho}_m \xrightarrow[n \rightarrow +\infty]{d} \mathcal{N}(0, \Sigma_{\hat{\rho}_m}) \quad \text{où,}$$

$$\Sigma_{\hat{\gamma}_m} = \Sigma_{\gamma_m} + \Phi_m \Sigma \Phi'_m + \Phi_m \Sigma_{\hat{\theta}_n, \gamma_m} + \Sigma'_{\hat{\theta}_n, \gamma_m} \Phi'_m \quad \text{et} \quad \Sigma_{\hat{\rho}_m} = \frac{1}{\sigma_0^4} \Sigma_{\hat{\gamma}_m}.$$

La preuve de cette proposition est similaire à celle donnée par [20, théorème 3]. Les matrices de variance asymptotique  $\Sigma_{\hat{\gamma}_m}$  et  $\Sigma_{\hat{\rho}_m}$  dépendent des matrices inconnues  $\Xi$ ,  $\Phi_m$  et du scalaire  $\sigma_0^2$ . Les matrices  $\Phi_m$  et  $\sigma_0^2$  peuvent être estimées empiriquement par

$$\hat{\Phi}_m = \frac{1}{n} \sum_{t=1}^n \left\{ (\hat{\epsilon}_{t-1}, \dots, \hat{\epsilon}_{t-m})' \frac{\partial \hat{\epsilon}_t}{\partial \theta'} \right\} \quad \text{and} \quad \hat{\sigma}^2 = \hat{\gamma}(0) = \frac{1}{n} \sum_{t=1}^n \hat{\epsilon}_t^2.$$

Notons que la matrice  $(2\pi)^{-1}\Xi$  peut être interprété comme la densité spectrale du processus stationnaire  $(w_t)$  évalué à la fréquence zero, ainsi un estimateur de  $\Xi$  est donné dans le Théorème 6 de [20]. d'autres estimateurs de cette variance, basée sur l'estimation de la densité spectrale par une autorégression paramétrique existent dans la littérature (voir par exemple [5], [10], [28], [63], pour les références générales). Dans ce chapitre, nous avons utilisé un estimateur spectral VAR (Vector AR) donné dans le Théorème 6 de [20] consistant à : i) ajuster les modèles VAR( $r$ ) pour  $r = 1, \dots, r_{\max}$  à la série  $\hat{w}_t$ ,  $t = 1, \dots, n$ , où  $\hat{w}_t$  est obtenu en remplaçant  $\theta_0$  par  $\hat{\theta}_n$  dans  $w_t$ ; ii) en sélectionnant l'ordre  $r$  qui minimise un critère d'information et en approximant  $\Xi$  par  $(2\pi)$  fois la densité spectrale à la fréquence zéro du modèle VAR( $r$ ) estimé. Par la suite, nous avons utilisé le critère de sélection du modèle AIC avec  $r_{\max} = 5$ .

### 1.3.3 Version modifiée du test portmanteau

A partir de la proposition 1.3.1 nous déduisons le résultat suivant, qui donne la distribution asymptotique exacte des statistiques (1.20) sous des hypothèses faibles sur l'innovation du modèle SARMA.

**Theorem 1.3.2** *Sous les hypothèses de la proposition 1.3.1 et  $(\mathcal{H}0)$ , les statistiques  $Q_m^{\text{LB}}$  et  $Q_m^{\text{BP}}$  introduites dans (1.20) convergent en distribution, quand  $n \rightarrow +\infty$ , vers*

$$Z_m(\xi_m) = \sum_{i=1}^m \xi_{i,m} Z_i^2$$

où  $\xi_m = (\xi_{1,m}, \dots, \xi_{m,m})'$  est le vecteur des valeurs propres de la matrice  $\Sigma_{\hat{\rho}_m} = \sigma_0^{-4} \Sigma_{\hat{\gamma}_m}$  et  $Z_1, \dots, Z_m$  sont des variables indépendantes et identiquement distribuées de la loi  $\mathcal{N}(0, 1)$ .

**Remarque 1.3.1** *Dans le cas de modèles SARMA forts, sous l'hypothèse nulle  $(\mathcal{H}0)$ , les distributions asymptotiques des statistiques  $Q_m^{\text{BP}}$  et  $Q_m^{\text{LB}}$  sont approximées par un  $\chi_{m-(p+q+P+Q)}^2$  où  $m > p + q + P + Q$  et  $\chi_k^2$  est la distribution de chi-deux à  $k$  degrés de liberté. Le théorème 1.3.2 montre que cette approximation n'est plus valable et est très différente dans le cas de modèles SARMA( $p, q$ )( $P, Q$ )<sub>s</sub> faibles.*

Les versions modifiées des tests portmanteau de Box–Pierce (BP) et de Ljung–Box (LB), que nous nous apprêtons à introduire, visent à détecter si les ordres  $p, P, q$  et  $Q$  d'un modèle SARMA( $p, q$ )( $P, Q$ )<sub>s</sub> faible sont bien choisis. Les versions modifiées des tests portmanteau seront notées  $\text{BP}_w$  et  $\text{LB}_w$ , l'indice  $w$  (pour weak en anglais) fait référence au terme faible.

La vraie distribution asymptotique dépend de paramètres de nuisance  $\sigma_0^2$ , la matrice  $\Phi_m$  et les éléments de la matrice  $\Xi$ . Par conséquent, afin d'obtenir la distribution asymptotique des statistiques de Portmanteau (1.20) sous des hypothèses faibles sur le bruit, il faut un estimateur consistant de la matrice de covariance asymptotique  $\Sigma_{\hat{\rho}_m}$ . Notons par  $\hat{\Sigma}_{\hat{\rho}_m}$  la matrice obtenue en remplaçant  $\Xi$  par  $\hat{\Xi}$  et  $\sigma_0^2$  par  $\hat{\sigma}^2$  dans  $\Sigma_{\hat{\rho}_m}$ . Soit  $\hat{\xi}_m = (\hat{\xi}_{1,m}, \dots, \hat{\xi}_{m,m})'$  le vecteur des valeurs propres de  $\hat{\Sigma}_{\hat{\rho}_m}$ .

Au seuil asymptotique  $\alpha$ , le test  $\text{LB}_w$  (respectivement le test  $\text{BP}_w$ ) consiste à rejeter l'hypothèse nulle d'un modèle SARMA( $p, q$ )( $P, Q$ )<sub>s</sub> faible quand

$$Q_m^{\text{LB}} > S_m(1 - \alpha) \quad (\text{respectivement } Q_m^{\text{BP}} > S_m(1 - \alpha)),$$

où  $S_m(1 - \alpha)$  est tel que  $\mathbb{P}\{Z_m(\hat{\xi}_m) > S_m(1 - \alpha)\} = \alpha$ . Nous soulignons le fait que les versions modifiées des tests portmanteau BP et LB sont plus difficiles à mettre en œuvre. En effet, les valeurs critiques de loi limite de leurs statistiques doivent être calculées à partir des données. Alors que les valeurs critiques de la méthode standard sont, quand à elles, simplement déduites d'une table  $\chi^2$ .

### 1.3.4 La distribution asymptotique auto-normalisée des autocorrélations résiduelles

Nous avons vu précédemment que l'obtention d'un estimateur convergent de la matrice asymptotique  $\Sigma_{\hat{\rho}_m}$  du vecteur des autocorrélations résiduelles nécessite le développement d'un estimateur au moins faiblement consistant de la matrice  $\Xi$ . Cette matrice, contrairement à  $\Phi_m$  et au scalaire  $\sigma_0^2$ , ne peut pas être estimée empiriquement. Nous avons proposé une approche paramétrique basée sur une estimation AR de la densité spectrale ( $w_t$ ) étudié par [10], [18] et [15]. Cependant, cette approche pose le problème du choix du paramètre de troncature. En effet, cette méthode est basée sur une représentation AR( $\infty$ ) du processus  $(w_t)_{t \in \mathbb{Z}}$ . Le choix de l'ordre de troncature est donc crucial et difficile.

Dans cette sous-section, comme dans [22], nous proposons une méthode alternative où nous n'estimons pas la matrice de covariance asymptotique. Elle est basée sur une approche d'auto-normalisation pour construire une statistique de test dont la distribution asymptotique ne dépend pas de paramètres inconnus (voir [22] qui ont étudié cette approche dans le cas de modèles ARMA faibles). L'idée vient de [54] et a déjà été étendue par [51, 73, 74, 75] dans un cadre



plus général. [22] ont proposé des versions modifiées des statistiques portmanteau standards de Box–Pierce et de Ljung–Box. En plus du fait qu'elle permet de contourner le problème de l'estimation de la matrice de variance asymptotique  $\Xi$ , cette technique nous épargne le calcul (à partir des données) des valeurs critiques de la distribution limite des statistiques de tests. En effet, les valeurs critiques de la loi limite de la statistique de test proposée par la méthode basée sur l'approche d'auto-normalisation sont tabulées dans [54]. Dans un certain sens, cette méthode est finalement plus proche de la méthode standard dans laquelle les valeurs critiques sont simplement déduites d'une table  $\chi^2$ .

Notons par  $\Lambda$  la matrice par blocs de  $\mathbb{R}^{m \times (k_0+m)}$  définis par  $\Lambda = (\Phi_m | I_m)$  où  $I_k$  est la matrice d'identité d'ordre  $k$ . De (1.19) et (1.21), nous déduisons que

$$\sqrt{n} \hat{\Upsilon}_m = n^{-1/2} \sum_{t=1}^n \Lambda w_t + o_{\mathbb{P}}(1). \quad (1.25)$$

A ce stade, nous ne considérons pas la méthode qui consiste à estimer la matrice de covariance asymptotique  $\Lambda w_t$ . Nous avons besoin d'appliquer le théorème central limite fonctionnel pour le processus  $\Lambda w = (\Lambda w_t)_{t \geq 1}$  (voir le lemme 1 dans [54]). Pour cela, nous définissons la matrice de normalisation  $C_m \in \mathbb{R}^{m \times m}$  par

$$C_m = \frac{1}{n^2} \sum_{t=1}^n S_t S_t' \text{ où } S_t = \sum_{j=1}^t (\Lambda w_j - \Upsilon_m).$$

Pour assurer l'inversibilité de la matrice  $C_m$ , qui est démontrée dans le lemme 6 dans [22], nous avons besoin de l'hypothèse technique suivante sur la distribution de  $\epsilon_t$ .

**(H5) :** Le processus  $\epsilon = (\epsilon_t)_{t \in \mathbb{Z}}$  a une densité strictement positive sur un voisinage de zéro.

Soit  $(B_K(r))_{r \geq 0}$  un mouvement brownien  $K$ -dimensionnel partant de 0. Pour  $K \geq 1$ , nous désignons  $\mathcal{U}_K$  la variable aléatoire définie par :

$$\mathcal{U}_K = B_K'(1) V_K^{-1} B_K(1) \text{ où } V_K = \int_0^1 (B_K(r) - r B_K(1)) (B_K(r) - r B_K(1))' dr. \quad (1.26)$$

Les valeurs critiques  $\mathcal{U}_K$  ont été tabulées par [54].

Le théorème suivant énonce la distribution asymptotique auto-normalisée des autocovariances et autocorrélations résiduelles.

**Theorem 1.3.3** *Supposons  $p > 0$ ,  $q > 0$ ,  $P > 0$  ou  $Q > 0$ . Sous les hypothèse de la Proposition 1.3.1, (H5) et sous l'hypothèse nulle ( $\mathcal{H}_0$ ), nous avons*

$$n \hat{\Upsilon}_m' C_m^{-1} \hat{\Upsilon}_m \xrightarrow[n \rightarrow \infty]{d} \mathcal{U}_m \text{ et } n \sigma_0^4 \hat{\rho}_m' C_m^{-1} \hat{\rho}_m \xrightarrow[n \rightarrow \infty]{d} \mathcal{U}_m$$

La preuve de ce théorème est similaire à celle donnée dans [22] dans le cas des modèles ARMA faibles.

Bien entendu, le théorème 1.3.3 est inutilisable dans la pratique car la matrice de normalisation  $C_m$  et le paramètre de nuisance  $\sigma_0^2$  ne sont pas observables. En pratique, nous remplaçons la matrice  $C_m$  et la variance du bruit  $\sigma_0^2$  par leurs parties empiriques ou observables. La matrice  $J$  peut être facilement estimé par sa partie empirique,

$$\hat{J} = \frac{2}{\hat{\sigma}^2} \frac{1}{n} \sum_{t=1}^n \frac{\partial e_t(\hat{\theta}_n)}{\partial \theta} \frac{\partial e_t(\hat{\theta}_n)}{\partial \theta'}.$$



Ainsi, nous définissons

$$\hat{\Lambda} = \left( \hat{\Phi}_m | I_m \right) \text{ et } \hat{w}_t = \left( -2\hat{J}_n^{-1} \frac{\partial e_t(\hat{\theta}_n)}{\partial \theta'} \frac{1}{\hat{\sigma}^2} \hat{e}_t, \hat{e}_t \hat{e}_{t-1}, \dots, \hat{e}_t \hat{e}_{t-m} \right)'.$$

Enfin, nous désignons la matrice de normalisation  $\hat{C}_m \in \mathbb{R}^{m \times m}$  par

$$\hat{C}_m = \frac{1}{n^2} \sum_{t=1}^n \hat{S}_t \hat{S}_t' \text{ où } \hat{S}_t = \sum_{j=1}^m (\hat{\Lambda} \hat{w}_j - \hat{Y}_m).$$

les quantités ci-dessus sont observables et nous pouvons donc énoncer la version applicable du théorème 1.3.3.

**Theorem 1.3.4** *Supposons  $p > 0$ ,  $q > 0$ ,  $P > 0$  ou  $Q > 0$ . Sous les hypothèses du théorème 1.3.3, nous avons*

$$n \hat{Y}_m' \hat{C}_m^{-1} \hat{Y}_m \xrightarrow[n \rightarrow \infty]{d} \mathcal{U}_m \text{ et } Q_m^{\text{SN}} = n \hat{\sigma}^4 \hat{\rho}_m' \hat{C}_m^{-1} \hat{\rho}_m \xrightarrow[n \rightarrow \infty]{d} \mathcal{U}_m.$$

La preuve de ce théorème est similaire à celle donnée dans [22] dans le cas des modèles ARMA faibles.

Sur la base du résultat ci-dessus, nous proposons la version modifiée de la statistique de Ljung-Box suivante :

$$\tilde{Q}_m^{\text{SN}} = n \hat{\sigma}^4 \hat{\rho}_m' D_{n,m}^{1/2} \hat{C}_m^{-1} D_{n,m}^{1/2} \hat{\rho}_m, \quad (1.27)$$

où  $D_{n,m} \in \mathbb{R}^{m \times m}$  est une matrice diagonale avec  $((n+2)/(n-1), \dots, (n+2)/(n-m))$  sur la diagonale.

## 1.4 Résultats du chapitre 3

Dans ce chapitre, nous étendons le travail de [56] au cas de modèles SVARMA faibles. L'objectif principal consiste à améliorer les résultats d'analyse statistique de modèle SVARMA en considérant l'étape de la validation du modèle avec des termes d'erreurs non corrélés mais qui peuvent contenir des dépendances non linéaires. Nous allons d'abord présenter brièvement les résultats sur l'estimateur du quasi-maximum de vraisemblance du modèle considéré. Nous énonçons des conditions assurant la consistance et la normalité asymptotique du QMLE. Nous obtenons les distributions asymptotiques des autocovariances et des autocorrélations résiduelles. Ces autocorrélations sont normalement distribuées et nous en déduisons le comportement asymptotique du test portmanteau. Il est démontré que les tests portmanteau standards (c'est-à-dire sous les hypothèses iid sur le bruit), peuvent être assez trompeurs. Dans le cas général, nous montrons que cette distribution asymptotique est celle d'une somme pondérée de chi-deux. Cette distribution peut être très différente de l'approximation de chi-deux usuelle du cas fort.

### 1.4.1 Estimation du quasi-maximum de vraisemblance

Soit  $X = (X_t)_{t \in \mathbb{Z}}$  un processus stationnaire de dimension  $d$ , satisfaisant une représentation vectorielle saisonnière multiplicative autorégressive moyenne mobile (SVARMA)

$$\Lambda_{\theta_0}(L^s) A_{\theta_0}(L) X_t = \Psi_{\theta_0}(L^s) B_{\theta_0}(L) \epsilon_t, \quad \forall t \in \mathbb{Z}, \quad (1.28)$$

où les polynômes AR et MA sont définis respectivement par  $A_{\theta_0}(L) = A_{00} - \sum_{i=1}^p A_{0i}L^i$  et  $B_{\theta_0}(L) = B_{00} - \sum_{i=1}^q B_{0i}L^i$ , tandis les polynômes saisonniers AR et MA sont donnés par  $\Lambda_{\theta_0}(L) = \Lambda_{00} - \sum_{i=1}^{p_s} \Lambda_{0i}L^{s_i}$  et  $\Psi_{\theta_0}(L) = \Psi_{00} - \sum_{i=1}^{q_s} \Psi_{0i}L^{s_i}$ . Le vrai paramètre inconnu est noté  $\theta_0$ . Sans perte de généralité, supposons les matrices  $A_{0p} \neq 0$  ou  $B_{0q} \neq 0$  ou  $\Lambda_{0p_s} \neq 0$  ou  $\Psi_{0q_s} \neq 0$ . Remarquons qu'en général, les polynômes matriciels saisonniers et non saisonniers ne commutent pas, et donc l'ordre dans lequel sont placées les matrices dans l'équation (1.28) est important.

L'équation (1.28) notée  $\text{SVARMA}(p, q)(p_s, q_s)$ , admet une représentation SVARMA faible si  $\epsilon_t$  vérifie l'hypothèse suivante :

$$\mathbf{(H6)}: \quad \mathbb{E}(\epsilon_t) = 0, \text{Var}(\epsilon_t) = \Sigma_0 \text{ et } \text{Cov}(\epsilon_t, \epsilon_{t-h}) = 0 \quad \forall h, t \in \mathbb{Z}, h \neq 0,$$

où la matrice de covariance  $\Sigma_0$  est non singulière. En général on dit que  $(X_t)$  est un modèle SVARMA fort si  $(\epsilon_t)$  est supposé être une suite iid. Un bruit blanc fort est évidemment un bruit blanc faible, car l'indépendance implique l'absence de corrélation, mais l'inverse n'est pas vrai. Soit  $I_d$  la matrice d'identité  $d \times d$ . Lorsqu'on suppose  $A_{00} = B_{00} = \Lambda_{00} = \Psi_{00} = I_d$ , la représentation (1.28) est dite réduite et structurelle dans le cas contraire. Les formes structurelles sont principalement utilisées en économétrie pour introduire des relations instantanées entre les variables économiques. Bien entendu, des contraintes sont nécessaires pour l'identifiabilité de ces représentations.

Soit  $[A_{00} \dots A_{0p} B_{00} \dots B_{0q} \Lambda_{00} \dots \Lambda_{0p_s} \Psi_{00} \dots \Psi_{0q_s}]$  la matrice de taille  $d^2(p + q + p_s + q_s + 4)$  constituée des coefficients de polynômes VAR et MA dans le modèle (1.28). Nous considérons la matrice  $\Sigma_0$  comme un paramètre de nuisance. Le vecteur des vraies valeurs des paramètres noté  $\theta_0$  appartient à l'espace compact des paramètres  $\Theta \subset \mathbb{R}^{k_0}$ , et  $k_0$  est le nombre de paramètres inconnus qui est inférieur à  $d^2(p + q + p_s + q_s + 4)$ .

Les matrices  $A_{00}, \dots, A_{0p}, B_{00}, \dots, B_{0q}, \Lambda_{00}, \dots, \Lambda_{0p_s}, \Psi_{00}, \dots, \Psi_{0q_s}$  contenues dans le modèle (1.28) sont paramétrées suivant le vecteur des vrais paramètres  $\theta_0$ . Plus précisément, nous écrivons  $A_{0i} = A_i(\theta_0)$ ,  $B_{0j} = B_j(\theta_0)$ ,  $\Lambda_{0k} = \Lambda_k(\theta_0)$  et  $\Psi_{0\ell} = \Psi_\ell(\theta_0)$  pour  $i = 0, \dots, p$ ,  $j = 0, \dots, q$ ,  $k = 0, \dots, p_s$  et  $\ell = 0, \dots, q_s$ .

Supposons que cette paramétrisation vérifie les conditions de régularité suivantes.

$$\mathbf{(H7)}: \quad \text{Les applications } \theta \mapsto A_i(\theta) \quad i = 0, \dots, p, \theta \mapsto B_j(\theta) \quad j = 0, \dots, q, \theta \mapsto \Lambda_k(\theta) \quad k = 0, \dots, p_s \text{ et } \theta \mapsto \Psi_\ell(\theta) \quad \ell = 0, \dots, q_s \text{ admettent des dérivées continues d'ordres 3 pour tout } \theta \in \Theta.$$

Par souci de simplicité, nous écrivons  $A_i$ ,  $B_j$ ,  $\Lambda_k$  et  $\Psi_\ell$  au lieu de  $A_i(\theta)$ ,  $B_j(\theta)$ ,  $\Lambda_k(\theta)$  et  $\Psi_\ell(\theta)$ . Soit  $A_\theta(z) = A_0 - \sum_{i=1}^p A_i z^i$ ,  $B_\theta(z) = B_0 - \sum_{i=1}^q B_i z^i$ ,  $\Lambda_\theta(z) = \Lambda_0 - \sum_{i=1}^{p_s} \Lambda_i z^{s_i}$  et  $\Psi_\theta(z) = \Psi_0 - \sum_{i=1}^{q_s} \Psi_i z^{s_i}$ . Nous supposons que  $\Theta$  correspond à des représentations stables et inversibles, à savoir

$$\mathbf{(H8)}: \quad \text{pour tout } \theta \in \Theta, \text{ nous avons } \det A_\theta(z) \det B_\theta(z) \det \Lambda_\theta(z) \det \Psi_\theta(z) \neq 0 \text{ pour tout } |z| \leq 1, \text{ Pour l'estimation des modèles SVARMA faibles, nous utiliserons la méthode de l'estimation du quasi-maximum de vraisemblance. Nous disposons des observations } X_1, \dots, X_n \text{ d'un processus satisfaisant la représentation (1.28). Notons que (H8) nous assure l'inversibilité des matrices } A_{00}, B_{00}, \Lambda_{00} \text{ et } \Psi_{00}. \text{ Nous introduisons le processus d'innovation}$$

$$e_t = A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{00} B_{00} \epsilon_t,$$

et écrivons la représentation (1.28) SVARMA structurelle sous forme SVARMA réduite

$$\begin{aligned} X_t & - \sum_{i=1}^p A_{00}^{-1} A_{0i} X_{t-i} - \sum_{j=1}^{p_s} A_{00}^{-1} \Lambda_{00}^{-1} \Lambda_{0j} A_{00} X_{t-s_j} + \sum_{i=1}^p \sum_{j=1}^{p_s} A_{00}^{-1} \Lambda_{00}^{-1} \Lambda_{0j} A_{0i} X_{t-i-s_j} \\ & = e_t - \sum_{i=1}^q A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{00} B_{0i} B_{00}^{-1} \Psi_{00}^{-1} \Lambda_{00} A_{00} e_{t-i} - \sum_{j=1}^{q_s} A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{0j} \Psi_{00}^{-1} \Lambda_{00} A_{00} e_{t-s_j} \\ & + \sum_{i=1}^q \sum_{j=1}^{q_s} A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{0j} B_{0i} B_{00}^{-1} \Psi_{00}^{-1} \Lambda_{00} A_{00} e_{t-i-s_j}. \end{aligned} \quad (1.29)$$

Pour  $t = 1, \dots, n$  et pour tout  $\theta \in \Theta$ , les variables aléatoires  $\tilde{e}_t(\theta)$  sont définies récursivement par

$$\begin{aligned} \tilde{e}_t(\theta) = & X_t - \sum_{i=1}^p A_0^{-1} A_i X_{t-i} - \sum_{j=1}^{p_s} A_0^{-1} \Lambda_0^{-1} \Lambda_j A_0 X_{t-sj} + \sum_{i=1}^p \sum_{j=1}^{p_s} A_0^{-1} \Lambda_0^{-1} \Lambda_j A_i X_{t-i-sj} \\ & + \sum_{i=1}^q A_0^{-1} \Lambda_0^{-1} \Psi_0 B_i B_0^{-1} \Psi_0^{-1} \Lambda_0 A_0 \tilde{e}_{t-i}(\theta) + \sum_{j=1}^{q_s} A_0^{-1} \Lambda_0^{-1} \Psi_j \Psi_0^{-1} \Lambda_0 A_0 \tilde{e}_{t-sj}(\theta) \\ & - \sum_{i=1}^q \sum_{j=1}^{q_s} A_0^{-1} \Lambda_0^{-1} \Psi_j B_i B_0^{-1} \Psi_0^{-1} \Lambda_0 A_0 \tilde{e}_{t-i-sj}(\theta). \end{aligned} \quad (1.30)$$

avec des valeurs initiales  $\tilde{e}_0(\theta) = \tilde{e}_{-1}(\theta) = \dots = \tilde{e}_{-q-q_s+1}(\theta) = X_0 = X_{-1} = \dots = X_{-p-p_s+1} = 0$ . Nous montrons que ces valeurs initiales sont asymptotiquement négligeables et en particulier, que

$\tilde{e}_t(\theta_0) - e_t \rightarrow 0$  presque sûrement quand  $t \rightarrow +\infty$  (voir [18]). La quasi-vraisemblance gaussienne s'écrit

$$\tilde{L}_n(\theta, \Sigma_e) = \prod_{t=1}^n \frac{1}{(2\pi)^{d/2} \sqrt{\det \Sigma_e}} \exp \left\{ -\frac{1}{2} \tilde{e}'_t(\theta) \Sigma_e^{-1} \tilde{e}_t(\theta) \right\},$$

où

$$\Sigma_e = \Sigma_e(\theta) = A_0^{-1} \Lambda_0^{-1} \Psi_0 B_0 \Sigma B_0' \Psi_0' \Lambda_0^{-1'} A_0^{-1'}.$$

Il est à noter que la variance de  $e_t$  est  $\Sigma_{e0} = \Sigma_e(\theta_0) = A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{00} B_{00} \Sigma_0 B_{00}' \Psi_{00}' \Lambda_{00}^{-1'} A_{00}^{-1'}$ . Des estimateur du QMLE de  $\theta$  et  $\Sigma_e$  sont définis comme toute solution mesurable  $(\hat{\theta}_n, \hat{\Sigma}_e)$  de

$$(\hat{\theta}_n, \hat{\Sigma}_e) = \arg \max_{\theta \in \Theta, \Sigma_e} \tilde{L}_n(\theta, \Sigma_e) = \arg \min_{\theta \in \Theta, \Sigma_e} \tilde{\ell}_n(\theta, \Sigma_e), \quad \tilde{\ell}_n(\theta, \Sigma_e) = \frac{-2}{n} \log \tilde{L}_n(\theta, \Sigma_e).$$

## 1.4.2 Propriétés asymptotiques de l'estimateur

Pour établir les propriétés asymptotiques de QMLE,  $\hat{\theta}_n$ , nous utiliserons les hypothèses suivantes

**(H9) :** Nous avons  $\theta_0 \in \Theta$ , où  $\Theta$  est compact.

**(H10) :** Le processus  $(\epsilon_t)$  est stationnaire et ergodique.

**(H11) :** Pour tout  $\theta \in \Theta$  tel que  $\theta_0 \neq \theta$ , soient les fonctions de transfert

$$A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{00} B_{00} B_{00}^{-1}(z) \Psi_{00}^{-1}(z^s) \Lambda_{\theta_0}(z^s) A_{\theta_0}(z) \neq A_0^{-1} \Lambda_0^{-1} \Psi_0 B_0 B_0^{-1}(z) \Psi_0^{-1}(z^s) \Lambda_{\theta}(z^s) A_{\theta}(z)$$

$$\text{pour un } z \in \mathbb{C}, \text{ or } A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{00} B_{00} \Sigma_0 B_{00}' \Psi_{00}' \Lambda_{00}^{-1'} A_{00}^{-1'} \neq A_0^{-1} \Lambda_0^{-1} \Psi_0 B_0 \Sigma B_0' \Psi_0' \Lambda_0^{-1'} A_0^{-1'}.$$

**Remarque 1.4.1** Dans la représentation SVARMA réduite (1.29), la condition

$$A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{00} B_{00} \Sigma_0 B_{00}' \Psi_{00}' \Lambda_{00}^{-1'} A_{00}^{-1'} \neq A_0^{-1} \Lambda_0^{-1} \Psi_0 B_0 \Sigma B_0' \Psi_0' \Lambda_0^{-1'} A_0^{-1'}$$

peut être supprimée dans **(H11)** mais peut être importante pour les formes SVARMA structurelles (1.28). En particulier, la condition **(H11)** est satisfaite lorsque l'on impose que  $A_0 = B_0 = \Lambda_0 = \Psi_0 = I_d$ , que **(H8)** est vérifié, que les facteurs commun de gauche de  $A_{\theta}(L)$ ,  $\Lambda_{\theta}(L^s)$ ,  $B_{\theta}(L)$  et  $\Psi_{\theta}(L^s)$  sont unimodulaires (i.e. avec un déterminant constant non nul), et enfin que la matrice  $[A_p : \Lambda_{p_s} : B_q : \Psi_{q_s}]$  est de plein rang (voir par exemple [40] qui a donné plusieurs procédures assurant l'identifiabilité).

**(H12) :** Nous avons  $\theta_0 \in \overset{\circ}{\Theta}$ , où  $\overset{\circ}{\Theta}$  est l'intérieure  $\Theta$ .

Nous utilisons la matrice  $M_{\theta_0}$  des coefficients de la forme réduite (1.29) où

$$\begin{aligned} M_{\theta_0} = & [A_{00}^{-1}A_{01} : \dots : A_{00}^{-1}A_{0p} : A_{00}^{-1}\Lambda_{00}^{-1}\Lambda_{01}A_{00} : \dots : A_{00}^{-1}\Lambda_{00}^{-1}\Lambda_{0p_s}A_{00} : A_{00}^{-1}\Lambda_{00}^{-1}\Lambda_{01}A_{01} : \\ & \dots : A_{00}^{-1}\Lambda_{00}^{-1}\Lambda_{01}A_{0p} : \dots : A_{00}^{-1}\Lambda_{00}^{-1}\Lambda_{0p_s}A_{01} : \dots : A_{00}^{-1}\Lambda_{00}^{-1}\Lambda_{0p_s}A_{0p} : \\ & A_{00}^{-1}\Lambda_{00}^{-1}\Psi_{00}B_{01}B_{00}^{-1}\Psi_{00}^{-1}\Lambda_{00}A_{00} : \dots : A_{00}^{-1}\Lambda_{00}^{-1}\Psi_{00}B_{0q}B_{00}^{-1}\Psi_{00}^{-1}\Lambda_{00}A_{00} : \\ & A_{00}^{-1}\Lambda_{00}^{-1}\Psi_{01}\Psi_{00}^{-1}\Lambda_{00}A_{00} : \dots : A_{00}^{-1}\Lambda_{00}^{-1}\Psi_{0q_s}\Psi_{00}^{-1}\Lambda_{00}A_{00} : \\ & A_{00}^{-1}\Lambda_{00}^{-1}\Psi_{01}B_{01}B_{00}^{-1}\Psi_{00}^{-1}\Lambda_{00}A_{00} : \dots : A_{00}^{-1}\Lambda_{00}^{-1}\Psi_{01}B_{0q}B_{00}^{-1}\Psi_{00}^{-1}\Lambda_{00}A_{00} : \dots : \\ & A_{00}^{-1}\Lambda_{00}^{-1}\Psi_{0q_s}B_{01}B_{00}^{-1}\Psi_{00}^{-1}\Lambda_{00}A_{00} : \dots : A_{00}^{-1}\Lambda_{00}^{-1}\Psi_{0q_s}B_{0q}B_{00}^{-1}\Psi_{00}^{-1}\Lambda_{00}A_{00}]. \end{aligned}$$

Maintenant, nous avons besoin d'une hypothèse d'identifiabilité locale qui complète (H11) et précise comment cette matrice  $M_{\theta_0}$  dépend du paramètre  $\theta_0$ . Nous notons par  $A \otimes B$  le produit Kronecker de deux matrices  $A$  et  $B$ , et par  $\text{vec}(A)$  le vecteur obtenu par l'empilement des colonnes de  $A$ . Soit  $\dot{M}_{\theta_0}$  la matrice  $\partial \text{vec}(M_{\theta}) / \partial \theta'$  évaluée en  $\theta_0$ .

**(H13) :** La matrice  $\dot{M}_{\theta_0}$  est de plein rang  $k_0$ .

**Proposition 1.4.1** *Sous les hypothèses (H6)–(H11), soit  $(X_t)$  une solution causal SVARMA de l'équation (1.28) et  $\hat{\theta}_n$  un QMLE. Alors  $\hat{\theta}_n \rightarrow \theta_0$  presque sûrement quand  $n \rightarrow +\infty$ . Si en plus les hypothèses (H4), (H12) et (H13) sont vérifiées, nous avons*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathcal{N}(0, \Omega := J^{-1}IJ^{-1}),$$

où  $J = J(\theta_0, \Sigma_{e0})$  et  $I = I(\theta_0, \Sigma_{e0})$ , avec

$$J(\theta, \Sigma_e) = \lim_{n \rightarrow \infty} \frac{-2}{n} \frac{\partial^2}{\partial \theta \partial \theta'} \log \tilde{L}_n(\theta, \Sigma_e) \quad \text{presque sûrement et } I(\theta, \Sigma_e) = \lim_{n \rightarrow \infty} \text{Var} \frac{2}{\sqrt{n}} \frac{\partial}{\partial \theta} \log \tilde{L}_n(\theta, \Sigma_e).$$

**Remarque 1.4.2** *Dans le cas standard de modèles SVARMA forts, i.e. quand l'hypothèse (H10) est remplacée par celle où les termes d'erreurs  $(\epsilon_t)$  sont iid, nous avons  $I = 2J$ , ainsi  $\Omega = 2J^{-1}$ . Nous retrouvons également le résultat obtenu par [81] (voir Lemme 1, p. 2048) pour le modèles SVARMA forts.*

### 1.4.3 La distribution asymptotique des autocovariances et autocorrélations résiduelles

Pour tout  $t \geq 1$ , soient  $\hat{e}_t := \hat{e}_t = \tilde{e}_t(\hat{\theta}_n)$  les résidus du quasi-maximum vraisemblance quand  $p > 0$ ,  $p_s > 0$ ,  $q > 0$  où  $q_s > 0$ . Nous utilisons (1.30) pour remarquer que nous avons  $\hat{e}_t = 0$  pour  $t \leq 0$  et  $t > n$ . A partir de (1.29), nous obtenons que

$$\begin{aligned} \hat{e}_t = & \hat{X}_t - \sum_{i=1}^p \hat{A}_0^{-1} \hat{A}_i \hat{X}_{t-i} - \sum_{j=1}^{p_s} \hat{A}_0^{-1} \hat{\Lambda}_0^{-1} \hat{\Lambda}_j \hat{A}_0 \hat{X}_{t-sj} + \sum_{i=1}^p \sum_{j=1}^{p_s} \hat{A}_0^{-1} \hat{\Lambda}_0^{-1} \hat{\Lambda}_j \hat{A}_i \hat{X}_{t-i-sj} \\ & + \sum_{i=1}^q \hat{A}_0^{-1} \hat{\Lambda}_0^{-1} \hat{\Psi}_0 \hat{B}_i \hat{B}_0^{-1} \hat{\Psi}_0^{-1} \hat{\Lambda}_0 \hat{A}_0 \hat{e}_{t-i} + \sum_{j=1}^{q_s} \hat{A}_0^{-1} \hat{\Lambda}_0^{-1} \hat{\Psi}_j \hat{\Psi}_0^{-1} \hat{\Lambda}_0 \hat{A}_0 \hat{e}_{t-sj} \\ & - \sum_{i=1}^q \sum_{j=1}^{q_s} \hat{A}_0^{-1} \hat{\Lambda}_0^{-1} \hat{\Psi}_j \hat{B}_i \hat{B}_0^{-1} \hat{\Psi}_0^{-1} \hat{\Lambda}_0 \hat{A}_0 \hat{e}_{t-i-sj}, \end{aligned}$$

pour  $t = 1, \dots, n$ , avec  $\hat{X}_t = 0$  pour  $t \leq 0$  et  $\hat{X}_t = X_t$  pour  $t \geq 1$ . Sous les conditions citées dans [57] (voir p. 301), nous notons par

$$\Upsilon(sh) = \frac{1}{n} \sum_{t=1+sh}^n e_t e'_{t-sh} \quad \text{et} \quad \hat{\Gamma}_e(sh) = \frac{1}{n} \sum_{t=1+sh}^n \hat{e}_t \hat{e}'_{t-sh}$$

les autocovariances empiriques du bruit blanc et les autocovariances résiduelles multiples de la saison  $s$ . Pour un entier  $m \geq 1$  fixé, nous considérons

$$\Upsilon_m(s) = \left( \{\text{vec}\Upsilon(1s)\}' , \dots , \{\text{vec}\Upsilon(ms)\}' \right)' \text{ et } \hat{\Gamma}_m(s) = \left( \{\text{vec}\hat{\Gamma}_e(1s)\}' , \dots , \{\text{vec}\hat{\Gamma}_e(ms)\}' \right)' .$$

Soient les matrices diagonales

$$S_e = \text{Diag}(\sigma_e(1), \dots, \sigma_e(d)) \quad \text{et} \quad \hat{S}_e = \text{Diag}(\hat{\sigma}_e(1), \dots, \hat{\sigma}_e(d)) ,$$

où  $\sigma_e^2(i)$  est la variance de la  $i$ -ème coordonnées de  $e_t$  et  $\hat{\sigma}_e^2(i)$  est son estimateur (*i.e.*  $\sigma_e^2(i) = \mathbb{E}e_{it}^2$  et  $\hat{\sigma}_e^2(i) = n^{-1} \sum_{t=1}^n \hat{e}_{it}^2$ ). Les autocorrélations théoriques et résiduelles du retard  $\ell$  sont définies respectivement par  $R_e(\ell) = S_e^{-1} \Gamma_e(\ell) S_e^{-1}$  et  $\hat{R}_e(\ell) = \hat{S}_e^{-1} \hat{\Gamma}_e(\ell) \hat{S}_e^{-1}$ , avec  $\Gamma_e(\ell) := \mathbb{E}e_t e_{t-\ell}' = 0$  pour tout  $\ell \neq 0$ . Nous considérons des vecteurs des  $m$  (pour  $m \geq 1$ ) premières autocorrélations résiduelles

$$\hat{\rho}_m(s) = \left( \{\text{vec}\hat{R}_e(1s)\}' , \dots , \{\text{vec}\hat{R}_e(ms)\}' \right)' .$$

[26] et [48] ont étendu les tests portmanteau univariés du modèle ARMA fort de [24, 53], au cas des modèles ARMA multivariés forts. Les statistiques de test multivariées introduites sont définies par :

$$Q_m^C = n \sum_{h=1}^m \text{Tr}(\hat{\Gamma}'(h) \hat{\Gamma}^{-1}(0) \hat{\Gamma}(h) \hat{\Gamma}^{-1}(0)) \text{ et } Q_m^H = \sum_{h=1}^m \frac{n^2}{(n-h)} \text{Tr}(\hat{\Gamma}'(h) \hat{\Gamma}^{-1}(0) \hat{\Gamma}(h) \hat{\Gamma}^{-1}(0)) , \quad (1.31)$$

où  $n$  est la taille de la série,  $m$  est un entier fixé et  $\hat{\Gamma}(h)$  est la fonction des matrices d'autocovariances résiduelles du processus multivarié  $\hat{e}$ , on note  $\text{Tr}(\cdot)$  la trace de la matrice  $A$  et  $A'$  la transposée de la matrice  $A$ . [46, 47] a donné plusieurs formes équivalentes des statistiques (1.31). Ces dernières (1.31) sont généralement utilisées pour vérifier la signification globale des autocorrélations résiduelles d'un modèle VARMA (voir [1, 52]).

Ces statistiques sont souvent utilisées pour tester l'hypothèse nulle

( $\mathcal{H}0$ ) :  $(X_t)_{t \in \mathbb{Z}}$  satisfait une représentation SVARMA( $p, q$ )( $p_s, q_s$ );

contre l'alternative

( $\mathcal{H}1$ ) :  $(X_t)_{t \in \mathbb{Z}}$  n'admet pas de représentation SVARMA( $p, q$ )( $p_s, q_s$ ) où  $(X_t)_{t \in \mathbb{Z}}$  satisfait une représentation SVARMA( $p, q$ )( $p'_s, q'_s$ ) avec  $p'_s > p_s$  ou  $q'_s > q_s$ .

Récemment, [56] a introduit les versions suivantes des statistiques de portmanteau pour les modèles SARMA univariés forts

$$Q_m^{\text{BP}}(s) = n \sum_{h=1}^m \hat{\rho}^2(sh) \text{ et } Q_m^{\text{LB}}(s) = n(n+2) \sum_{h=1}^m \frac{\hat{\rho}^2(sh)}{n-sh} . \quad (1.32)$$

En nous inspirant de (1.31) et des statistiques univariées de portmanteau (1.32), nous avons introduit les versions suivantes des statistiques de portmanteau.

$$\begin{aligned} Q_m^C(s) &= n \sum_{h=1}^m \text{Tr}(\hat{\Gamma}'(sh) \hat{\Gamma}^{-1}(0) \hat{\Gamma}(sh) \hat{\Gamma}^{-1}(0)) \text{ et} \\ Q_m^H(s) &= \sum_{h=1}^m \frac{n^2}{(n-sh)} \text{Tr}(\hat{\Gamma}'(sh) \hat{\Gamma}^{-1}(0) \hat{\Gamma}(sh) \hat{\Gamma}^{-1}(0)) . \end{aligned} \quad (1.33)$$

Cette modification est justifiée par les conditions indiquées dans [57] (voir p. 301). La distribution limite des tests qui découle de (1.33) est obtenue par une extension de ceux obtenus dans [24, 26, 48, 53].

**Proposition 1.4.2** *Nous supposons que la représentation (1.28) vérifie un modèle SVARMA fort. Sous les hypothèses indiquées par [57] (voir p. 301), à savoir : si  $p \ll s$  et  $q \ll s$  et les racines de  $\det A_\theta(z) \det B_\theta(z) = 0$  sont à l'extérieur du disque unité et  $\mathcal{H}0$ , les statistiques définies dans (1.33) suivent une distribution chi-deux à  $d^2(m - p_s - q_s) > 0$  degré de liberté pour les  $n$  grands et  $ms \leq (n - 1)$ .*

Compte tenue de la proposition 1.4.2, la procédure de test consiste à rejeter l'hypothèse nulle d'un modèle SVARMA( $p, q$ )( $p_s, q_s$ ) si les statistiques (1.33) sont supérieures à un  $\chi^2_{d^2(m-p_s-q_s)}$ . Par conséquent, ces tests ne sont pas non plus applicables pour  $m \leq p_s + q_s$ . A partir de la distribution jointe du vecteur  $\sqrt{n}(\hat{\theta}'_n - \theta'_0, \gamma'_m(s))$ , nous en déduisons la distribution asymptotique de  $\sqrt{n}\hat{\rho}_m(s)$ . Les hypothèses de mélange (H4) impliquent la normalité asymptotique des autocovariances empiriques  $\gamma_m(s)$ . Soit

$$\Gamma(s\ell, s\ell') = \sum_{h=-\infty}^{\infty} \mathbb{E}(\{e_{t-s\ell} \otimes e_t\} \{e_{t-h-s\ell'} \otimes e_{t-h}\}'), \quad \text{pour } (\ell, \ell') \neq (0, 0).$$

Nous sommes maintenant en mesure d'énoncer le théorème suivant, qui est une extension d'un résultat donné dans [15].

**Theorem 1.4.3** *Supposons que  $p > 0$  et  $p_s > 0$  ou  $q > 0$  et  $q_s > 0$ . Sous les hypothèses (H4) et (H6)–(H13), quand  $n \rightarrow +\infty$ ,*

$$\sqrt{n}(\hat{\theta}'_n - \theta'_0, \gamma'_m(s))' = \frac{1}{\sqrt{n}} \sum_{t=1}^n w_t + o_{\mathbb{P}}(1) \xrightarrow[n \rightarrow +\infty]{d} \mathcal{N}(0, \Xi),$$

où

$$\Xi = \begin{pmatrix} \Omega & \Sigma_{\hat{\theta}_n, \gamma_m(s)} \\ \Sigma'_{\hat{\theta}_n, \gamma_m(s)} & \Sigma_{\gamma_m(s)} \end{pmatrix} = \sum_{h=-\infty}^{+\infty} \mathbb{E} w_t w'_{t-h}, \quad \text{avec } w_t = \begin{pmatrix} w_{1t} \\ w_{2t} \end{pmatrix} \quad (1.34)$$

où  $w_{1t} = -2J^{-1}(\partial e'_t / \partial \theta) \Sigma_{e0}^{-1} e_t$  et  $w_{2t} = (\epsilon_{t-1s}, \dots, \epsilon_{t-ms})' \epsilon_t$  et où  $\Sigma_{\gamma_m(s)} = \{\Gamma(s\ell, s\ell')\}_{1 \leq \ell, \ell' \leq m}$ ,  $\Sigma'_{\gamma_m(s), \hat{\theta}_n} = \lim_{n \rightarrow \infty} \text{Cov}(n^{-1/2} \sum_{t=1}^n J^{-1} \Upsilon_t, \sqrt{n} \gamma_m(s))$  et  $\Omega = J^{-1} I J^{-1}$ . Les matrices  $I$  et  $J$  sont définies dans la proposition 1.4.1.

Définissons la matrice

$$\Phi_m(s) = \mathbb{E} \left\{ \begin{pmatrix} e_{t-1s} \\ \vdots \\ e_{t-ms} \end{pmatrix} \otimes \frac{\partial e_t}{\partial \theta'} \right\}. \quad (1.35)$$

Le théorème suivant donne le comportement asymptotique des autocovariances et autocorrélations résiduelles.

**Theorem 1.4.4** *Sous les hypothèses du théorème 1.4.3, nous avons*

$$\sqrt{n} \hat{\Gamma}_m(s) \xrightarrow[n \rightarrow +\infty]{d} \mathcal{N}(0, \Sigma_{\hat{\Gamma}_m(s)}) \quad \text{et} \quad \sqrt{n} \hat{\rho}_m(s) \xrightarrow[n \rightarrow +\infty]{d} \mathcal{N}(0, \Sigma_{\hat{\rho}_m(s)}) \quad \text{où,}$$

$$\Sigma_{\hat{\Gamma}_m(s)} = \Sigma_{\gamma_m(s)} + \Phi_m(s) \Omega \Phi'_m(s) + \Phi_m(s) \Sigma_{\hat{\theta}_n, \gamma_m(s)} + \Sigma'_{\hat{\theta}_n, \gamma_m(s)} \Phi'_m(s) \quad (1.36)$$

$$\Sigma_{\hat{\rho}_m(s)} = \{I_m \otimes (S_e \otimes S_e)^{-1}\} \Sigma_{\hat{\Gamma}_m(s)} \{I_m \otimes (S_e \otimes S_e)^{-1}\} \quad (1.37)$$

et  $\Phi_m(s)$  est donnée par (1.35).

**Remarque 1.4.3** *Quand  $s = 1$ , nous retrouvons le résultat obtenu par [15]. Dans le cas standard de modèles SVARMA forts, i.e. quand l'hypothèse (H10) est remplacée par celle où les termes d'erreurs ( $\epsilon_t$ ) sont iid et quand  $s = 1$ , nous avons  $\Sigma_{\gamma_m(1)} = I_m \otimes \Sigma_{e0} \otimes \Sigma_{e0}$ ,  $\Omega = 2J^{-1}$  (voir Remarque 1.4.2) et  $\Sigma'_{\hat{\theta}_n, \gamma_m(s)} = -\Phi_m(1)\Omega$ . Ainsi  $\Sigma_{\hat{\Gamma}_m(1)} = I_m \otimes \Sigma_{e0} \otimes \Sigma_{e0} - \Phi_m(1)\Omega\Phi'_m(1)$ , nous retrouvons le résultat obtenu par [81].*

### 1.4.4 Version modifiée de test portmanteau

Plusieurs formes équivalentes des statistiques (1.31) sont données par [48]. En utilisant les identités élémentaires  $\text{vec}(AB) = (I \otimes A) \text{vec} B$ ,  $(A \otimes B)(C \otimes D) = AC \otimes BD$  et  $\text{Tr}(ABC) = \text{vec}(A)'(C' \otimes I) \text{vec} B$ , on peut trouver des formes de base analogues pour les statistiques (1.33)

$$\begin{aligned} Q_m^C(s) &= n \sum_{h=1}^m \text{Tr}(\hat{\Gamma}'_e(sh) \hat{\Gamma}_e^{-1}(0) \hat{\Gamma}_e(sh) \hat{\Gamma}_e^{-1}(0)) \\ &= n \hat{\rho}'_m(s) (I_m \otimes \{\hat{R}_e^{-1}(0) \otimes \hat{R}_e^{-1}(0)\}) \hat{\rho}_m(s). \end{aligned}$$

Les mêmes formes équivalentes valent pour  $Q_m^H(s)$ . A partir du Théorème 1.4.4, nous en déduisons le résultat suivant, qui donne la distribution asymptotique exacte des statistiques standards de portmanteau (1.33), dans le cas de modèles SVARMA faibles. Nous verrons que cette distribution peut être très différente du  $\chi^2_{d^2(m-p_s-q_s)}$  obtenu dans le cas des modèles SVARMA forts.

**Theorem 1.4.5** *Sous les hypothèses du théorème 1.4.4 et  $(\mathcal{H}0)$ , les statistiques  $Q_m^C$  et  $Q_m^H$  introduites dans (1.33) convergent en distribution, quand  $n \rightarrow +\infty$ , vers*

$$Z_m(\xi_m) = \sum_{i=1}^{d^2 m} \xi_{i,d^2 m} Z_i^2$$

où  $\xi_m = (\xi_{1,d^2 m}, \dots, \xi_{d^2 m,d^2 m})'$  est le vecteur des valeurs propres de la matrice

$$\Omega_m = (I_m \otimes \Sigma_e^{-1/2} \otimes \Sigma_e^{-1/2}) \Sigma_{\hat{\Gamma}_m} (I_m \otimes \Sigma_e^{-1/2} \otimes \Sigma_e^{-1/2}),$$

et  $Z_1, \dots, Z_{d^2 m}$  sont des variables indépendantes  $\mathcal{N}(0, 1)$ .

Au vu du théorème 1.4.4, les matrices de variance asymptotique  $\Sigma_{\hat{\Gamma}_m(s)}$  et  $\Sigma_{\hat{\rho}_m(s)}$  dépendent des matrices inconnues  $\Xi$ ,  $\Phi_m(s)$  et de paramètre de nuisance  $\Sigma_e$ . Les matrices  $\Phi_m(s)$  et  $\Sigma_e$  peuvent être respectivement estimées empiriquement par

$$\hat{\Phi}_m(s) = \frac{1}{n} \sum_{t=1}^n \left\{ (\hat{e}_{t-1s}, \dots, \hat{e}_{t-ms})' \frac{\partial \hat{e}_t}{\partial \theta'} \right\} \text{ et } \hat{\Sigma}_e = \hat{\Gamma}_e(0) = \frac{1}{n} \sum_{t=1}^n \hat{e}_t \hat{e}_t'.$$

Dans la littérature économétrique, l'estimateur non paramétrique du noyau (voir [63], ou [5]), aussi appelé HAC (Heteroscedasticity and Autocorrelation Consistant) est largement utilisé pour estimer une matrices de covariance de la forme  $\Xi$ . La matrice  $\Xi/2\pi$  peut être interprétée comme la densité spectrale du processus stationnaire  $(w_t) = (w_{1t}', w_{2t}')'_{t \in \mathbb{Z}}$  évaluée à la fréquence 0 (voir [25], p. 459), une autre méthode consiste à utiliser une estimation de la densité spectrale de  $w_t$  par une autorégression. Cette approche, qui a été étudiée par [10] (voir aussi [28]), repose sur l'expression

$$\Xi = \Phi^{-1}(1) \Sigma_u \Phi'^{-1}(1)$$

lorsque  $(w_t)$  satisfait une représentation  $\text{AR}(\infty)$  de la forme paramétrique

$$\Phi(L) w_t = w_t + \sum_{i=1}^{\infty} \Phi_i w_{t-i} = u_t, \quad (1.38)$$

où  $u_t$  est un bruit blanc faible de variance  $\Sigma_u$ . Puisque  $w_t$  est non observable, nous posons  $\hat{w}_t$  le vecteur obtenu en remplaçant  $\theta_0$  par  $\hat{\theta}_n$  dans l'expression  $w_t$ . Nous définissons le polynôme  $\hat{\Phi}_r(z) = I_{k_0+m} + \sum_{i=1}^r \hat{\Phi}_{r,i} z^i$ , où  $\hat{\Phi}_{r,1}, \dots, \hat{\Phi}_{r,r}$  sont les coefficients de la régression des moindres carrés de  $\hat{w}_t$  sur  $\hat{w}_{t-1}, \dots, \hat{w}_{t-r}$ . Notons  $\hat{u}_{r,t}$  les résidus de cette régression, et soit  $\hat{\Sigma}_{\hat{u}_r}$  la variance empirique de  $\hat{u}_{r,1}, \dots, \hat{u}_{r,n}$ . Nous énonçons le théorème suivant qui est une extension du résultat de [33].



**Theorem 1.4.6** *Sous les hypothèses du théorème 1.4.5, nous supposons que le processus  $(w_t)$  défini dans (1.38) admet une représentation AR( $\infty$ ) dont les racines de  $\det \Phi(z) = 0$  sont à l'extérieur du disque unité,  $\|\Phi_i\| = o(i^{-2})$ , et que la matrice  $\Sigma_u = \text{Var}(u_t)$  est non-singulière. De plus, nous supposons que  $\mathbb{E}|\epsilon_t|^{8+4\nu} < \infty$  et  $\sum_{k=0}^{\infty} \{\alpha_\epsilon(k)\}^{\nu/(2+\nu)} < \infty$  pour un réel  $\nu > 0$ . Alors l'estimateur spectral de  $\Xi$*

$$\hat{\Xi}^{\text{SP}} = \hat{\Phi}_r^{-1}(1) \hat{\Sigma}_{\hat{u}_r} \hat{\Phi}_r'^{-1}(1) \rightarrow \Xi$$

en probabilité quand  $r = r(n) \rightarrow +\infty$  et  $r^3/n \rightarrow 0$  quand  $n \rightarrow +\infty$ .

Soit  $\hat{\Omega}_m$  la matrice obtenue en remplaçant  $\Xi$  par  $\hat{\Xi}$  et  $\Sigma_e$  par  $\hat{\Sigma}_e$  dans  $\Omega_m$ . Définissons  $\hat{\xi}_m = (\hat{\xi}_{1,d^2m}, \dots, \hat{\xi}_{d^2m,d^2m})'$  le vecteur des valeurs propres de  $\hat{\Omega}_m$ . Pour un niveau asymptotique de risque  $\alpha$ , le test consiste à rejeter l'adéquation de modèles SVARMA faibles quand

$$Q_m^{\text{C}}(s) > S_m(1 - \alpha) \quad (\text{resp.} \quad Q_m^{\text{H}}(s) > S_m(1 - \alpha))$$

où  $S_m(1 - \alpha)$  est tel que  $\mathbb{P}\{Z_m(\hat{\xi}_m) > S_m(1 - \alpha)\} = \alpha$ .

## 1.5 Résultats du chapitre 4

Dans cet chapitre nous rappelons dans un premier temps, les résultats sur les estimateurs de moindres carrés ordinaires, généralisés et quasi-généralisés des paramètres des modèles PARMA obtenus par [32]. Dans un second temps, nous étudions le comportement asymptotique du test portmanteau dans le cadre du modèle PARMA dont les termes d'erreur sont non corrélés mais peuvent contenir de dépendances non linéaires.

L'étude de la distribution asymptotique jointe de l'estimateur des moindres carrés quasi-généralisés et des autocovariances empiriques du bruit périodique, nous permet d'obtenir les distributions asymptotiques des autocovariances et autocorrélations résiduelles périodiques pour un  $\nu$  donnée. Ceci nous permet d'avoir par la suite les distributions asymptotiques des autocovariances et autocorrélations résiduelles globales. Nous déduisons le comportement asymptotique des statistiques portmanteau pour un  $\nu$  donnée et le comportement asymptotique des statistiques portmanteau globales. De la même manière que le chapitre 2, une deuxième méthode alternative basée sur une approche d'auto-normalisation est proposée dans ce chapitre pour tester l'adéquation des modèles PARMA faibles.

### 1.5.1 Propriétés asymptotiques des estimateurs

Soient  $\phi(\nu) = (\phi_1(\nu), \dots, \phi_p(\nu))'$  et  $\theta(\nu) = (\theta_1(\nu), \dots, \theta_q(\nu))'$  respectivement les vecteurs des paramètres autorégressifs et des moyennes mobiles pour une saison donnée  $\nu$ . L'ensemble de dimension  $T(p + q)$  de tous les paramètres PARMA est noté par

$$\alpha := (\phi(1)', \dots, \phi(T)', \theta(1)', \dots, \theta(T))'$$

dont la dimension de  $\alpha$  est  $(p + q)T \times 1$ . La variance  $\sigma^2 = (\sigma^2(1), \dots, \sigma^2(T))'$  périodiques du bruit blanc seront considérée comme un paramètre de nuisance.

Le paramètre inconnu  $\alpha_0$  appartient à l'espace des paramètres

$$\Delta = \left\{ \alpha = (\phi(1)', \dots, \phi(T)', \theta(1)', \dots, \theta(T))' \text{ tel que } (\mathbf{H1}'') \text{ est vérifiée} \right\}.$$

Sans perte de généralité, nous supposons que  $\Delta$  est compact. Pour  $\alpha \in \Delta$ , soit  $\epsilon_{nT+\nu}(\alpha)$  la solution stationnaire périodique de seconde ordre de

$$\epsilon_{nT+\nu}(\alpha) = X_{nT+\nu} - \sum_{k=1}^p \phi_k(\nu) X_{nT+\nu-k} + \sum_{l=1}^q \theta_l(\nu) \epsilon_{nT+\nu-l}(\alpha). \quad (1.39)$$



Les variables  $\epsilon_{nT+v}(\alpha)$  peuvent être approximées, pour  $0 < nT+v \leq NT$  par  $e_{nT+v}(\alpha)$  définies récursivement par

$$e_{nT+v}(\alpha) = X_{nT+v} - \sum_{k=1}^p \phi_k(v)X_{nT+v-k} + \sum_{l=1}^q \theta_l(v)e_{nT+v-l}(\alpha), \quad (1.40)$$

où les valeurs initiales inconnues sont mises à zero :  $e_0(\alpha) = \dots = e_{1-q}(\alpha) = X_0 = \dots = X_{1-p} = 0$ .

Comme le montre [32, Lemme 7], ces valeurs initiales sont asymptotiquement négligeables uniformément en  $\alpha$ .

La variable aléatoire  $\hat{\alpha}_{OLS}$  est appelée l'estimateur de moindres carrés ordinaires (OLS pour Ordinary Least Squares en anglais) de  $\alpha$  si elle satisfait presque sûrement,

$$S_N(\hat{\alpha}_{OLS}) = \min_{\alpha \in \Delta} S_N(\alpha) \quad \text{où} \quad S_N(\alpha) = \frac{1}{2N} \sum_{n=0}^{N-1} \sum_{v=1}^T e_{nT+v}^2(\alpha).$$

En raison de la présence d'innovations hétéroscédastiques, l'estimateur OLS pourrait être inefficace. Étant donné certains vecteurs de poids,  $\omega^2 = (\omega_1^2, \dots, \omega_T^2)'$ , [32] (voir également [6]) ont montré que l'estimateur OLS est inefficace par rapport à l'estimateur des moindres carrés pondérés (WLS pour Weighted LS)  $\hat{\alpha}_{WLS} = \hat{\alpha}_{WLS}^{\omega^2}$  défini par

$$Q_N^{\omega^2}(\hat{\alpha}_{WLS}) = \min_{\alpha \in \Delta} Q_N^{\omega^2}(\alpha) \quad \text{où} \quad Q_N^{\omega^2}(\alpha) = \frac{1}{2N} \sum_{n=0}^{N-1} \sum_{v=1}^T \omega_v^{-2} e_{nT+v}^2(\alpha).$$

[32] et [6] ont montré qu'un estimateur WLS optimal est l'estimateur des moindres carrés généralisés (GLS pour Generalized LS) défini par

$$\hat{\alpha}_{GLS} = \hat{\alpha}_{WLS}^{\sigma^2}, \quad \text{où} \quad \sigma^2 = (\sigma_1^2, \dots, \sigma_T^2)'. \quad (1.41)$$

Comme mentionné par [32], l'estimateur GLS suppose que  $\sigma^2$  est connu. En pratique, ce paramètre doit également être estimé. Pour tout estimateur consistant  $\hat{\sigma}^2 = (\hat{\sigma}_1^2, \dots, \hat{\sigma}_T^2)'$  de  $\sigma^2$ , un estimateur des moindres carrés quasi-généralisé (QLS pour Quasi-generalized LS) de  $\alpha_0$  est défini par

$$\hat{\alpha}_{QLS} = \hat{\alpha}_{WLS}^{\hat{\sigma}^2}. \quad (1.42)$$

Un estimateur consistant possible de  $\sigma_v^2$  peut être estimé par

$$\hat{\sigma}_v^2 = \frac{1}{N} \sum_{n=0}^{N-1} e_{nT+v}^2(\hat{\alpha}_{OLS}).$$

Pour établir la consistance des estimateurs des moindres carrés (LSE), une hypothèse supplémentaire est nécessaire.

**(H2'')** : Le bruit blanc  $\{\epsilon_n = (\epsilon_{nT+1}, \dots, \epsilon_{nT+v}, \dots, \epsilon_{nT+T})', n \in \mathbb{Z}\}$  de dimension T est strictement stationnaire et ergodique.

Pour la normalité asymptotique de LSE, nous avons besoin des hypothèses supplémentaires. Il est d'abord nécessaire de supposer que le paramètre  $\alpha_0$  n'est pas situé sur le bord de l'espace de paramètre  $\Delta$ .

**(H3'')** : Nous avons  $\alpha_0 \in \overset{\circ}{\Delta}$ , où  $\overset{\circ}{\Delta}$  est à l'intérieur  $\Delta$ .

Pour contrôler la dépendance dans le temps du processus stationnaire  $(\epsilon_n)$ , nous introduisons les coefficients de mélange fort  $\alpha_\epsilon(h)$  définis par

$$\alpha_\epsilon(h) = \sup_{A \in \mathcal{F}_{-\infty}^n, B \in \mathcal{F}_{n+h}^{+\infty}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|,$$

où  $\mathcal{F}_{-\infty}^n = \sigma(\epsilon_u, u \leq n)$  et  $\mathcal{F}_{n+h}^{+\infty} = \sigma(\epsilon_u, u \geq n+h)$ .

Nous utilisons  $\|\cdot\|$  pour noter la norme Euclidienne du vecteur. Nous ferons une hypothèse d'intégrabilité sur le moment du bruit et une condition de sommabilité sur les coefficients de mélange fort  $(\alpha_\epsilon(k))_{k \geq 0}$ .

**(H4'')** : Nous avons  $\mathbb{E}\|\epsilon_n\|^{4+2\kappa} < \infty$  et  $\sum_{k=0}^{\infty} \{\alpha_\epsilon(k)\}^{\frac{\kappa}{2+\kappa}} < \infty$  pour certaines  $\kappa > 0$ .

Sous les hypothèses ci-dessus [32] ont montré que

$\hat{\alpha}_{\text{OLS}} \rightarrow \alpha_0$ ,  $\hat{\alpha}_{\text{WLS}} \rightarrow \alpha_0$ ,  $\hat{\alpha}_{\text{GLS}} \rightarrow \alpha_0$  et  $\hat{\alpha}_{\text{QLS}} \rightarrow \alpha_0$  quand  $N \rightarrow +\infty$  et

$$\sqrt{N}(\hat{\alpha}_{\text{LS}} - \alpha_0) \xrightarrow[N \rightarrow +\infty]{d} \mathcal{N}(0, \Omega^{\text{LS}}), \quad (1.43)$$

où l'exposant LS signifie OLS, WLS, GLS et QLS, et où

$$\Omega^{\text{OLS}} = \Omega(\alpha_0, (1, \dots, 1)'), \quad \Omega^{\text{WLS}} = \Omega(\alpha_0, \omega^2), \quad \Omega^{\text{GLS}} = \Omega(\alpha_0, \sigma^2), \quad \Omega^{\text{QLS}} = \Omega(\alpha_0, \hat{\sigma}^2),$$

avec

$$\Omega(\alpha_0, \omega^2) = (J(\alpha_0, \omega^2))^{-1} I(\alpha_0, \omega^2) (J(\alpha_0, \omega^2))^{-1}.$$

Les matrices  $I(\alpha_0, \omega^2)$  et  $J(\alpha_0, \omega^2)$  sont définies comme suit :

$$J(\alpha_0, \omega^2) = \sum_{v=1}^T \omega_v^{-2} \mathbb{E} \left[ \left( \frac{\partial \epsilon_{nT+v}(\alpha)}{\partial \alpha} \right) \left( \frac{\partial \epsilon_{nT+v}(\alpha)}{\partial \alpha} \right)' \right]_{\alpha=\alpha_0}, \quad (1.44)$$

$$I(\alpha_0, \omega^2) = \sum_{v=1}^T \sum_{v'=1}^T \omega_v^{-2} \omega_{v'}^{-2} \sum_{k=-\infty}^{\infty} \mathbb{E} \left[ \epsilon_v(\alpha) \epsilon_{kT+v'}(\alpha) \left( \frac{\partial \epsilon_v(\alpha)}{\partial \alpha} \right) \left( \frac{\partial \epsilon_{kT+v'}(\alpha)}{\partial \alpha} \right)' \right]_{\alpha=\alpha_0}. \quad (1.45)$$

## 1.5.2 La distribution asymptotique des autocovariances et autocorrélations résiduelles

Soient  $\hat{\epsilon}_{nT+v} = \hat{e}_{nT+v} = e_{nT+v}(\hat{\alpha}_{\text{LS}})$  les résidus périodiques de moindres carrés quand  $p+q > 0$ . Nous utilisons (1.40) pour remarquer que nous avons  $\hat{e}_t = 0$  pour  $t := nT+v \leq 0$  et  $t > NT$ . De (1.6), pour  $n = 0, 1, \dots, N-1$ , nous déduisons que

$$\hat{e}_{nT+v} = X_{nT+v} - \sum_{k=1}^p \hat{\phi}_k(v) X_{nT+v-k} + \sum_{l=1}^q \hat{\theta}_l(v) \hat{e}_{nT+v-l},$$

pour  $1 \leq t \leq NT$ , avec  $\hat{X}_t = 0$  pour  $t \leq 0$  et  $\hat{X}_t = X_t$  pour  $t \geq 1$ . Nous notons par

$$\gamma(v, l) = \frac{1}{N} \sum_{n=l+1}^N \epsilon_{nT+v} \epsilon_{nT+v-l} \text{ et } \hat{\gamma}(v, l) = \frac{1}{N} \sum_{n=l+1}^N \hat{e}_{nT+v} \hat{e}_{nT+v-l}, \quad (1.46)$$

les autocovariances empiriques du bruit blanc périodique et les autocovariances périodiques résiduelles. Il est à noter que  $\gamma(v, l)$  ne sont pas de statistiques puisqu'elles dépendent des innovations théoriques  $\epsilon_{nT+v} = \epsilon_{nT+v}(\alpha_0)$ .

Pour un entier  $m \geq 1$  fixé, soient

$$\gamma_m(v) := (\gamma(v, 1), \dots, \gamma(v, m))' \text{ et } \hat{\gamma}_m(v) := (\hat{\gamma}(v, 1), \dots, \hat{\gamma}(v, m))',$$

les vecteurs des autocovariances empiriques et autocovariances résiduelles pour un  $v$  donné.

De la même manière, les autocovariances empiriques globales et les autocovariances résiduelles sur toutes les saisons sont données par

$$\Upsilon_{mT} := (\Upsilon_m(1), \dots, \Upsilon_m(T))' \text{ et } \hat{\Upsilon}_{mT} := (\hat{\Upsilon}_m(1), \dots, \hat{\Upsilon}_m(T))'$$

Notons tout d'abord que l'hypothèse de mélange (**H4''**) impliquera la normalité asymptotique de  $\Upsilon_m(v)$  et  $\Upsilon_m(T)$ . Les autocorrélations théoriques et résiduelles de retard  $l$  sont respectivement définis par

$$\rho(v, l) = \frac{\Gamma(v, l)}{\sqrt{\Gamma(v, 0)\Gamma(v-l, 0)}} \text{ et } \hat{\rho}(v, l) = \frac{\hat{\Upsilon}(v, l)}{\sqrt{\hat{\Upsilon}(v, 0)\hat{\Upsilon}(v-l, 0)}} \text{ avec } \hat{\Upsilon}(v-l, 0) = \frac{1}{N} \sum_{n=1}^N \hat{e}_{nT+v-l}^2,$$

et où  $\Gamma(v, l) := \text{Cov}(\epsilon_{nT+v}, \epsilon_{nT+v-l})$  avec  $\Gamma(v, 0) := \sigma_v^2$ . Voir par exemple [72] pour certaines propriétés de  $\Gamma(v, l)$ .

Dans la suite, nous aurons également besoin de ces expression :

$$\hat{\rho}_m(v) = (\hat{\rho}(v, 1), \dots, \hat{\rho}(v, m))' \text{ et } \hat{\rho}_{mT} = (\hat{\rho}_m(1), \dots, \hat{\rho}_m(T))',$$

où  $\hat{\rho}_m(v)$  représente le vecteur de  $m$  premiers autocorrélations résiduelles pour un  $v$  donné et  $\hat{\rho}_m(T)$  est le vecteur global de autocorrélations résiduelles de toutes les saisons.

Sur la base des autocorrélations empiriques résiduelles, [24] ont proposé un test pour mesurer la qualité d'ajustement, le test dit portmanteau pour les modèles ARMA forts. Une modification de leur test a été proposée par [53]. Pour tester simultanément si toutes les autocorrélations résiduelles aux retards  $1, \dots, m$  d'un modèle PARMA sont égaux à zéros pour une période  $v$  donnée, le test portmanteau de [24, 53] peut être adapté comme proposé par [59, 60] (voir également [43]). Sur la base de ces résultats, [59, 60] a suggéré les statistiques suivantes définies par

$$Q_m^{\text{BP}}(v) = N \sum_{h=1}^m \hat{\rho}^2(v, h) \text{ et } Q_m^{\text{LBM}}(v) = N \sum_{h=1}^m \varrho(h, v, N, T) \hat{\rho}^2(v, h), \quad (1.47)$$

où  $\varrho(h, v, N, T)$  est appelé la correction de Ljung-Box-McLeod (LBM) et est définie par

$$\varrho(h, v, N, T) = \begin{cases} (N+2)/(N-h/T) & \text{si } h \equiv 0 \pmod{T} \\ N/(N - \lfloor (h-v+T)/T \rfloor) & \text{sinon.} \end{cases}$$

La statistique  $Q_m^{\text{LBM}}(v)$  a la même distribution chi-deux que la statistique  $Q_m^{\text{BP}}(v)$  et est adaptée pour les échantillons de petites et moyennes taille (voir [59, 60]). Si le modèle  $\text{PARMA}_T(p; q)$  fort est vérifié pour un  $v$  donné,  $Q_m^{\text{BP}}(v)$  ou  $Q_m^{\text{LBM}}(v)$  sera distribué asymptotiquement comme une variable chi-deux avec  $m - (p + q)$  (et  $m - (p_v + q_v)$  degrés de liberté pour le PARMA fort (1.7)) degrés de liberté pour le modèle (1.6).

Il convient de noter que le retard  $m$  utilisé dans (1.47) pourrait être choisi pour être différent selon les saisons mais dans la plupart des applications, il est raisonnable d'utiliser la même valeur de  $m$  pour toutes les saisons (voir par exemple [43]). Entre les saisons, les statistiques du test portmanteau sont asymptotiquement indépendant pour  $v = 1, \dots, T$ . Par conséquent, dans le cas des statistiques du test portmanteau dans (1.47), un test global pour vérifier si les résidus de toutes les saisons sont des estimations de bruits blancs est donné par

$$Q_m^{\text{BP}} = \sum_{v=1}^T Q_m^{\text{BP}}(v) \text{ et } Q_m^{\text{LBM}} = \sum_{v=1}^T Q_m^{\text{LBM}}(v). \quad (1.48)$$

Ainsi, si le modèle  $\text{PARMA}_T(p; q)$  fort est vérifié, les statistiques globales (1.48) seront également distribuées de manière asymptotique comme une variable chi-deux avec  $mT - T(p + q)$  (et  $mT - \sum_{v=1}^T (p_v + q_v)$  pour le modèle (1.7) fort) degrés de liberté (voir [43]).

Les statistiques (1.48) (resp. (1.47)) sont généralement utilisées pour tester l'hypothèse nulle suivante

( $\mathcal{H}0$ ) :  $(X_{nT+v})$  satisfait une représentation  $\text{PARMA}_T(p; q)$  (resp. pour une période  $v$  donnée);  
contre l'alternative

( $\mathcal{H}1$ ) :  $(X_{nT+v})$  n'admet pas une représentation  $\text{PARMA}_T(p; q)$  (resp. pour une période  $v$  donnée)  
ou  $(X_{nT+v})$  satisfait une représentation  $\text{PARMA}_T(p'; q')$  (resp. pour une période  $v$  donnée)  
avec  $p' > p$  ou  $q' > q$ .

Pour les modèles  $\text{PARMA}_T(p; q)$  faibles, nous montrons que les distributions asymptotiques des statistiques définies dans (1.47) et (1.48) ne sont plus des distributions de chi-deux mais un mélange de distribution chi-deux pondéré par les valeurs propres de la matrice de covariance asymptotique du vecteur d'autocorrélation.

Un développement limité de Taylor, nous donne  $h = 1, \dots, m$

$$\sqrt{N}\hat{\gamma}(v, h) = \sqrt{N}\gamma(v, h) + \left( \mathbb{E} \left[ \epsilon_{nT+v-h} \frac{\partial}{\partial \alpha'} \epsilon_{nT+v}(\alpha_0) \right] \right) \sqrt{N}(\hat{\alpha}_{LS} - \alpha_0) + o_{\mathbb{P}}(1). \quad (1.49)$$

Nous remarquons maintenant que dans l'équation (1.49),

$\mathbb{E}[\epsilon_{nT+v-h}(\partial \epsilon_{nT+v}(\alpha_0) / \partial \alpha')]$   $\mathbb{E}[\epsilon_{nT+v-h}(\partial \epsilon_{nT+v}(\alpha_0) / \partial \alpha')]$  est la ligne  $h$  de la matrice  $\Psi_m(v) \in \mathbb{R}^{m \times (p+q)T}$  suivante définie par

$$\Psi_m(v) = \mathbb{E} \left\{ \left( \begin{array}{c} \epsilon_{nT+v-1} \\ \vdots \\ \epsilon_{nT+v-m} \end{array} \right) \frac{\partial \epsilon_{nT+v}}{\partial \alpha'} \right\}. \quad (1.50)$$

Donc, pour  $h = 1, \dots, m$ , l'équation (1.49) dévient

$$\sqrt{N}\hat{\gamma}_m(v) = \left( \sqrt{N}\hat{\gamma}(v, 1), \dots, \sqrt{N}\hat{\gamma}(v, m) \right)' = \sqrt{N}\gamma_m(v) + \Psi_m(v) \sqrt{N}(\hat{\alpha}_{LS} - \alpha_0) + o_{\mathbb{P}}(1). \quad (1.51)$$

De (1.51), il est clair que la distribution asymptotique des autocovariances résiduelles  $\sqrt{N}\hat{\gamma}_m(v)$  est lié au comportement asymptotique de  $\sqrt{N}(\hat{\alpha}'_{LS} - \alpha'_0, \gamma'_m(v))'$ .

Au vu de [32, Théorème 1] et (H3''), nous avons presque sûrement

$\hat{\alpha}_{LS} \rightarrow \alpha_0 \in \overset{\circ}{\Theta}$ . Ainsi,  $\partial Q_N^{\omega^2}(\hat{\alpha}_{LS}) / \partial \alpha = 0$  pour de  $N$  suffisamment grand et un développement de Taylor donne

$$\sqrt{N} \frac{\partial}{\partial \alpha} O_N^{\omega^2}(\alpha_0) + J(\alpha_0, \omega^2) \sqrt{N}(\hat{\alpha}_{LS} - \alpha_0) = o_{\mathbb{P}}(1), \quad (1.52)$$

où

$$O_N^{\omega^2}(\alpha) = \frac{1}{2N} \sum_{n=0}^{N-1} \sum_{v=1}^T \omega_v^{-2} \epsilon_{nT+v}^2(\alpha).$$

En conséquence de (1.52), nous avons

$$\sqrt{N}(\hat{\alpha}_{LS} - \alpha_0) = -J(\alpha_0, \omega^2)^{-1} \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \sum_{v=1}^T \omega_v^{-2} \epsilon_{nT+v} \frac{\partial \epsilon_{nT+v}(\alpha_0)}{\partial \alpha} + o_{\mathbb{P}}(1). \quad (1.53)$$

**Proposition 1.5.1** *Supposons  $p + q > 0$ . Sous les hypothèses (H0''), (H1''), (H2''), (H3'') et (H4''), le vecteur aléatoire*

$$\sqrt{N} \left( (\hat{\alpha}_{LS} - \alpha_0)', \gamma'_m(v) \right)'$$

a une distribution limite normale centrée avec la matrice de covariance

$$\Xi(\nu) = \begin{pmatrix} \Omega^{\text{LS}} & \Sigma_{\hat{\alpha}, \gamma_m(\nu)} \\ \Sigma'_{\hat{\alpha}, \gamma_m(\nu)} & \Sigma_{\gamma_m(\nu)} \end{pmatrix} = \sum_{h=-\infty}^{\infty} \mathbb{E} \left[ W_{nT+\nu} W'_{(n-h)T+\nu} \right],$$

à partir de (1.46) et (1.53) nous avons

$$W_{nT+\nu} = \begin{pmatrix} W_{1,nT+\nu} \\ W_{2,nT+\nu} \end{pmatrix} = \begin{pmatrix} -J(\alpha_0, \omega^2)^{-1} \sum_{v=1}^T \omega_v^{-2} \epsilon_{nT+\nu} \frac{\partial \epsilon_{nT+\nu}(\alpha_0)}{\partial \alpha} \\ (\epsilon_{nT+\nu-1}, \dots, \epsilon_{nT+\nu-m})' \epsilon_{nT+\nu} \end{pmatrix}. \quad (1.54)$$

Le théorème suivant est une extension des résultats donnés dans [33], qui fournit la distribution limite des autocovariances et autocorrélations résiduelles des modèles PARMA faibles pour une saison donnée  $\nu$ . Nous fournissons également la distribution limite des autocovariances et autocorrélations résiduelles globales des modèles PARMA faibles pour toutes les saisons. Soient

$$D_m(\nu) = \text{Diag}(\sigma_\nu \sigma_{\nu-1}, \dots, \sigma_\nu \sigma_{\nu-m})$$

et

$$D_{mT} = \text{Diag}(\sigma_1 \sigma_{1-1}, \dots, \sigma_1 \sigma_{1-m}, \dots, \sigma_T \sigma_{T-1}, \dots, \sigma_T \sigma_{T-m}).$$

**Theorem 1.5.2** *Sous les hypothèses de la proposition 1.5.1, nous avons*

$$\sqrt{N} \hat{\gamma}_m(\nu) \xrightarrow[N \rightarrow +\infty]{d} \mathcal{N}(0, \nabla_{\hat{\gamma}}(\nu, \nu)) \quad (1.55)$$

où

$$\begin{aligned} \nabla_{\hat{\gamma}}(\nu, \nu) &= \Sigma_{\gamma_m(\nu)} + \Psi_m(\nu) \Omega^{\text{LS}} \Psi'_m(\nu) + \Psi_m(\nu) \Sigma_{\hat{\alpha}, \gamma_m(\nu)} + \Sigma'_{\hat{\alpha}, \gamma_m(\nu)} \Psi'_m(\nu) \text{ quand } p + q > 0 \\ \text{et } \nabla_{\hat{\gamma}}(\nu, \nu) &= \Sigma_{\gamma_m(\nu)} \text{ quand } p = q = 0. \end{aligned}$$

Alors :

$$\sqrt{N} \hat{\gamma}_{mT} \xrightarrow[N \rightarrow +\infty]{d} \mathcal{N}(0, \nabla_{\hat{\gamma}}) \quad (1.56)$$

où la matrice de covariance asymptotique  $\nabla_{\hat{\gamma}}$  est une matrice en bloc, avec les variances asymptotiques données par  $\nabla_{\hat{\gamma}}(\nu, \nu)$ ,  $\nu = 1, \dots, T$  et les covariances asymptotiques données par

$$\nabla_{\hat{\gamma}}(\nu, \nu') = \Sigma_{\gamma_m}(\nu, \nu') + \Psi_m(\nu) \Omega^{\text{LS}} \Psi'_m(\nu') + \Psi_m(\nu) \Sigma_{\hat{\alpha}, \gamma_m}(\nu, \nu') + \Sigma'_{\hat{\alpha}, \gamma_m}(\nu) \Psi'_m(\nu'), \quad \nu \neq \nu'.$$

Nous avons également

$$\sqrt{N} \hat{\rho}_m(\nu) \xrightarrow[N \rightarrow +\infty]{d} \mathcal{N}(0, \nabla_{\hat{\rho}}(\nu, \nu)) \quad \text{où } \nabla_{\hat{\rho}}(\nu, \nu) = D_m^{-1}(\nu) \nabla_{\hat{\gamma}}(\nu, \nu) D_m^{-1}(\nu) \quad (1.57)$$

et

$$\sqrt{N} \hat{\rho}_{mT} \xrightarrow[N \rightarrow +\infty]{d} \mathcal{N}(0, \nabla_{\hat{\rho}}) \quad \text{où } \nabla_{\hat{\rho}} = D_{mT}^{-1} \nabla_{\hat{\gamma}} D_{mT}^{-1}. \quad (1.58)$$

Les matrices de variances asymptotiques  $\nabla_{\hat{\gamma}}(\nu, \nu)$  et  $\nabla_{\hat{\rho}}(\nu, \nu)$  dépendent des matrices inconnues  $\Xi(\nu)$ ,  $\Psi_m(\nu)$  et du scalaire  $\sigma_{\nu-h}$  pour  $h = 1, \dots, m$ . La matrice  $\Psi_m(\nu)$  et la variance du bruit  $\sigma_{\nu-h}$  peuvent être respectivement estimées empiriquement par

$$\hat{\Psi}_m(\nu) = \frac{1}{N} \sum_{n=0}^{N-1} \left\{ (\hat{\epsilon}_{nT+\nu-1}, \dots, \hat{\epsilon}_{nT+\nu-m})' \frac{\partial \hat{\epsilon}_{nT+\nu}}{\partial \alpha'} \right\} \text{ et } \hat{\sigma}_{\nu-h} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} \hat{\epsilon}_{nT+\nu-h}^2}.$$

De la même manière la matrice  $J(\alpha_0, \omega^2)$  peut être estimée par  $J_N(\hat{\alpha}_{\text{LS}}, \omega^2)$  d'ordre  $(p+q)T$  définie par

$$J_N(\hat{\alpha}_{\text{LS}}, \omega^2) = \sum_{v=1}^T \omega_v^2 \frac{1}{N} \sum_{n=0}^{N-1} \left\{ \frac{\partial}{\partial \alpha} e_{nT+v}(\hat{\alpha}_{\text{LS}}) \right\} \left\{ \frac{\partial}{\partial \alpha} e_{nT+v}(\hat{\alpha}_{\text{LS}}) \right\}'. \quad (1.59)$$

### 1.5.3 Version modifiée du test portmanteau

Pour l'estimation de la matrice  $\Xi$ , nous utilisons l'approche basée sur l'estimation de la densité spectrale de  $(W_{nT+v})_{n \in \mathbb{Z}}$  par un VAR paramétrique, définie dans la partie 1.4.4. Le processus stationnaire  $(W_{nT+v})_{n \in \mathbb{Z}}$  admet la décomposition de Wold

$$W_{nT+v} = u_{nT+v} + \sum_{i=1}^{\infty} \omega_i(v) u_{nT+v-i},$$

où  $u_{nT+v}$  est un bruit blanc périodique faible  $((p+q)T+m)$ -dimensionnel. Supposons que la matrice de covariance  $\Sigma_u(v) := \text{Var}(u_{nT+v})$  est non singulière, que  $\sum_{i=1}^{\infty} \|\omega_i(v)\| < \infty$ , où  $\|\cdot\|$  désigne toute norme sur l'espace des matrices réelles  $((p+q)T+m) \times ((p+q)T+m)$ , et que  $\det\{I_{(p+q)T+m} + \sum_{i=1}^{\infty} \omega_i(v) z^i\} \neq 0$  si  $|z| \leq 1$ . Ensuite,  $(W_{nT+v})_{n \in \mathbb{Z}}$  admet une représentation AR( $\infty$ ) (voir [3]) de la forme

$$\Phi_v(L)W_{nT+v} = W_{nT+v} + \sum_{i=1}^{\infty} \Phi_i(v)W_{nT+v-i} = u_{nT+v}, \quad (1.60)$$

tel que  $\sum_{i=1}^{\infty} \|\Phi_i(v)\| < \infty$  et  $\det\{\Phi_v(z)\} \neq 0$  si  $|z| \leq 1$ .

Cette approche paramétrique a été étudiée par [10] (voir aussi [28]) est basée sur l'expression

$$\Xi(v) = \Phi_v^{-1}(1)\Sigma_u(v)\Phi_v^{-1'}(1). \quad (1.61)$$

Puisque  $W_{nT+v}$  n'est pas observé, nous introduisons  $\hat{W}_{nT+v} \in \mathbb{R}^{T(p+q)+m}$  obtenu en remplaçant  $\epsilon_{nT+v}(\alpha_0)$  par  $e_{nT+v}(\hat{\alpha}_{LS})$  et  $J(\alpha_0, \omega^2)$  par sa partie empirique ou observable  $J_N(\hat{\alpha}_{LS}, \omega^2)$  dans (1.54). Nous définissons  $\hat{\Phi}_{v,r}(z) = I_{(p+q)T+m} + \sum_{i=1}^r \hat{\Phi}_{r,i}(v)z^i$ , où  $\hat{\Phi}_{r,1}(v), \dots, \hat{\Phi}_{r,r}(v)$  sont les coefficients de la régression des moindres carrés de  $\hat{W}_{nT+v}$  sur  $\hat{W}_{nT+v-1}, \dots, \hat{W}_{nT+v-r}$  pour une période  $v$  donnée. Notons  $\hat{u}_{r,nT+v}$  le résidus de cette régression, et soit  $\hat{\Sigma}_{\hat{u}_r}(v)$  la variance empirique de  $\hat{u}_{r,1}, \dots, \hat{u}_{r,N}$ .

Dans le cas de processus linéaires avec des innovations indépendantes, [10] a montré que la densité spectrale peut être estimée de manière consistante en ajustant des modèles autorégressifs d'ordre  $r = r(N)$ , lorsque  $r$  tend vers l'infini et  $r^3/N$  tend vers 0 quand  $N$  tend vers l'infini. Il existe des différences avec [10] :  $(W_{nT+v})_{n \in \mathbb{Z}}$  est multivarié, il n'est pas observé directement et est remplacé par  $(\hat{W}_{nT+v})_{n \in \mathbb{Z}}$ . Il est montré que ce résultat reste valable pour le processus linéaire multivarié  $(W_{nT+v})_{n \in \mathbb{Z}}$  avec des innovations non indépendantes (voir [17, 18, 33] pour les références dans les modèles ARMA faibles (multivariés)). Nous étendrons ces résultats aux modèles PARMA faibles. Nous sommes maintenant en mesure d'énoncer le théorème suivant.

**Theorem 1.5.3** *Sous les hypothèses du théorème 1.5.2, nous supposons que le processus  $(W_{nT+v})$  défini dans (1.54) admet une représentation périodique VAR( $\infty$ ) dont les racines de  $\det\Phi(z) = 0$  sont à l'extérieur du disque unité,  $\|\Phi_i\| = o(i^{-2})$ , et que la matrice  $\Sigma_u(v) = \text{Var}(u_{nT+v})$  est non-singulière. De plus nous supposons que  $\mathbb{E}\|\epsilon_n\|^{8+4\kappa} < \infty$  et  $\sum_{k=0}^{\infty} \{\alpha_{\epsilon}(k)\}^{\kappa/(2+\kappa)} < \infty$  pour un réel  $\kappa > 0$ . Alors l'estimateur de  $\Xi(v)$*

$$\hat{\Xi}^{SP}(v) := \hat{\Phi}_{v,r}^{-1}(1)\hat{\Sigma}_{\hat{u}_r}(v)\hat{\Phi}_{v,r}^{-1'}(1) \rightarrow \Xi(v)$$

en probabilité quand  $r = r(N) \rightarrow +\infty$  et  $r^3/N \rightarrow 0$  quand  $N \rightarrow +\infty$ .

Du théorème 1.5.2 nous pouvons déduire le résultat suivant qui donne la distribution limite exacte des statistiques standards de portmanteau (1.47) sous des hypothèses générales sur le processus d'innovation du modèle PARMA( $p, q$ ) ajusté.

**Theorem 1.5.4** *Sous les hypothèses du théorème 1.5.2 et  $(\mathcal{H}0)$ , les statistiques  $Q_m^{\text{LB}}(\mathbf{v})$  et  $Q_m^{\text{LBM}}(\mathbf{v})$  définies dans (1.47) convergent en distribution, quand  $N \rightarrow +\infty$ , vers*

$$Z_m^{\mathbf{v}}(\xi_m^{\mathbf{v}}) = \sum_{i=1}^m \xi_{i,m}^{\mathbf{v}} Z_i^2$$

où  $\xi_m^{\mathbf{v}} = (\xi_{1,m}^{\mathbf{v}}, \dots, \xi_{m,m}^{\mathbf{v}})'$  est le vecteur de valeurs propres de la matrice  $\nabla_{\hat{\rho}}(\mathbf{v}, \mathbf{v})$  définie dans (1.57) et  $Z_1, \dots, Z_m$  sont des variables indépendantes de la loi  $\mathcal{N}(0, 1)$ .

La distribution asymptotiques des statistiques globales du test portmanteau (qui prend en compte toutes les saisons)  $Q_m^{\text{LB}}$  et  $Q_m^{\text{LBM}}$  définie dans (1.48) est également une somme pondérée de variables aléatoires de chi-deux :

$$Z_m(\xi_m) = \sum_{i=1}^{mT} \xi_{i,mT} Z_i^2$$

où  $\xi_{mT} = (\xi_{1,mT}, \dots, \xi_{mT,mT}^{\mathbf{v}})'$  est le vecteur de valeurs propres de la matrice  $\nabla_{\hat{\rho}}$  donnée (1.58) et  $Z_1, \dots, Z_{mT}$  sont également des variables indépendantes de la loi  $\mathcal{N}(0, 1)$ .

Au vue du théorème 1.5.4, les approximations  $\chi_{m-(p+q)}^2$  et  $\chi_{mT-(p+q)T}^2$  standard ne sont plus valables pour approximer la distribution asymptotique de (1.47) et (1.48) dans le cadre des modèles PARMA( $p; q$ ) faibles.

La distribution limite dépend du paramètre de nuisance périodique  $\sigma_{\mathbf{v}-h}$  pour  $h = 1, \dots, m$  et les éléments de  $\Xi(\mathbf{v})$ . Par conséquent, afin d'obtenir la distribution asymptotique des statistiques de portmanteau (1.47) (resp. (1.48)) sous des hypothèses faibles sur le bruit, il faut un estimateur convergent de la matrice de covariance asymptotique  $\nabla_{\hat{\rho}}(\mathbf{v}, \mathbf{v})$  (resp.  $\nabla_{\hat{\rho}}$ ).

Notons par  $\hat{\nabla}_{\hat{\rho}}(\mathbf{v}, \mathbf{v})$  (resp.  $\hat{\nabla}_{\hat{\rho}}$ ) la matrice obtenue en remplaçant  $\Xi(\mathbf{v})$  par  $\hat{\Xi}(\mathbf{v})$ ,  $\Psi_m(\mathbf{v})$  par  $\hat{\Psi}_m(\mathbf{v})$  et  $\sigma_{\mathbf{v}-h}$  par  $\hat{\sigma}_{\mathbf{v}-h}$  dans  $\nabla_{\hat{\rho}}(\mathbf{v}, \mathbf{v})$  (resp.  $\nabla_{\hat{\rho}}$ ). Notons  $\hat{\xi}_m^{\mathbf{v}} = (\hat{\xi}_{1,m}^{\mathbf{v}}, \dots, \hat{\xi}_{m,m}^{\mathbf{v}})'$  (resp.  $\hat{\xi}_m = (\hat{\xi}_{1,mT}, \dots, \hat{\xi}_{mT,mT})'$ ) le vecteur de valeurs propres de  $\hat{\nabla}_{\hat{\rho}}(\mathbf{v}, \mathbf{v})$  (resp.  $\hat{\nabla}_{\hat{\rho}}$ ).

Au seuil asymptotique  $\alpha$  et pour une saison  $\mathbf{v}$  donnée, le test LB (resp. le test BP) consiste à rejeter l'adéquation du modèle PARMA( $p, q$ ) faible quand

$$\lim_{N \rightarrow \infty} \mathbb{P}(Q_m^{\text{BP}}(\mathbf{v}) > S_m^{\mathbf{v}}(1 - \alpha)) = \lim_{N \rightarrow \infty} \mathbb{P}(Q_m^{\text{LBM}}(\mathbf{v}) > S_m^{\mathbf{v}}(1 - \alpha)) = \alpha,$$

où  $S_m^{\mathbf{v}}(1 - \alpha)$  est tel que  $\mathbb{P}\{Z_m^{\mathbf{v}}(\hat{\xi}_m^{\mathbf{v}}) > S_m^{\mathbf{v}}(1 - \alpha)\} = \alpha$ .

Nous soulignons le fait que les versions modifiées des tests BP et LBM sont plus difficiles à mettre en œuvre. En effet, les valeurs critiques doivent être calculées à partir des données alors que les valeurs critiques de la méthode standard sont, quand à elles, simplement déduites d'une table  $\chi^2$ .

Nous évaluerons les  $p$ -valeurs

$$\mathbb{P}\{Z_m(\hat{\xi}_m^{\mathbf{v}}) > Q_m^{\text{BP}}(\mathbf{v})\} \text{ et } \mathbb{P}\{Z_m(\hat{\xi}_m^{\mathbf{v}}) > Q_m^{\text{LBM}}(\mathbf{v})\}, \text{ avec } Z_m(\hat{\xi}_m^{\mathbf{v}}) = \sum_{i=1}^m \hat{\xi}_{i,m}^{\mathbf{v}} Z_i^2,$$

par l'algorithme de Imhof (voir [49]). Les procédures de test pour  $Q_m^{\text{BP}}$  et  $Q_m^{\text{LBM}}$  définies dans (1.48) sont similaires, mais sont basées sur les  $mT$  valeurs propres empiriques de  $\hat{\nabla}_{\hat{\rho}}$ .

## 1.5.4 La distribution asymptotique auto-normalisée des autocorrélations résiduelles

Dans cette section nous proposons la méthode d'auto-normalisation introduite précédemment.



Soit  $(B_K(r))_{r \geq 0}$  un mouvement brownien  $K$ -dimensionnel à partir de 0. Pour  $K \geq 1$ , nous désignons  $\mathcal{U}_K$  la variable aléatoire définie par

$$\mathcal{U}_K = B'_K(1)V_K^{-1}B_K(1) \text{ où } V_K = \int_0^1 (B_K(r) - rB_K(1))(B_K(r) - rB_K(1))' dr.$$

Notons par  $\Lambda(v)$  la matrice par blocs de  $\mathbb{R}^{m \times (T(p+q)+m)}$  définis par  $\Lambda(v) = (\Psi_m(v)|I_m)$  où  $I_k$  est la matrice d'identité d'ordre  $k$ . Compte tenu de (1.51) et (1.53), nous en déduisons que  $\sqrt{N}\hat{\Upsilon}_m(v) = N^{-1/2} \sum_{n=1}^{N-1} \Lambda(v)W_{nT+v} + o_{\mathbb{P}}(1)$ .

Nous avons besoin d'appliquer le théorème centrale limite fonctionnel pour le processus  $W = (W_{nT+v})_{nT+v \geq 1}$  (voir le Lemme 1 dans [54]).

Enfin, nous définissons la matrice de normalisation  $C_m(v, v) \in \mathbb{R}^{m \times m}$  pour un  $v$  donnée par

$$C_m(v, v) = \frac{1}{N^2} \sum_{n=0}^{N-1} S_n(v)S'_n(v) \text{ où } S_n(v) = \sum_{j=0}^n (\Lambda(v)W_{jT+v} - \Upsilon_m(v)). \quad (1.62)$$

De la même manière nous notons par  $C_m \in \mathbb{R}^{mT \times mT}$  la matrice globale de normalisation par blocs pour toutes les saisons avec le bloc diagonal donné par  $C_m(v, v)$ , pour  $v = 1, \dots, T$  et

$$C_m(v, v') = \frac{1}{N^2} \sum_{n=0}^{N-1} S_n(v)S'_n(v') \text{ où } S_n(v') = \sum_{j=0}^n (\Lambda(v)W_{jT+v'} - \Upsilon_m(v')) \text{ pour } v \neq v'.$$

Pour assurer l'inversibilité de la matrice de normalisation  $C_m(v, v)$ , nous avons besoin de l'hypothèse technique suivante sur la distribution  $\epsilon_n$ .

**(H5'')** : Le processus  $(\epsilon_n)_{n \in \mathbb{Z}}$  a une densité strictement positive sur un voisinage de zéro.

La proposition suivante donne l'inversibilité des matrices de normalisation  $C_m(v, v)$  et  $C_m$ .

**Proposition 1.5.5** *Sous les hypothèses du théorème 1.5.2 et (H5''), les matrices  $C_m(v, v)$  et  $C_m$  sont presque sûrement non singulière.*

Le théorème suivant énonce la distribution asymptotiques des autocovariances et autocorrélations résiduelles.

**Theorem 1.5.6** *Nous supposons que  $p + q > 0$ . Sous les hypothèses du théorème 1.5.2, (H5'') et sous l'hypothèse nulle  $\mathcal{H}0$ , pour une saison  $v$  donnée nous avons*

$$N\hat{\Upsilon}'_m(v)C_m^{-1}(v, v)\hat{\Upsilon}_m(v) \xrightarrow[N \rightarrow +\infty]{d} \mathcal{U}_m \text{ et } N\hat{\rho}'_m(v)D_m(v)C_m^{-1}(v, v)D_m(v)\hat{\rho}_m(v) \xrightarrow[N \rightarrow +\infty]{d} \mathcal{U}_m.$$

*Pour les autocovariances et autocorrélations résiduelles globales sur toutes les saisons, nous avons également*

$$N\hat{\Upsilon}'_{mT}C_{mT}^{-1}\hat{\Upsilon}_{mT} \xrightarrow[N \rightarrow +\infty]{d} \mathcal{U}_{mT} \text{ et } N\hat{\rho}'_{mT}D_{mT}C_{mT}^{-1}D_{mT}\hat{\rho}_{mT} \xrightarrow[N \rightarrow +\infty]{d} \mathcal{U}_{mT}.$$

Bien entendu, le théorème est inutilisable dans la pratique car la matrice de normalisation  $C_m(v, v)$ ,  $C_{mT}$  et le paramètre de nuisance  $\sigma_{v-h}$  ne sont pas observables. Ce problème sera corrigé lorsque nous remplaçons la matrice  $C_m(v, v)$ ,  $C_{mT}$  et la variance du bruit  $\sigma_{v-h}$  par leurs parties empiriques ou observables. Ainsi nous notons

$$\hat{C}_m(v, v) = \frac{1}{N^2} \sum_{n=0}^{N-1} \hat{S}_n(v)\hat{S}'_n(v) \text{ où } \hat{S}_n(v) = \sum_{j=0}^n (\hat{\Lambda}(v)\hat{W}_{jT+v} - \hat{\Upsilon}_m(v)),$$

avec  $\hat{\Lambda}(v) = (\hat{\Psi}_m(v)|I_m)$  et où  $\hat{W}_{nT+v}$  et  $\hat{\sigma}_{v-h}$  sont définies dans 1.5.2.

Nous sommes donc en mesure d'énoncer la version applicable du théorème 1.5.6.



**Theorem 1.5.1** *Sous les hypothèses du théorème 1.5.6, pour une saison  $v$  donnée nous avons*

$$N\hat{\gamma}'_m(v)\hat{C}_m^{-1}(v,v)\hat{\gamma}_m(v) \xrightarrow[N \rightarrow +\infty]{d} \mathcal{U}_m \text{ et } Q_m^{\text{SN}}(v) =: N\hat{\rho}'_m(v)\hat{D}_m(v)\hat{C}_m^{-1}(v,v)\hat{D}_m(v)\hat{\rho}_m(v) \xrightarrow[N \rightarrow +\infty]{d} \mathcal{U}_m.$$

*Nous avons également pour toutes les saisons*

$$N\hat{\gamma}'_{mT}\hat{C}_{mT}^{-1}\hat{\gamma}_{mT} \xrightarrow[N \rightarrow +\infty]{d} \mathcal{U}_{mT} \text{ et } Q_m^{\text{SN}} =: N\hat{\rho}'_{mT}\hat{D}_{mT}\hat{C}_{mT}^{-1}\hat{D}_{mT}\hat{\rho}_{mT} \xrightarrow[N \rightarrow +\infty]{d} \mathcal{U}_{mT}.$$

Sur la base du résultat ci-dessus, nous proposons la version modifiée de la statistique de LBM pour une saison  $v$  donnée, lorsque l'on utilise la statistique

$$\tilde{Q}_m^{\text{SN}}(v) = N\hat{\rho}'_m(v)M_m^{1/2}(v, N, T)\hat{D}_m(v)\hat{C}_m^{-1}(v, v)\hat{D}_m(v)M_m^{1/2}(v, N, T)\hat{\rho}_m(v), \quad (1.63)$$

où  $M_m(v, N, T) \in \mathbb{R}^{m \times m}$  est une matrice diagonale avec  $(\varrho(1, v, N, T), \dots, \varrho(m, v, N, T))$  sur la diagonale.

De même, nous proposons une version modifiée de la statistique LBM pour toutes les saisons lorsque l'on utilise la statistique suivante

$$\tilde{Q}_m^{\text{SN}} = N\hat{\rho}'_{mT}M_{mT}^{1/2}\hat{D}_{mT}\hat{C}_{mT}^{-1}\hat{D}_{mT}M_{mT}^{1/2}\hat{\rho}_{mT}, \quad (1.64)$$

où  $M_{mT} \in \mathbb{R}^{mT \times mT}$  est une matrice diagonale avec

$$(\varrho(1, 1, N, T), \dots, \varrho(m, 1, N, T), \dots, \varrho(1, T, N, T), \dots, \varrho(m, T, N, T))$$

sur la diagonale.

Au seuil asymptotique  $\alpha$  et pour une saison  $v$  donnée, le test d'auto-normalisation LBM (resp. le test BP) consiste à rejeter l'adéquation du modèle  $\text{PARMA}_T(p; q)$  faible quand

$$\tilde{Q}_m^{\text{SN}}(v) > \mathcal{U}_m(1 - \alpha) \quad (\text{resp. } Q_m^{\text{SN}}(v) > \mathcal{U}_m(1 - \alpha)),$$

où les valeurs critiques  $\mathcal{U}_m(1 - \alpha)$  sont tabulées dans [54, Table 1]. De la même manière nous avons pour le test global

$$\tilde{Q}_m^{\text{SN}} > \mathcal{U}_{mT}(1 - \alpha) \quad (\text{resp. } Q_m^{\text{SN}} > \mathcal{U}_{mT}(1 - \alpha)).$$

## 1.6 Résultats du chapitre 5

Dans cet chapitre nous présentons un modèle saisonnier autorégressif moyenne mobile avec des paramètres périodiques (SPARMA). Nous allons étudier les propriétés asymptotiques de l'estimateur de moindres carrés quasi-généralisés des modèles SPARMA faibles. Nous estimons la matrice de la variance asymptotique de l'estimateur des moindres carrés quasi-généralisés des modèles SPARMA faibles. Nous proposons un estimateur convergent de cette matrice. Nous développons la distribution asymptotique des statistiques portmanteau des modèles SPARMA faibles. Nous en déduisons des tests portmanteau modifiés pour tester l'adéquation des modèles SPARMA faibles. Nous proposons également dans cet chapitre la méthode d'auto-normalisation pour les modèles SPARMA faibles.

### 1.6.1 Estimations des modèles SPARMA faibles

Soit  $a(v) = (a_1(v), \dots, a_p(v))'$  et  $b(v) = (b_1(v), \dots, b_q(v))'$  respectivement les vecteurs des paramètres non saisonniers autorégressifs et moyennes mobiles pour une saison  $v$  donnée. et soit  $\mathbf{a}(v) = (\mathbf{a}_1(v), \dots, \mathbf{a}_p(v))'$  et  $\mathbf{b}(v) = (\mathbf{b}_1(v), \dots, \mathbf{b}_q(v))'$  respectivement les vecteurs des paramètres saisonniers autorégressifs et moyennes mobiles pour une saison  $v$  donnée. L'ensemble de dimension  $T(p+q+P+Q)$  de tous les paramètres SPARMA est noté par

$$\alpha = (a(1)', \dots, a(T)', b(1)', \dots, b(T)', \mathbf{a}(1)', \dots, \mathbf{a}(T)', \mathbf{b}(1)', \dots, \mathbf{b}(T))' \in \mathbb{R}^{k_0 T}$$

avec  $k_0 = (p+q+P+Q)$ . La variance  $\sigma^2 = (\sigma^2(1), \dots, \sigma^2(T))'$  périodiques du bruit blanc est considérée comme un paramètre de nuisance. Soient les observations  $X_1, \dots, X_{NT}$  solution du modèle SPARMA défini en (1.14) avec la vraie valeur du paramètre  $\alpha = \alpha_0$  et  $\sigma^2 = \sigma_0^2$ . Le vrai paramètre inconnu  $\alpha_0$  appartient à l'espace des paramètres

$$\Delta = \left\{ \alpha = (a(1)', \dots, a(T)', b(1)', \dots, b(T)', \mathbf{a}(1)', \dots, \mathbf{a}(T)', \mathbf{b}(1)', \dots, \mathbf{b}(T))' \right. \\ \left. \text{tel que } (\mathbf{H1}'') \text{ est vérifiée} \right\}.$$

Sans perte de généralité, nous supposons que  $\Delta$  est compact. Pour  $\alpha \in \Delta$ , soit  $\epsilon_{nT+v}(\alpha)$  la solution stationnaire périodique de seconde ordre de

$$\begin{aligned} \epsilon_{nT+v}(\alpha) = & X_{nT+v} - \sum_{i=1}^p a_i X_{nT+v-i} - \sum_{j=1}^P \mathbf{a}_j X_{nT+v-Tj} + \sum_{i=1}^p \sum_{j=1}^P a_i \mathbf{a}_j X_{nT+v-Tj-i} + \sum_{i=1}^q b_i \epsilon_{nT+v-i}(\alpha) \\ & + \sum_{j=1}^Q \mathbf{b}_j \epsilon_{nT+v-Tj}(\alpha) - \sum_{i=1}^q \sum_{j=1}^Q b_i \mathbf{b}_j \epsilon_{nT+v-Tj-i}(\alpha), \end{aligned} \quad (1.65)$$

Les variables  $\epsilon_{nT+v}(\alpha)$  peuvent être approximées, pour  $0 < nT+v \leq NT$ , par  $e_{nT+v}(\alpha)$  définies récursivement par

$$\begin{aligned} e_{nT+v}(\alpha) = & X_{nT+v} - \sum_{i=1}^p a_i X_{nT+v-i} - \sum_{j=1}^P \mathbf{a}_j X_{nT+v-Tj} + \sum_{i=1}^p \sum_{j=1}^P a_i \mathbf{a}_j X_{nT+v-Tj-i} + \sum_{i=1}^q b_i e_{nT+v-i}(\alpha) \\ & + \sum_{j=1}^Q \mathbf{b}_j e_{nT+v-Tj}(\alpha) - \sum_{i=1}^q \sum_{j=1}^Q b_i \mathbf{b}_j e_{nT+v-Tj-i}(\alpha), \end{aligned} \quad (1.66)$$

où les valeurs initiales sont mises à zéro :  $e_0(\alpha) = \dots = e_{1-q-TQ}(\alpha) = X_0 = \dots = X_{1-p-TP} = 0$ . Ces valeurs initiales sont asymptotiquement négligeables uniformément en  $\alpha$ .

La variable aléatoire  $\hat{\alpha}_{OLS}$  est appelée l'estimateur de moindres carrés ordinaires de  $\alpha$  si elle satisfait, presque sûrement,

$$S_N(\hat{\alpha}_{OLS}) = \min_{\alpha \in \Delta} S_N(\alpha) \quad \text{où} \quad S_N(\alpha) = \frac{1}{2N} \sum_{n=0}^{N-1} \sum_{v=1}^T e_{nT+v}^2(\alpha).$$

En raison de la présence d'innovation hétéroscédastiques, [32] (voir également [6]) ont montré que l'estimateur OLS est inefficace par rapport aux estimateur des moindres carrés pondérés. Donc dans cet chapitre nous avons choisi de considérer l'estimateur de moindres carrés quasi-généralisé (QLS). Nous définissons d'abord l'estimateur de moindres carrés généralisés (GLS) qui est optimal parmi les estimateurs de moindres carrés pondérés (voir [32] )

$$Q_N^{\sigma^2}(\hat{\alpha}_{GLS}) = \min_{\alpha \in \Delta} Q_N^{\sigma^2}(\alpha) \quad \text{où} \quad Q_N^{\sigma^2}(\alpha) = \frac{1}{2N} \sum_{n=0}^{N-1} \sum_{v=1}^T \sigma_v^{-2} e_{nT+v}^2(\alpha) \quad \text{avec} \quad \sigma^2 = (\sigma_1^2, \dots, \sigma_T^2)'. \quad (1.67)$$

Cet estimateur suppose que le paramètre  $\sigma$  soit connu mais dans la pratique on doit l'estimer. D'où un estimateur consistant possible de  $\sigma_v$  est donné par

$$\hat{\sigma}_v^2 = \frac{1}{N} \sum_{n=0}^{N-1} e_{nT+v}^2(\hat{\alpha}_{OLS}).$$

Cela, nous permet de définir pour tout estimateur consistant  $\hat{\sigma}^2 = (\hat{\sigma}_1^2, \dots, \hat{\sigma}_T^2)'$  de  $\sigma^2$ , l'estimateur QLS comme suit :

$$\hat{\alpha}_{QLS} = \hat{\alpha}_{GLS}^{\hat{\sigma}^2}. \quad (1.68)$$

## 1.6.2 Propriétés asymptotiques de l'estimateur des moindres carrés quasi-généralisés

Pour établir la consistance et la normalité asymptotique de l'estimateur QLS, nous utiliserons les hypothèses définies dans la partie 1.5.

**Theorem 1.6.1** Nous supposons le processus  $X_{nT+v}$  admet une représentation SPARMA. Soit  $\hat{\alpha}_{QLS}$  l'estimateur QLS défini en (1.68). Alors, sous certaines hypothèses nous avons :

$$\hat{\alpha}_{QLS} \xrightarrow{p} \alpha_0$$

**Theorem 1.6.2** Sous les hypothèses du théorème 1.6.1 et (H3''), nous avons :

$$\sqrt{n}(\hat{\alpha}_{QLS} - \alpha_0) \rightarrow \mathcal{N}(0, \Omega)$$

où

$$\Omega = (J(\alpha_0, \sigma^2))^{-1} I(\alpha_0, \sigma^2) (J(\alpha_0, \sigma^2))^{-1} \text{ avec}$$

$$J(\alpha_0, \sigma^2) = \sum_{v=1}^T \sigma_v^{-2} \mathbb{E} \left[ \left( \frac{\partial \epsilon_{nT+v}(\alpha)}{\partial \alpha} \right) \left( \frac{\partial \epsilon_{nT+v}(\alpha)}{\partial \alpha} \right)' \right]_{\alpha=\alpha_0}, \quad (1.69)$$

$$I(\alpha_0, \sigma^2) = \sum_{v=1}^T \sum_{v'=1}^T \sigma_v^{-2} \sigma_{v'}^{-2} \sum_{k=-\infty}^{\infty} \mathbb{E} \left[ \epsilon_v(\alpha) \epsilon_{kT+v'}(\alpha) \left( \frac{\partial \epsilon_v(\alpha)}{\partial \alpha} \right) \left( \frac{\partial \epsilon_{kT+v'}(\alpha)}{\partial \alpha} \right)' \right]_{\alpha=\alpha_0}. \quad (1.70)$$

**Remarque 1.6.1** Dans le cas standard des modèles SPARMA forts, i.e. quand  $\epsilon_{nT+v}$  est considéré comme une suite iid, nous avons  $I(\alpha_0, \sigma^2) = J(\alpha_0, \sigma^2)$ . La matrice de covariance asymptotique est donc réduite à :  $\Omega = (J(\alpha_0, \sigma^2))^{-1}$

**Remarque 1.6.2** Généralement, quand le bruit est non indépendant, nous avons  $I(\alpha_0, \sigma^2) \neq J(\alpha_0, \sigma^2)$ . La variance asymptotique  $\Omega = (J(\alpha_0, \sigma^2))^{-1} I(\alpha_0, \sigma^2) (J(\alpha_0, \sigma^2))^{-1}$  obtenue dans le cas des modèles SPARMA faibles peut être différente à celle obtenue dans le cas standard.

## 1.6.3 Estimation de la matrice de variance asymptotique

La variance asymptotique  $\Omega$  doit être estimée pour l'inférence statistique. La matrice  $J(\alpha_0, \sigma^2)$  peut être estimée empiriquement par la matrice carrée  $J_N(\hat{\alpha}, \hat{\sigma}^2)$  d'ordre  $(p + q + P + Q)T$  définie par :

$$J_N(\hat{\alpha}, \hat{\sigma}^2) = \sum_{v=1}^T \sigma_v^{-2} \frac{1}{N} \sum_{n=0}^{N-1} \left\{ \frac{\partial}{\partial \alpha} e_{nT+v}(\hat{\alpha}) \right\} \left\{ \frac{\partial}{\partial \alpha} e_{nT+v}(\hat{\alpha}) \right\}'. \quad (1.71)$$

L'estimateur  $\hat{\Omega}$  de  $\Omega$  est fortement consistant dans le cas d'un modèle SPARMA fort. En général, dans le cas des modèles SPARMA faibles, cet estimateur n'est pas consistant, nous devons donc disposer d'un estimateur consistant de  $I(\alpha_0)$ .

### 1.6.4 Estimation de la matrice asymptotique $I(\alpha_0)$

Soit le processus,

$$Y_n = \sum_{v=1}^T \sigma_v^{-2} \epsilon_n(\alpha_0) \frac{\partial \epsilon_n(\alpha_0)}{\partial \alpha} \quad (1.72)$$

Pour tout  $\alpha \in \Delta$ ,  $Y_n$  est une fonction mesurable de  $\{\epsilon_s, s \leq n\}$ .  
Le processus stationnaire  $(Y_n)$  admet une décomposition de Wold

$$Y_n = u_n + \sum_{i=1}^{\infty} \Psi_i(v) u_{n-i}$$

, où  $(u_{nT+v})$  est un bruit blanc périodique faible  $Tk_0$ -dimensionnel de matrice de variance  $\Sigma_u(v) := \text{Var}(u_n)$ . Supposons que  $\Sigma_u(v)$  est non singulière, que  $\sum_{i=1}^{\infty} \|\Psi_i(v)\| < \infty$ , et que  $\det(I_{Tk_0} + \sum_{i=1}^{\infty} \Psi_i(v) z^i) \neq 0$  quand  $|z| \leq 1$ . Alors  $(Y_n)$  admet une représentation VAR( $\infty$ ) (voir [3]) de la forme

$$\Phi_v(L) Y_n = Y_n - \sum_{i=1}^{\infty} \Phi_i(v) Y_{n-i} = u_n \quad (1.73)$$

tel que  $\sum_{i=1}^{\infty} \|\Phi_i(v)\| < \infty$  et  $\det\{\Phi_v(z)\} \neq 0$  pour tout  $|z| \leq 1$ , et nous obtenons

$$I(\theta_0) = \Phi_v^{-1}(1) \Sigma_u \Phi_v^{-1}(1)' \quad (1.74)$$

Considérons la régression de  $Y_n$  sur  $Y_{n-1}, \dots, Y_{n-r}$

$$Y_n = \sum_{i=1}^r \Phi_{r,i}(v) Y_{n-i} + u_{r,n} \quad (1.75)$$

où  $u_{r,n}$  ne sont pas corrélés avec  $Y_{n-1}, \dots, Y_{n-r}$ . Puisque  $Y_n(\alpha_0)$  n'est pas observé, nous introduisons  $\hat{Y}_n \in \mathbb{R}^{Tk_0}$  obtenue en remplaçant  $\epsilon_n$  par  $e_n$  et  $\alpha_0$  par  $\hat{\alpha}_n$  dans (1.72) :

$$\hat{Y}_n = \sum_{v=1}^T \sigma_v^{-2} e_n(\hat{\alpha}_n) \frac{\partial e_n(\hat{\alpha}_n)}{\partial \alpha} \quad (1.76)$$

Nous définissons  $\hat{\Phi}_{v,r}(z) = I_{Tk_0} + \sum_{k=1}^r \hat{\Phi}_{r,k}(v) z^k$ , où  $\hat{\Phi}_{r,1}(v), \dots, \hat{\Phi}_{r,r}(v)$  sont les coefficients de la régression des moindres carrés de  $\hat{Y}_n$  sur  $\hat{Y}_{n-1}, \dots, \hat{Y}_{n-r}$ . Notons  $\hat{u}_{r,n}$  le résidu de cette régression, et soit  $\hat{\Sigma}_{\hat{u}_r}(v)$  la variance empirique de  $\hat{u}_{r,1}, \dots, \hat{u}_{r,N}$ . Les estimateurs des moindres carrés de  $\underline{\Phi}_r(v) = (\Phi_{r,1}(v) \dots \Phi_{r,r}(v))$  et de  $\Sigma_{u_r} = \text{Var}(u_{r,n})$  sont définis par

$$\hat{\underline{\Phi}}_r(v) = \hat{\Sigma}_{\hat{Y}, \hat{Y}_r}(v) \hat{\Sigma}_{\hat{Y}_r}^{-1}(v) \quad \text{and} \quad \hat{\Sigma}_{\hat{u}_r}(v) = \frac{1}{N} \sum_{n=0}^{N-1} \left( \hat{Y}_n - \hat{\underline{\Phi}}_r(v) \hat{Y}_{r,n} \right) \left( \hat{Y}_n - \hat{\underline{\Phi}}_r(v) \hat{Y}_{r,n} \right)' \quad (1.77)$$

où

$$\hat{Y}_{r,n} = (\hat{Y}'_{n-1} \dots \hat{Y}'_{n-r})', \quad \hat{\Sigma}_{\hat{Y}, \hat{Y}_r}(v) = \frac{1}{N} \sum_{n=0}^{N-1} \hat{Y}_r \hat{Y}'_{r,n}, \quad \hat{\Sigma}_{\hat{Y}_r}^{-1}(v) = \frac{1}{N} \sum_{n=0}^{N-1} \hat{Y}_{r,n} \hat{Y}'_{r,n}$$

Par convention  $\hat{Y}_n = 0$  quand  $n \leq 0$ , et nous supposons que  $\hat{\Sigma}_{\hat{Y}_r}(v)$  est inversible.

**Theorem 1.6.3** *Sous certaines hypothèses, nous supposons que le processus  $Y_n$  défini dans (1.72) admet une représentation périodique VAR( $\infty$ ) comme dans (1.73), où  $\|\Phi_i\| = o(i^{-2})$  quand  $i \rightarrow +\infty$ , les racines de  $\det(\Phi(z)) = 0$  sont à l'extérieur de disque unité, et  $\Sigma_u(v) := \text{Var}(u_n)$  est non singulière. De plus nous supposons que  $\mathbb{E} |\epsilon_t|^{8+4v} < \infty$ . Alors l'estimateur spectral*

$$\hat{I}_N^{\text{SP}} := \hat{\Phi}_{r,v}^{-1}(1) \hat{\Sigma}_{\hat{u}_r}(v) \hat{\Phi}_{r,v}^{-1}(1) \xrightarrow{N \rightarrow \infty} I(\theta_0) = \Phi_v^{-1}(1) \Sigma_u(v) \Phi_v^{-1}(1)'$$

en probabilité quand  $r = r(N) \rightarrow \infty$  et  $r^3/3 \rightarrow 0$  quand  $N \rightarrow \infty$

### 1.6.5 La distribution asymptotique des autocovariances et autocorrélations résiduelles

Nous utilisons les mêmes notions que la partie 1.5.2 et la méthodologie pour la validation du modèle SPARMA est pratiquement similaire à celle des modèles PARMA vue dans la partie 1.5.2. La seule différence est l'expression du résidus que nous définirons par la suite. Nous énoncerons dans la suite que les propositions et les théorèmes.

Soit  $\hat{\epsilon}_{nT+v} = \hat{e}_{nT+v} = e_{nT+v}(\hat{\alpha}_{\text{QLS}})$  les résidus périodiques de moindres carrés quasi-généralisé quand  $p + q + P + Q > 0$ . Nous utilisons (1.66) pour remarquer que nous avons  $\hat{e}_t = 0$  pour  $t := nT + v \leq 0$  et  $t > NT$ . De (1.14), pour  $n = 0, 1, \dots, N-1$ , nous déduisons que

$$\begin{aligned} \hat{e}_{nT+v} = & X_{nT+v} - \sum_{i=1}^p \hat{a}_i X_{nT+v-i} - \sum_{j=1}^P \hat{\mathbf{a}}_j X_{nT+v-Tj} + \sum_{i=1}^p \sum_{j=1}^P \hat{a}_i \hat{\mathbf{a}}_j X_{nT+v-Tj-i} + \sum_{i=1}^q \hat{b}_i e_{nT+v-i} \\ & + \sum_{j=1}^Q \hat{\mathbf{b}}_j \hat{e}_{nT+v-Tj} - \sum_{i=1}^q \sum_{j=1}^Q \hat{b}_i \hat{\mathbf{b}}_j \hat{e}_{nT+v-Tj-i}, \end{aligned}$$

pour  $1 \leq t \leq NT$ , avec  $\hat{X}_t = 0$  pour  $t \leq 0$  et  $\hat{X}_t = X_t$  pour  $t \geq 1$ .

Les statistiques (1.48) (resp. (1.47)) sont généralement utilisés pour tester l'hypothèse nulle suivante

( $\mathcal{H}0$ ) :  $(X_{nT+v})$  satisfait une représentation  $\text{SPARMA}_T(p; q)(P, Q)$  (resp. pour une période  $v$  donnée);

contre l'alternative

( $\mathcal{H}1$ ) :  $(X_{nT+v})$  n'admet pas une représentation  $\text{SPARMA}_T(p; q)(P, Q)$  (resp. pour une période  $v$  donnée) ou  $(X_{nT+v})$  satisfait une représentation  $\text{PARMA}_T(p'; q')(P', Q')$  (resp. pour une période  $v$  donnée) avec  $p' > p$  ou  $q' > q$  ou  $P' > P$  ou  $Q' > Q$ .

Pour les modèles  $\text{SPARMA}_T(p; q)(P, Q)$  faibles, nous montrons que les distributions asymptotiques des statistiques définies dans (1.47) et (1.48) ne sont plus des distributions de chi-deux mais un mélange de distribution chi-deux pondéré par les valeurs propres de la matrice de covariance asymptotique du vecteur d'autocorrélation.

**Proposition 1.6.4** *Supposons  $p + q + P + Q > 0$ . Sous les hypothèses (H0''), (H1''), (H2''), (H3'') et (H4''), le vecteur aléatoire*

$$\sqrt{N} \left( (\hat{\alpha}_{\text{LS}} - \alpha_0)', \gamma'_m(v) \right)'$$

*a une distribution limite normale centré avec la matrice de covariance*

$$\Xi(v) = \begin{pmatrix} \Omega^{\text{LS}} & \Sigma_{\hat{\alpha}, \gamma_m(v)} \\ \Sigma'_{\hat{\alpha}, \gamma_m(v)} & \Sigma_{\gamma_m(v)} \end{pmatrix} = \sum_{h=-\infty}^{\infty} \mathbb{E} \left[ W_{nT+v} W'_{(n-h)T+v} \right],$$

*à partir de (1.46) et (1.53) nous avons*

$$W_{nT+v} = \begin{pmatrix} W_{1, nT+v} \\ W_{2, nT+v} \end{pmatrix} = \begin{pmatrix} -J(\alpha_0, \omega^2)^{-1} \sum_{v=1}^T \omega_v^{-2} \epsilon_{nT+v} \frac{\partial \epsilon_{nT+v}(\alpha_0)}{\partial \alpha} \\ (\epsilon_{nT+v-1}, \dots, \epsilon_{nT+v-m})' \epsilon_{nT+v} \end{pmatrix}. \quad (1.78)$$

Le théorème suivant est une extension des résultats donnés dans [33], qui fournit la distribution limite des autocovariances et autocorrélations résiduelles des modèles PARMA faibles pour une saison donnée  $v$ . Nous fournissons également la distribution limite des autocovariances et autocorrélations résiduelles globales des modèles SPARMA faibles pour toutes les saisons.

**Theorem 1.6.5** *Sous les hypothèses de la proposition 1.6.4, nous avons*

$$\sqrt{N}\hat{\gamma}_m(\mathbf{v}) \xrightarrow[N \rightarrow +\infty]{d} \mathcal{N}(0, \nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v})) \quad (1.79)$$

où

$$\begin{aligned} \nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v}) &= \Sigma_{\gamma_m(\mathbf{v})} + \Psi_m(\mathbf{v})\Omega^{\text{LS}}\Psi'_m(\mathbf{v}) + \Psi_m(\mathbf{v})\Sigma_{\hat{\alpha}, \gamma_m(\mathbf{v})} + \Sigma'_{\hat{\alpha}, \gamma_m(\mathbf{v})}\Psi'_m(\mathbf{v}) \text{ quand } p + q > 0 \\ \text{et } \nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v}) &= \Sigma_{\gamma_m(\mathbf{v})} \text{ quand } p = q = P = Q = 0. \end{aligned}$$

Alors :

$$\sqrt{N}\hat{\gamma}_{mT} \xrightarrow[N \rightarrow +\infty]{d} \mathcal{N}(0, \nabla_{\hat{\gamma}}) \quad (1.80)$$

où la matrice de covariance asymptotique  $\nabla_{\hat{\gamma}}$  est une matrice en bloc, avec les variances asymptotiques données par  $\nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v})$ ,  $\mathbf{v} = 1, \dots, T$  et les covariances asymptotiques données par

$$\nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v}') = \Sigma_{\gamma_m}(\mathbf{v}, \mathbf{v}') + \Psi_m(\mathbf{v})\Omega^{\text{LS}}\Psi'_m(\mathbf{v}') + \Psi_m(\mathbf{v})\Sigma_{\hat{\alpha}, \gamma_m}(\mathbf{v}, \mathbf{v}') + \Sigma'_{\hat{\alpha}, \gamma_m(\mathbf{v})}\Psi'_m(\mathbf{v}'), \quad \mathbf{v} \neq \mathbf{v}'.$$

Nous avons également

$$\sqrt{N}\hat{\rho}_m(\mathbf{v}) \xrightarrow[N \rightarrow +\infty]{d} \mathcal{N}(0, \nabla_{\hat{\rho}}(\mathbf{v}, \mathbf{v})) \quad \text{où } \nabla_{\hat{\rho}}(\mathbf{v}, \mathbf{v}) = D_m^{-1}(\mathbf{v})\nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v})D_m^{-1}(\mathbf{v}) \quad (1.81)$$

et

$$\sqrt{N}\hat{\rho}_{mT} \xrightarrow[N \rightarrow +\infty]{d} \mathcal{N}(0, \nabla_{\hat{\rho}}) \quad \text{où } \nabla_{\hat{\rho}} = D_{mT}^{-1}\nabla_{\hat{\gamma}}D_{mT}^{-1}. \quad (1.82)$$

Les matrices de variance asymptotique  $\nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v})$  et  $\nabla_{\hat{\rho}}(\mathbf{v}, \mathbf{v})$  dépendent des matrices inconnues  $\Xi(\mathbf{v})$ ,  $\Psi_m(\mathbf{v})$  et du scalaire  $\sigma_{\mathbf{v}-h}$  pour  $h = 1, \dots, m$ . La matrice  $\Psi_m(\mathbf{v})$  et la variance du bruit  $\sigma_{\mathbf{v}-h}$  peuvent être respectivement estimées empiriquement par

$$\hat{\Psi}_m(\mathbf{v}) = \frac{1}{N} \sum_{n=0}^{N-1} \left\{ (\hat{\epsilon}_{nT+\mathbf{v}-1}, \dots, \hat{\epsilon}_{nT+\mathbf{v}-m})' \frac{\partial \hat{\epsilon}_{nT+\mathbf{v}}}{\partial \alpha'} \right\} \text{ et } \hat{\sigma}_{\mathbf{v}-h} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} \hat{\epsilon}_{nT+\mathbf{v}-h}^2}$$

### 1.6.6 Version modifiée du test portmanteau

Nous énonçons le théorème ci-dessus, qui nous donne une estimation de la matrice  $\Xi$ . C'est une approche basée sur l'estimation de la densité spectrale de  $(W_{nT+\mathbf{v}})_{n \in \mathbb{Z}}$  par VAR paramétrique définie en 1.5.3.

**Theorem 1.6.6** *Sous les hypothèses du théorème 1.6.5, nous supposons que le processus  $(W_{nT+\mathbf{v}})$  défini dans (1.54) admet une représentation périodique VAR( $\infty$ ) dont les racines de  $\det \Phi(z) = 0$  sont à l'extérieur du disque unité,  $\|\Phi_i\| = o(i^{-2})$ , et que la matrice  $\Sigma_u(\mathbf{v}) = \text{Var}(u_{nT+\mathbf{v}})$  est non-singulière. De plus nous supposons que  $\mathbb{E}\|\epsilon_n\|^{8+4\kappa} < \infty$  et  $\sum_{k=0}^{\infty} \{\alpha_{\epsilon}(k)\}^{\kappa/(2+\kappa)} < \infty$  pour un réel  $\kappa > 0$ . Alors l'estimateur de  $\Xi(\mathbf{v})$*

$$\hat{\Xi}^{\text{SP}}(\mathbf{v}) := \hat{\Phi}_{\mathbf{v},r}^{-1}(1)\hat{\Sigma}_{\hat{u}_r}(\mathbf{v})\hat{\Phi}_{\mathbf{v},r}^{-1'}(1) \rightarrow \Xi(\mathbf{v})$$

en probabilité quand  $r = r(N) \rightarrow +\infty$  et  $r^3/N \rightarrow 0$  quand  $N \rightarrow +\infty$ .

Du théorème 1.6.5 nous pouvons déduire le résultat suivant qui donne la distribution limite exacte des statistiques standards de portmanteau (1.47) sous des hypothèses générales sur le processus d'innovation du modèle PARMA( $p, q$ )(P, Q) ajusté.

**Theorem 1.6.7** *Sous les hypothèses du théorème 1.6.5 et  $(\mathcal{H}0)$ , les statistiques  $Q_m^{\text{LB}}(\mathbf{v})$  et  $Q_m^{\text{LBM}}(\mathbf{v})$  définies dans (1.47) convergent en distribution, quand  $N \rightarrow +\infty$ , vers*

$$Z_m^{\mathbf{v}}(\xi_m^{\mathbf{v}}) = \sum_{i=1}^m \xi_{i,m}^{\mathbf{v}} Z_i^2$$

où  $\xi_m^{\mathbf{v}} = (\xi_{1,m}^{\mathbf{v}}, \dots, \xi_{m,m}^{\mathbf{v}})'$  est le vecteur de valeurs propres de la matrice  $\nabla_{\hat{\rho}}(\mathbf{v}, \mathbf{v})$  définie dans (1.81) et  $Z_1, \dots, Z_m$  sont des variables indépendantes de la loi  $\mathcal{N}(0, 1)$ .

La distribution asymptotiques des statistiques globales du test portmanteau (qui prend en compte toutes les saisons)  $Q_m^{\text{LB}}$  et  $Q_m^{\text{LBM}}$  définie dans (1.48) est également une somme pondéré de variables aléatoires de chi-deux :

$$Z_m(\xi_m) = \sum_{i=1}^{mT} \xi_{i,mT} Z_i^2$$

où  $\xi_{mT} = (\xi_{1,mT}, \dots, \xi_{mT,mT}^{\mathbf{v}})'$  est le vecteur de valeurs propres de la matrice  $\nabla_{\hat{\rho}}$  donnée (1.82) et  $Z_1, \dots, Z_{mT}$  sont également des variables indépendantes de la loi  $\mathcal{N}(0, 1)$ .

Au vue du théorème 1.6.7, les approximations  $\chi_{m-(k_0)}^2$  et  $\chi_{mT-(k_0)T}^2$  standard ne sont plus valables pour approximer la distribution asymptotique de (1.47) et (1.48) dans le cadre des modèles PARMA( $p; q$ )(P, Q) faibles.

La distribution limite dépend du paramètre de nuisance périodique  $\sigma_{\mathbf{v}-h}$  pour  $h = 1, \dots, m$  et les éléments de  $\Xi(\mathbf{v})$ . Par conséquent, afin d'obtenir la distribution asymptotique des statistiques de portmanteau (1.47) (resp. (1.48)) sous des hypothèses faibles sur le bruit, il faut un estimateur convergent de la matrice de covariance asymptotique  $\nabla_{\hat{\rho}}(\mathbf{v}, \mathbf{v})$  (resp.  $\nabla_{\hat{\rho}}$ ).

Notons par  $\hat{\nabla}_{\hat{\rho}}(\mathbf{v}, \mathbf{v})$  (resp.  $\hat{\nabla}_{\hat{\rho}}$ ) la matrice obtenue en remplaçant  $\Xi(\mathbf{v})$  par  $\hat{\Xi}(\mathbf{v})$ ,  $\Psi_m(\mathbf{v})$  par  $\hat{\Psi}_m(\mathbf{v})$  et  $\sigma_{\mathbf{v}-h}$  par  $\hat{\sigma}_{\mathbf{v}-h}$  dans  $\nabla_{\hat{\rho}}(\mathbf{v}, \mathbf{v})$  (resp.  $\nabla_{\hat{\rho}}$ ). Notons  $\hat{\xi}_m^{\mathbf{v}} = (\hat{\xi}_{1,m}^{\mathbf{v}}, \dots, \hat{\xi}_{m,m}^{\mathbf{v}})'$  (resp.  $\hat{\xi}_m = (\hat{\xi}_{1,mT}, \dots, \hat{\xi}_{mT,mT})'$ ) le vecteur de valeurs propres de  $\hat{\nabla}_{\hat{\rho}}(\mathbf{v}, \mathbf{v})$  (resp.  $\hat{\nabla}_{\hat{\rho}}$ ).

Au seuil asymptotique  $\alpha$  et pour une saison  $\mathbf{v}$  donnée, le test LB (resp. le test BP) consiste à rejeter l'adéquation de modèle PARMA( $p; q$ ) faible quand

$$\lim_{N \rightarrow \infty} \mathbb{P}(Q_m^{\text{BP}}(\mathbf{v}) > S_m^{\mathbf{v}}(1 - \alpha)) = \lim_{N \rightarrow \infty} \mathbb{P}(Q_m^{\text{LBM}}(\mathbf{v}) > S_m^{\mathbf{v}}(1 - \alpha)) = \alpha,$$

où  $S_m^{\mathbf{v}}(1 - \alpha)$  est tel que  $\mathbb{P}\{Z_m^{\mathbf{v}}(\hat{\xi}_m^{\mathbf{v}}) > S_m^{\mathbf{v}}(1 - \alpha)\} = \alpha$ .

## 1.6.7 La distribution asymptotique auto-normalisée des autocorrélations résiduelles

Dans cette section nous proposons également la méthode d'autonormalisation définie dans la partie 1.5.4 et on garde les mêmes notations.

Notons par  $\Lambda(\mathbf{v})$  la matrice par blocs de  $\mathbb{R}^{m \times (\Gamma(p+q+p+Q)+m)}$  définis par  $\Lambda(\mathbf{v}) = (\Psi_m(\mathbf{v})|I_m)$  où  $I_k$  est la matrice d'identité d'ordre  $k$ .

La proposition suivante donne l'inversibilité des matrices de normalisation  $C_m(\mathbf{v}, \mathbf{v})$  et  $C_m$  définies en 1.62.

**Proposition 1.6.8** *Sous les hypothèses du théorème 1.6.5 et  $(\mathbf{H5}'')$ , les matrices  $C_m(\mathbf{v}, \mathbf{v})$  et  $C_m$  sont presque sûrement non singulière.*

Le théorème suivant énonce la distribution asymptotiques des autocovariances et autocorrélations résiduelles.



**Theorem 1.6.9** *Nous supposons que  $p + q > 0$ . Sous les hypothèses du théorème 1.6.5, (H5'') et sous l'hypothèse nulle  $\mathcal{H}0$ , pour une saison  $\nu$  donnée nous avons*

$$N\hat{\gamma}'_m(\nu)C_m^{-1}(\nu, \nu)\hat{\gamma}_m(\nu) \xrightarrow[N \rightarrow +\infty]{d} \mathcal{U}_m \text{ et } N\hat{\rho}'_m(\nu)D_m(\nu)C_m^{-1}(\nu, \nu)D_m(\nu)\hat{\rho}_m(\nu) \xrightarrow[N \rightarrow +\infty]{d} \mathcal{U}_m.$$

*Pour les autocovariances et autocorrélations résiduelles globales sur toutes les saisons, nous avons également*

$$N\hat{\gamma}'_{mT}C_{mT}^{-1}\hat{\gamma}_{mT} \xrightarrow[N \rightarrow +\infty]{d} \mathcal{U}_{mT} \text{ et } N\hat{\rho}'_{mT}D_{mT}C_{mT}^{-1}D_{mT}\hat{\rho}_{mT} \xrightarrow[N \rightarrow +\infty]{d} \mathcal{U}_{mT}.$$

Bien entendu, le théorème est inutilisable dans la pratique car la matrice de normalisation  $C_m(\nu, \nu)$ ,  $C_{mT}$  et le paramètre de nuisance  $\sigma_{\nu-h}$  ne sont pas observables. Ce problème sera corrigé lorsque nous remplaçons la matrice  $C_m(\nu, \nu)$ ,  $C_{mT}$  et la variance du bruit  $\sigma_{\nu-h}$  par leurs parties empiriques ou observables. Ainsi nous notons

$$\hat{C}_m(\nu, \nu) = \frac{1}{N^2} \sum_{n=0}^{N-1} \hat{S}_n(\nu)\hat{S}'_n(\nu) \text{ où } \hat{S}_n(\nu) = \sum_{j=0}^n (\hat{\Lambda}(\nu)\hat{W}_{jT+\nu} - \hat{\gamma}_m(\nu)),$$

avec  $\hat{\Lambda}(\nu) = (\hat{\Psi}_m(\nu)|I_m)$  et où  $\hat{W}_{nT+\nu}$  et  $\hat{\sigma}_{\nu-h}$  sont définies dans 1.5.2.

Nous sommes donc en mesure d'énoncer la version applicable du théorème 1.6.9.

**Theorem 1.6.1** *Sous les hypothèses du théorème 1.6.9, pour une saison  $\nu$  donnée nous avons*

$$N\hat{\gamma}'_m(\nu)\hat{C}_m^{-1}(\nu, \nu)\hat{\gamma}_m(\nu) \xrightarrow[N \rightarrow +\infty]{d} \mathcal{U}_m \text{ et } Q_m^{\text{SN}}(\nu) =: N\hat{\rho}'_m(\nu)\hat{D}_m(\nu)\hat{C}_m^{-1}(\nu, \nu)\hat{D}_m(\nu)\hat{\rho}_m(\nu) \xrightarrow[N \rightarrow +\infty]{d} \mathcal{U}_m.$$

*Nous avons également pour toutes les saisons*

$$N\hat{\gamma}'_{mT}\hat{C}_{mT}^{-1}\hat{\gamma}_{mT} \xrightarrow[N \rightarrow +\infty]{d} \mathcal{U}_{mT} \text{ et } Q_m^{\text{SN}} =: N\hat{\rho}'_{mT}\hat{D}_{mT}\hat{C}_{mT}^{-1}\hat{D}_{mT}\hat{\rho}_{mT} \xrightarrow[N \rightarrow +\infty]{d} \mathcal{U}_{mT}.$$

Sur la base du résultat ci-dessus, nous proposons la version modifiée de la statistique de LBM pour une saison  $\nu$  donnée, lorsque l'on utilise la statistique

$$\tilde{Q}_m^{\text{SN}}(\nu) = N\hat{\rho}'_m(\nu)M_m^{1/2}(\nu, N, T)\hat{D}_m(\nu)\hat{C}_m^{-1}(\nu, \nu)\hat{D}_m(\nu)M_m^{1/2}(\nu, N, T)\hat{\rho}_m(\nu), \quad (1.83)$$

où  $M_m(\nu, N, T) \in \mathbb{R}^{m \times m}$  est une matrice diagonale avec  $(\varrho(1, \nu, N, T), \dots, \varrho(m, \nu, N, T))$  sur la diagonale.

De même, nous proposons une version modifiée de la statistique LBM pour toutes les saisons lorsque l'on utilise la statistique suivante

$$\tilde{Q}_m^{\text{SN}} = N\hat{\rho}'_{mT}M_{mT}^{1/2}\hat{D}_{mT}\hat{C}_{mT}^{-1}\hat{D}_{mT}M_{mT}^{1/2}\hat{\rho}_{mT}, \quad (1.84)$$

où  $M_{mT} \in \mathbb{R}^{mT \times mT}$  est une matrice diagonale avec

$$(\varrho(1, 1, N, T), \dots, \varrho(m, 1, N, T), \dots, \varrho(1, T, N, T), \dots, \varrho(m, T, N, T))$$

sur la diagonale.

Au seuil asymptotique  $\alpha$  et pour une saison  $\nu$  donnée, le test d'auto-normalisation LBM (resp. le test BP) consiste à rejeter l'adéquation du modèle  $\text{PARMA}_T(p; q)(P, Q)$  faible quand

$$\tilde{Q}_m^{\text{SN}}(\nu) > \mathcal{U}_m(1 - \alpha) \quad (\text{resp. } Q_m^{\text{SN}}(\nu) > \mathcal{U}_m(1 - \alpha)),$$

où les valeurs critiques  $\mathcal{U}_m(1 - \alpha)$  sont tabulées dans [54, Table 1]. De la même manière nous avons pour le test global

$$\tilde{Q}_m^{\text{SN}} > \mathcal{U}_{mT}(1 - \alpha) \quad (\text{resp. } Q_m^{\text{SN}} > \mathcal{U}_{mT}(1 - \alpha)).$$



## Bibliographie

- [1] Ahn, S. K. (1988). Distribution for residual autocovariances in multivariate autoregressive models with structured parameterization. *Biometrika*, 75(3) :590–593. [28](#)
- [2] Aknouche, A. and Bibi, A. (2009). Quasi-maximum likelihood estimation of periodic garch and periodic arma-garch processes. *Journal of Time Series Analysis*, 30(1) :19–46. [11](#), [14](#)
- [3] Akutowicz, E. J. (1957). On an explicit formula in linear least squares prediction. *Math. Scand.*, 5 :261–266. [37](#), [43](#)
- [4] Andrews, B., Davis, R. A., and Breidt, F. J. (2006). Maximum likelihood estimation for all-pass time series models. *Journal of Multivariate Analysis*, 97(7) :1638–1659. [6](#)
- [5] Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3) :817–858. [21](#), [30](#)
- [6] Basawa, I. V. and Lund, R. (2001). Large sample properties of parameter estimates for periodic ARMA models. *J. Time Ser. Anal.*, 22(6) :651–663. [13](#), [32](#), [41](#)
- [7] Basawa, I. V., Lund, R., and Shao, Q. (2004). First-order seasonal autoregressive processes with periodically varying parameters. *Statist. Probab. Lett.*, 67(4) :299–306. [16](#)
- [8] Battaglia, F., Cucina, D., and Rizzo, M. (2018). A generalization of periodic autoregressive models for seasonal time series. Technical report, Tech. Rep. 2, Department of Statistical Sciences, University La Sapienza . . . . [13](#)
- [9] Battaglia, F., Cucina, D., and Rizzo, M. (2020). Parsimonious periodic autoregressive models for time series with evolving trend and seasonality. *Statistics and Computing*, 30(1) :77–91. [13](#)
- [10] Berk, K. N. (1974). Consistent autoregressive spectral estimates. *Ann. Statist.*, 2 :489–502. Collection of articles dedicated to Jerzy Neyman on his 80th birthday. [21](#), [22](#), [30](#), [37](#)
- [11] Bloomfield, P., Hurd, H. L., and Lund, R. B. (1994). Periodic correlation in stratospheric ozone data. *Journal of Time Series Analysis*, 15(2) :127–150. [10](#)
- [12] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3) :307–327. [6](#)
- [13] Bollerslev, T. and Ghysels, E. (1996). Periodic autoregressive conditional heteroscedasticity. *Journal of Business & Economic Statistics*, 14(2) :139–151. [11](#)
- [14] Boubacar Mainassara, Y. (2009). *Estimation, validation et identification des modèles ARMA faibles multivariés*. PhD thesis. [13](#)
- [15] Boubacar Mainassara, Y. (2011). Multivariate portmanteau test for structural VARMA models with uncorrelated but non-independent error terms. *J. Statist. Plann. Inference*, 141(8) :2961–2975. [13](#), [22](#), [29](#)
- [16] Boubacar Mainassara, Y. (2012). Selection of weak varma models by modified akaike’s information criteria. *Journal of Time Series Analysis*, 33(1) :121–130. [13](#)
- [17] Boubacar Mainassara, Y., Carbon, M., and Francq, C. (2012). Computing and estimating information matrices of weak ARMA models. *Comput. Statist. Data Anal.*, 56(2) :345–361. [37](#)

- [18] Boubacar Mainassara, Y. and Francq, C. (2011). Estimating structural VARMA models with uncorrelated but non-independent error terms. *J. Multivariate Anal.*, 102(3) :496–505. [13](#), [22](#), [26](#), [37](#)
- [19] Boubacar Mainassara, Y. and Ilmi Amir, A. (2019). Distribution of residual autocorrelations for multiplicative seasonal arma models with uncorrelated but nonindependent error terms. *Statistica Neerlandica*, 73(4) :454–474. [9](#)
- [20] Boubacar Mainassara, Y. and Ilmi Amir, A. (2020). Multivariate portmanteau tests for weak multiplicative seasonal VARMA models. *Statist. Papers*, 61(6) :2529–2560. [9](#), [17](#), [19](#), [21](#)
- [21] Boubacar Mainassara, Y. and Kokonendji, C. C. (2016). Modified schwarz and hannan-quinn information criteria for weak varma models. *Statistical Inference for Stochastic Processes*, 19(2) :199–217. [13](#)
- [22] Boubacar Mainassara, Y. and Sausseureau, B. (2018). Diagnostic checking in multivariate arma models with dependent errors using normalized residual autocorrelations. *J. Amer. Statist. Assoc.*, 113(524) :1813–1827. [13](#), [17](#), [22](#), [23](#), [24](#)
- [23] Box, G. E. P. and Jenkins, G. M. (1970). *Times series analysis. Forecasting and control*. Holden-Day, San Francisco, Calif.-London-Amsterdam. [7](#)
- [24] Box, G. E. P. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J. Amer. Statist. Assoc.*, 65 :1509–1526. [7](#), [20](#), [28](#), [34](#)
- [25] Brockwell, P. J. and Davis, R. A. (1991). *Time series : theory and methods*. Springer Series in Statistics. Springer-Verlag, New York, second edition. [6](#), [13](#), [30](#)
- [26] Chitturi, R. V. (1974). Distribution of residual autocorrelations in multiple autoregressive schemes. *J. Amer. Statist. Assoc.*, 69 :928–934. [28](#)
- [27] Davydov, J. A. (1968). The convergence of distributions which are generated by stationary random processes. *Teor. Verojatnost. i Primenen.*, 13 :730–737. [19](#), [21](#)
- [28] den Haan, W. J. and Levin, A. T. (1997). A practitioner’s guide to robust covariance matrix estimation. In *Robust inference*, volume 15 of *Handbook of Statist.*, pages 299–342. North-Holland, Amsterdam. [21](#), [30](#), [37](#)
- [29] Duchesne, P. (2007). On consistent testing for serial correlation in seasonal time series models. *Canad. J. Statist.*, 35(2) :193–213. [9](#)
- [30] Dufour, J.-M. and Pelletier, D. (2008). Practical methods for modelling weak varma processes : Identification, estimation and specification with a macroeconomic application. *Manuscript, McGill University*. [13](#)
- [31] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica : Journal of the Econometric Society*, pages 987–1007. [6](#)
- [32] Francq, C., Roy, R., and Saidi, A. (2011). Asymptotic properties of weighted least squares estimation in weak PARMA models. *J. Time Series Anal.*, 32(6) :699–723. [11](#), [12](#), [13](#), [14](#), [17](#), [31](#), [32](#), [33](#), [35](#), [41](#)
- [33] Francq, C., Roy, R., and Zakoïan, J.-M. (2005). Diagnostic checking in ARMA models with uncorrelated errors. *J. Amer. Statist. Assoc.*, 100(470) :532–544. [7](#), [30](#), [36](#), [37](#), [44](#)

- [34] Francq, C. and Zakoïan, J.-M. (1998). Estimating linear representations of nonlinear processes. *J. Statist. Plann. Inference*, 68(1) :145–165. [6](#), [7](#), [9](#)
- [35] Francq, C. and Zakoïan, J.-M. (2005). Recent results for linear time series models with non independent innovations. In *Statistical modeling and analysis for complex data problems*, volume 1 of *GERAD 25th Anniv. Ser.*, pages 241–265. Springer, New York. [6](#), [9](#)
- [36] Franses, P. H. and Paap, R. (2004). *Periodic time series models*. OUP Oxford. [10](#)
- [37] Gallagher, C. M. and Fisher, T. J. (2015). On weighted portmanteau tests for time-series goodness-of-fit. *Journal of Time Series Analysis*, 36(1) :67–83. [7](#)
- [38] Giovanis, E. et al. (2014). The turn-of-the-month-effect : Evidence from periodic generalized autoregressive conditional heteroskedasticity (pgarch) model. *International Journal of Economic Sciences and Applied Research*, 7(3) :43–61. [11](#)
- [39] Hannan, E. (1975). The estimation of arma models. *The Annals of Statistics*, pages 975–981. [7](#)
- [40] Hannan, E. J. (1976). The identification and parametrization of ARMAX and state space forms. *Econometrica*, 44(4) :713–723. [26](#)
- [41] Hannan, E. J. and Rissanen, J. (1982). Recursive estimation of mixed autoregressive-moving average order. *Biometrika*, 69(1) :81–94. [7](#)
- [42] Herrndorf, N. (1984). A functional central limit theorem for weakly dependent sequences of random variables. *Ann. Probab.*, 12(1) :141–153. [21](#)
- [43] Hipel, K. and McLeod, A. I. (1994). *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier, Amsterdam. [10](#), [13](#), [34](#)
- [44] Hipel, K. W., McLeod, A. I., and Lennox, W. C. (1977). Advances in box-jenkins modeling : 1. model construction. *Water Resources Research*, 13(3) :567–575. [7](#)
- [45] Holan, S. H., Lund, R., Davis, G., et al. (2010). The arma alphabet soup : A tour of arma model variants. *Statistics Surveys*, 4 :232–274. [11](#)
- [46] Hosking, J. (1981). Equivalent forms of the multivariate portmanteau statistic. *Journal of the Royal Statistical Society : Series B (Methodological)*, 43(2) :261–262. [28](#)
- [47] Hosking, J. (1989). Corrigendum : Equivalent forms of the multivariate portmanteau statistic. *Journal of the Royal Statistical Society. Series B (Methodological)*, 51(2) :303–303. [28](#)
- [48] Hosking, J. R. M. (1980). The multivariate portmanteau statistic. *J. Amer. Statist. Assoc.*, 75(371) :602–608. [28](#), [30](#)
- [49] Imhof, J. P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika*, 48 :419–426. [38](#)
- [50] Jones, R. H. and Brelsford, W. M. (1967). Time series with periodic structure. *Biometrika*, 54(3-4) :403–408. [10](#)
- [51] Kuan, C.-M. and Lee, W.-M. (2006). Robust M tests without consistent estimation of the asymptotic covariance matrix. *J. Amer. Statist. Assoc.*, 101(475) :1264–1275. [22](#)

- [52] Li, W. and McLeod, A. (1981). Distribution of the residual autocorrelations in multivariate arma time series models. *Journal of the Royal Statistical Society : Series B (Methodological)*, 43(2) :231–239. [28](#)
- [53] Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2) :pp. 297–303. [20](#), [28](#), [34](#)
- [54] Lobato, I. N. (2001). Testing that a dependent process is uncorrelated. *J. Amer. Statist. Assoc.*, 96(455) :1066–1076. [17](#), [22](#), [23](#), [39](#), [40](#), [47](#)
- [55] Lund, R. and Basawa, I. V. (2000). Recursive prediction and likelihood evaluation for periodic ARMA models. *J. Time Ser. Anal.*, 21(1) :75–93. [11](#), [13](#)
- [56] Mahdi, E. (2016). Portmanteau test statistics for seasonal serial correlation in time series models. *SpringerPlus*, 5(1) :1485. [9](#), [17](#), [24](#), [28](#)
- [57] McLeod, A. I. (1978). On the distribution of residual autocorrelations in Box-Jenkins models. *J. Roy. Statist. Soc. Ser. B*, 40(3) :296–302. [7](#), [9](#), [27](#), [28](#), [29](#)
- [58] McLeod, A. I. (1993). Parsimony, model adequacy and periodic correlation in time series forecasting. *International Statistical Review/Revue Internationale de Statistique*, pages 387–393. [14](#)
- [59] McLeod, A. I. (1994). Diagnostic checking of periodic autoregression models with application. *J. Time Ser. Anal.*, 15(2) :221–233. [34](#)
- [60] McLeod, A. I. (1995). Addendum : Diagnostic checking of periodic autoregression models with application [J. Time Ser. Anal. **15** (1994), no. 2, 221–233]. *J. Time Ser. Anal.*, 16(6) :647–648. [34](#)
- [61] McLeod, A. I. and Li, W. K. (1983). Diagnostic checking arma time series models using squared-residual autocorrelations. *Journal of time series analysis*, 4(4) :269–273. [7](#)
- [62] Morgan, J. and Tatar, J. (1972). Calculation of the residual sum of squares for all possible regressions. *Technometrics*, 14(2) :317–325. [13](#)
- [63] Newey, W. K. and West, K. D. (1987). A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3) :703–708. [21](#), [30](#)
- [64] Noakes, D. J., McLeod, A. I., and Hipel, K. W. (1985). Forecasting monthly riverflow time series. *International Journal of Forecasting*, 1(2) :179–190. [13](#)
- [65] Osborn, D. R. and Smith, J. P. (1989). The performance of periodic autoregressive models in forecasting seasonal uk consumption. *Journal of Business & Economic Statistics*, 7(1) :117–127. [10](#)
- [66] Pagano, M. (1978). On periodic and multiple autoregressions. *Ann. Statist.*, 6(6) :1310–1317. [10](#)
- [67] Parzen, E. and Pagano, M. (1979). An approach to modeling seasonally stationary time series. *Journal of Econometrics*, 9(1-2) :137–153. [10](#)
- [68] Romano, J. P. and Thombs, L. A. (1996). Inference for autocorrelations under weak assumptions. *J. Amer. Statist. Assoc.*, 91(434) :590–600. [6](#), [7](#), [18](#)

- [69] Salas, J. and Obeysekera, J. (1982). Arma model identification of hydrologic time series. *Water Resources Research*, 18(4) :1011–1021. [7](#)
- [70] Salas, J. D. (1980). *Applied modeling of hydrologic time series*. Water Resources Publication. [13](#)
- [71] Salas, J. D., Boes, D. C., and Smith, R. A. (1982). Estimation of arma models with seasonal parameters. *Water Resources Research*, 18(4) :1006–1010. [13](#)
- [72] Shao, Q. and Lund, R. (2004). Computation and characterization of autocorrelations and partial autocorrelations in periodic ARMA models. *J. Time Ser. Anal.*, 25(3) :359–372. [34](#)
- [73] Shao, X. (2010a). A self-normalized approach to confidence interval construction in time series. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 72(3) :343–366. [22](#)
- [74] Shao, X. (2010b). Corrigendum : A self-normalized approach to confidence interval construction in time series. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 72(5) :695–696. [22](#)
- [75] Shao, X. (2012). Parametric inference in stationary time series models with dependent errors. *Scand. J. Stat.*, 39(4) :772–783. [22](#)
- [76] Slutsky, E. (1937). The summation of random causes as the source of cyclic processes. *Econometrica : Journal of the Econometric Society*, pages 105–146. [5](#)
- [77] Thompstone, R. M., Hipel, K. W., and McLeod, A. I. (1985). Grouping of periodic autoregressive models. *Time series analysis : Theory and practice*, 6 :35–49. [13](#)
- [78] Tiao, G. and Grupe, M. (1980). Hidden periodic autoregressive-moving average models in time series data. *Biometrika*, 67(2) :365–373. [10](#), [13](#)
- [79] Tiao, G. C. and Tsay, R. S. (1983). Consistency properties of least squares estimates of autoregressive parameters in arma models. *The Annals of Statistics*, pages 856–871. [7](#)
- [80] Ursu, E. and Duchesne, P. (2009a). Estimation and model adequacy checking for multivariate seasonal autoregressive time series models with periodically varying parameters. *Stat. Neerl.*, 63(2) :183–212. [12](#), [16](#)
- [81] Ursu, E. and Duchesne, P. (2009b). On multiplicative seasonal modelling for vector time series. *Statist. Probab. Lett.*, 79(19) :2045–2052. [9](#), [27](#), [29](#)
- [82] Ursu, E. and Pereau, J.-C. (2016). Application of periodic autoregressive process to the modeling of the garonne river flows. *Stochastic Environmental Research and Risk Assessment*, 30(7) :1785–1795. [13](#)
- [83] Vecchia, A. (1985a). Maximum likelihood estimation for periodic autoregressive moving average models. *Technometrics*, 27(4) :375–384. [13](#)
- [84] Vecchia, A. (1985b). Periodic autoregressive-moving average (parma) modeling with applications to water resources 1. *JAWRA Journal of the American Water Resources Association*, 21(5) :721–730. [10](#)
- [85] Wold, H. (1938). *A study in the analysis of stationary time series*. PhD thesis, Almqvist & Wiksell. [6](#)
- [86] Yule, G. U. (1927). Vii. on a method of investigating periodicities disturbed series, with special reference to wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 226(636-646) :267–298. [5](#)

# Distribution of residual autocorrelations for multiplicative seasonal ARMA models with uncorrelated but non-independent error terms

## Abstract

In this paper we consider portmanteau tests for testing the adequacy of multiplicative seasonal autoregressive moving-average (SARMA) models under the assumption that the errors are uncorrelated but not necessarily independent. We relax the standard independence assumption on the error term in order to extend the range of application of the SARMA models. We study the asymptotic distributions of residual and normalized residual empirical autocovariances and autocorrelations under weak assumptions on the noise. We establish the asymptotic behaviour of the proposed statistics. A set of Monte Carlo experiments and an application to monthly mean total sunspot number are presented.

**Keywords:** Goodness-of-fit test, quasi-maximum likelihood estimation, Box-Pierce and Ljung-Box portmanteau tests, residual autocorrelation, self-normalization, weak SARMA models

## 2.1 Introduction

The multiplicative seasonal autoregressive moving average (SARMA, for short) model of order  $(p, q)(P, Q)_s$  for the univariate time series  $X = (X_t)_{t \in \mathbb{Z}}$ , is defined by

$$a_{\theta_0}(L)\mathbf{a}_{\theta_0}(L)X_t = b_{\theta_0}(L)\mathbf{b}_{\theta_0}(L)\epsilon_t, \quad \forall t \in \mathbb{Z}, \quad (2.1)$$

where  $\theta_0 = (a_{01}, \dots, a_{0p}, b_{01}, \dots, b_{0q}, \mathbf{a}_{01}, \dots, \mathbf{a}_{0P}, \mathbf{b}_{01}, \dots, \mathbf{b}_{0Q})'$  and  $L$  stands for the backshift operator. The nonseasonal autoregressive (AR, for short) and moving average (MA, for short) operators are defined by  $a_{\theta_0}(L) = 1 - \sum_{i=1}^p a_{0i}L^i$  and  $b_{\theta_0}(L) = 1 - \sum_{i=1}^q b_{0i}L^i$ , respectively, while the seasonal AR and MA operators are given by  $\mathbf{a}_{\theta_0}(L) = 1 - \sum_{i=1}^P \mathbf{a}_{0i}L^{si}$  and  $\mathbf{b}_{\theta_0}(L) = 1 - \sum_{i=1}^Q \mathbf{b}_{0i}L^{si}$ , respectively, where  $s$  denotes the length of the seasonal period. It is assumed that the model defined by (2.1) is stationary, invertible and not redundant. Without loss of generality, we also assume that  $a_{0p}^2 + b_{0q}^2 + \mathbf{a}_{0P}^2 + \mathbf{b}_{0Q}^2 \neq 0$  (by convention  $a_{00} = b_{00} = \mathbf{a}_{00} = \mathbf{b}_{00} = 1$ ).

In the standard situation  $\epsilon = (\epsilon_t)_{t \in \mathbb{Z}}$  is assumed to be a sequence of independent and identically distributed (iid for short) random variables with zero mean and common variance. In this standard framework,  $(\epsilon_t)$  is said to be a *strong white noise* and the representation (2.1) is called a



strong SARMA( $p, q$ )( $P, Q$ ) $_s$  process. In contrast with this previous definition, the representation (2.1) is said to be a weak SARMA( $p, q$ )( $P, Q$ ) $_s$  if the noise process ( $\epsilon_t$ ) is a *weak white noise*, that is, if it satisfies

$$(A0): \quad \mathbb{E}(\epsilon_t) = 0, \text{Var}(\epsilon_t) = \sigma_0^2 \text{ and } \text{Cov}(\epsilon_t, \epsilon_{t-h}) = 0 \text{ for all } t \in \mathbb{Z} \text{ and all } h \neq 0.$$

A strong white noise is obviously a weak white noise, because independence entails uncorrelatedness, but the converse is not true. It is clear from these definitions that the following inclusions hold:

$$\{\text{strong SARMA}(p, q)(P, Q)_s\} \subset \{\text{weak SARMA}(p, q)(P, Q)_s\}.$$

After estimating the SARMA process, the next important step in the modeling consists in checking if the estimated model fits satisfactorily the data. Thus, under the null hypothesis that the model has been correctly identified, the residuals ( $\hat{\epsilon}_t$ ) are approximately a white noise. This adequacy checking step validates or invalidates the choice of the orders ( $p, q$ ) and ( $P, Q$ ) $_s$ .

Based on the residual empirical autocorrelations  $\hat{\rho}(h) = \sum_{t=1+h}^n \hat{\epsilon}_t \hat{\epsilon}_{t-h} / \sum_{t=1}^n \hat{\epsilon}_t^2$ , where  $n$  is the length of the series, [9] have proposed a goodness-of-fit test, the so-called portmanteau test, for strong ARMA models. A modification of their test has been proposed by [22]. It is nowadays one of the most popular diagnostic checking tools in ARMA modeling of time series. These tests are defined by

$$Q_m^{\text{BP}} = n \sum_{h=1}^m \hat{\rho}^2(h) \text{ and } Q_m^{\text{LB}} = n(n+2) \sum_{h=1}^m \frac{\hat{\rho}^2(h)}{n-h}, \quad (2.2)$$

where  $m$  is a fixed integer. The statistic  $Q_m^{\text{LB}}$  has the same asymptotic chi-squared distribution as  $Q_m^{\text{BP}}$  and has the reputation of doing better for small or medium sized sample (see [22]). For weak ARMA models, [14] show that the asymptotic distributions of the statistics defined in (2.2) are no longer chi-square distributions but a mixture of chi-squared distributions, weighted by eigenvalues of the asymptotic covariance matrix of the vector of autocorrelations. Recently, [8] proposed an alternative method based on a self-normalization approach to construct a new test statistic which is asymptotically distribution-free under the null hypothesis.

In many situations, these tests are implemented to check the lack of fit of SARMA models. However, the traditional methodology of Box and Jenkins cannot be extended to the case of SARMA models when  $s > 1$  because of the multiplicative constraints on the parameters. This standard methodology needs to be adapted to take into account the possible lack of independence of the error terms. See for instance [13] and [24] who considered serial correlation testing in multiplicative seasonal univariate time series models. Duchesne (see [13]) proposed his test statistic based on a kernel-based spectral density estimator, whose weighting scheme is more adapted to autocorrelations associated to seasonal lags. The standard test procedures consist in rejecting the null hypothesis of a SARMA( $p, q$ )( $P, Q$ ) $_s$  model if the statistics (2.2) are larger than a certain quantile of a chi-squared distribution with  $m - (p + q + P + Q) > 0$  degrees of freedom. Consequently, these standard tests are not applicable for  $m \leq p + q + P + Q$ .

The works on the portmanteau statistic of SARMA models are generally performed under the assumption that the errors  $\epsilon_t$  are independent. This independence assumption is often considered too restrictive by practitioners. It precludes conditional heteroscedasticity and/or other forms of nonlinearity (see [18], for a review on weak univariate ARMA models). In this framework, we relax the standard independence assumption on the error term in order to be able to cover SARMA representations of general nonlinear models. For the asymptotic theory of weak SARMA, notable exception are [6] where the consistency and the asymptotic normality of the quasi-maximum likelihood estimator (QMLE) for weak multivariate SARMA models are studied. They also study a particular case of the asymptotic distributions of residual autocovariances and autocorrelations at the seasonal lags  $1s, 2s, 3s, \dots, ms$ .

This paper is devoted to the problem of the validation step of weak SARMA representations. We consider portmanteau test statistics based on the residual empirical autocorrelations but not necessarily at multiple lags  $s$  as in [6]. For such models, we show that the asymptotic distributions of the statistics defined in (2.2) are no longer chi-square distributions but a mixture of chi-squared distributions, weighted by eigenvalues of the asymptotic covariance matrix of the vector of autocorrelations as in [6]. We also proposed another modified statistics based on a self-normalization approach which are asymptotically distribution-free under the null hypothesis and generalize the result of [8].

In Monte Carlo experiments, we illustrate that the proposed test statistics have reasonable finite sample performance. Under nonindependent errors, it appears that the standard test statistics are generally non reliable, overrejecting severely, while the proposed tests statistics offer satisfactory levels. Even for independent errors, they seem preferable to the standard ones, when the number  $m$  of autocorrelations is small. Moreover, the error of first kind is well controlled. Contrarily to the standard tests (2.2), the proposed tests can be used safely for  $m$  small (see for instance Figure 2.1). For all these above reasons, we think that the modified versions that we propose in this paper are preferable to the standard ones for diagnosing SARMA models under nonindependent errors. Other contribution is to improve the results concerning the statistical analysis of weak SARMA models by considering the adequacy problem.

The article is organized as follows. In the next section, we briefly recall the results on the QMLE asymptotic distribution obtained by [6] when  $(\epsilon_t)$  satisfies mild mixing assumptions. We study the asymptotic behaviour of the residuals autocovariances and autocorrelations under weak assumptions on the noise in Section 2.3.1. It is also shown how the standard portmanteau tests (2.2) must be adapted in the case of multiplicative seasonal ARMA models with non-independent innovations. In Section 2.3.2 we derive the asymptotic distribution of residuals autocovariances and autocorrelations using self-normalization approach and we establish the asymptotic behaviour of the proposed statistics. Section 2.4 proposes numerical illustrations and an illustrative application on real data. We provide a conclusion in Section 2.5. The technical proofs are relegated to the appendix.

## 2.2 Estimating weak SARMA models

In this section, we recall the results on the QMLE asymptotic distribution obtained by [6] when  $(\epsilon_t)$  satisfies mild mixing assumptions in order to have a self-containing paper.

The unknown parameter of interest  $\theta_0$  is supposed to belong to the parameter space

$$\Theta = \left\{ \theta = (a_1, \dots, a_p, b_1, \dots, b_q, \mathbf{a}_1, \dots, \mathbf{a}_P, \mathbf{b}_1, \dots, \mathbf{b}_Q)' \in \mathbb{R}^{k_0}, \text{ where } k_0 = p + q + P + Q, \right.$$

$$\left. \begin{aligned} a_\theta(z) = 1 - \sum_{i=1}^p a_i z^i, \mathbf{a}_\theta(z) = 1 - \sum_{j=1}^P \mathbf{a}_j z^{s_j}, b_\theta(z) = 1 - \sum_{i=1}^q b_i z^i \text{ and } \mathbf{b}_\theta(z) = 1 - \sum_{j=1}^Q \mathbf{b}_j z^{s_j} \\ \text{have all their zeros outside the unit disk} \end{aligned} \right\}.$$

To ensure the asymptotic theory of the QMLE, we assume that the parametrization satisfies the following smoothness conditions. Without loss of generality, we may assume that  $\Theta$  is compact.

**(A1):** The process  $\epsilon = (\epsilon_t)_{t \in \mathbb{Z}}$  is ergodic and strictly stationary.

For the asymptotic normality of the QMLE, additional assumptions are required. It is necessary to assume that  $\theta_0$  is not on the boundary of the parameter space  $\Theta$ .

**(A2):** We have  $\theta_0 \in \overset{\circ}{\Theta}$ , where  $\overset{\circ}{\Theta}$  denotes the interior of  $\Theta$ .



To control the serial dependence of the stationary process  $(\epsilon_t)$ , we introduce the strong mixing coefficients  $\alpha_\epsilon(h)$  defined by

$$\alpha_\epsilon(h) = \sup_{A \in \mathcal{F}_{-\infty}^t, B \in \mathcal{F}_{t+h}^{+\infty}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|,$$

where  $\mathcal{F}_{-\infty}^t = \sigma(\epsilon_u, u \leq t)$  and  $\mathcal{F}_{t+h}^{+\infty} = \sigma(\epsilon_u, u \geq t+h)$ . We use  $|\cdot|$  to denote the Euclidian norm  $|z| = (\sum_{i=1}^r z_i^2)^{1/2}$  of a column vector  $z = (z_1, \dots, z_r)'$ . We will make an integrability assumption on the moment of the noise and a summability condition on the strong mixing coefficients  $(\alpha_\epsilon(k))_{k \geq 0}$ .

**(A3):** We have  $\mathbb{E}|\epsilon_t|^{4+2\nu} < \infty$  and  $\sum_{k=0}^{\infty} \{\alpha_\epsilon(k)\}^{\frac{\nu}{2+\nu}} < \infty$  for some  $\nu > 0$ .

Assumption **(A3)** from [15, 18] is a technical condition for proving the asymptotic theory of the QMLE. The integrability assumption on the moment of the noise is not very restrictive in this framework because the innovation process  $(\epsilon_t)$  is directly observed (see [26]).

For the estimation of SARMA and multivariate SARMA models, the commonly used estimation method is the quasi-maximum likelihood estimation, which can be also viewed as a nonlinear least squares estimation (LSE). Given a realization  $X_1, X_2, \dots, X_n$  satisfying (2.1), the variable  $\epsilon_t(\theta)$  can be approximated, for  $0 < t \leq n$ , by  $e_t(\theta)$  defined recursively by

$$\begin{aligned} e_t(\theta) = & X_t - \sum_{i=1}^p a_i X_{t-i} - \sum_{j=1}^p \mathbf{a}_j X_{t-sj} + \sum_{i=1}^p \sum_{j=1}^p a_i \mathbf{a}_j X_{t-sj-i} + \sum_{i=1}^q b_i e_{t-i}(\theta) \\ & + \sum_{j=1}^q \mathbf{b}_j e_{t-sj}(\theta) - \sum_{i=1}^q \sum_{j=1}^q b_i \mathbf{b}_j e_{t-sj-i}(\theta), \end{aligned} \quad (2.3)$$

where the unknown initial values are set to zero:  $e_0(\theta) = \dots = e_{1-q-sQ}(\theta) = X_0 = \dots = X_{1-p-sP} = 0$ . The Gaussian quasi-likelihood is given by

$$L_n(\theta, \sigma^2) = \prod_{t=1}^n \frac{1}{(2\pi)^{1/2} \sqrt{\sigma^2}} \exp \left\{ -\frac{e_t^2(\theta)}{2\sigma^2} \right\}.$$

A QMLE of  $(\theta, \sigma^2)$  is a measurable solution  $(\hat{\theta}_n, \hat{\sigma}^2)$  of

$$(\hat{\theta}_n, \hat{\sigma}^2) = \operatorname{argmin}_{\theta, \sigma^2} \left\{ \log(\sigma^2) + \frac{1}{2\sigma^2} Q_n(\theta) \right\} \text{ where } Q_n(\theta) = \frac{1}{n} \sum_{t=1}^n e_t^2(\theta).$$

In all the sequel, we denote by  $\xrightarrow{d}$ , the convergence in distribution. The symbol  $o_{\mathbb{P}}(1)$  is used for a sequence of random variables that converges to zero in probability. Under the above assumptions, [6] showed that  $\hat{\theta}_n \rightarrow \theta_0$  *a.s.* as  $n \rightarrow \infty$  and

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = -J^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^n Y_t + o_{\mathbb{P}}(1) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Sigma = J^{-1} I J^{-1}), \quad (2.4)$$

where

$$J = J(\theta_0) = \frac{2}{\sigma_0^2} \mathbb{E} \left[ \frac{\partial \epsilon_t(\theta_0)}{\partial \theta} \frac{\partial \epsilon_t(\theta_0)}{\partial \theta'} \right], \quad I = I(\theta_0) = \sum_{h=-\infty}^{+\infty} \operatorname{Cov}(Y_t, Y_{t-h}) \text{ and } Y_t = \frac{2}{\sigma_0^2} \epsilon_t(\theta_0) \frac{\partial \epsilon_t(\theta_0)}{\partial \theta}.$$

Note that, the existence of the matrix  $I(\theta_0)$  is a consequence of **(A3)** and of Davydov's inequality (see [11]).

## 2.3 Diagnostic checking in weak SARMA models

In order to check the validity of the SARMA( $p, q$ )( $P, Q$ ) $_s$  model, it is a common practice to examine the QMLE residuals  $\hat{\epsilon}_t = \hat{e}_t = e_t(\hat{\theta}_n)$  where  $e_t(\theta)$  is given by (2.3) for all  $\theta \in \mathbb{R}^{k_0}$ . For a fixed integer  $m \geq 1$ , consider the vector of residual autocovariances

$$\hat{\gamma}_m = (\hat{\gamma}(1), \dots, \hat{\gamma}(m))' \text{ where } \hat{\gamma}(h) = \frac{1}{n} \sum_{t=h+1}^n \hat{e}_t \hat{e}_{t-h} \text{ for } 0 \leq h < n.$$

In the sequel, we will also need the vector of the first  $m$  sample autocorrelations

$$\hat{\rho}_m = (\hat{\rho}(1), \dots, \hat{\rho}(m))' \text{ where } \hat{\rho}(h) = \hat{\gamma}(h)/\hat{\gamma}(0).$$

The statistics (2.2) are usually used to test the following null hypothesis

**(H0)** :  $(X_t)_{t \in \mathbb{Z}}$  satisfies a SARMA( $p, q$ )( $P, Q$ ) $_s$  representation;

against the alternative

**(H1)** :  $(X_t)_{t \in \mathbb{Z}}$  does not admit a SARMA( $p, q$ )( $P, Q$ ) $_s$  representation or  $(X_t)_{t \in \mathbb{Z}}$  satisfies a SARMA( $p', q'$ )( $P', Q'$ ) $_s$  representation with  $p' > p$  or  $q' > q$  or  $P' > P$  or  $Q' > Q$ .

### 2.3.1 Asymptotic distribution of the residual autocorrelations

First note that the mixing assumptions **(A3)** entail the asymptotic normality of the "empirical" autocovariances

$$\gamma_m = (\gamma(1), \dots, \gamma(m))' \text{ where } \gamma(h) = \frac{1}{n} \sum_{t=h+1}^n \epsilon_t \epsilon_{t-h} \text{ for } 0 \leq h < n.$$

It should be noted that  $\gamma(h)$  is not a computable statistic because it depends on the unobserved innovations  $\epsilon_t = \epsilon_t(\theta_0)$  except when  $p = q = P = Q = 0$ . Define the matrix

$$\Xi = \begin{pmatrix} \Sigma & \Sigma_{\hat{\theta}_n, \gamma_m} \\ \Sigma'_{\hat{\theta}_n, \gamma_m} & \Sigma_{\gamma_m} \end{pmatrix} = \sum_{h=-\infty}^{+\infty} \mathbb{E} w_t w'_{t-h}, \text{ where } w_t = \begin{pmatrix} w_{1t} \\ w_{2t} \end{pmatrix} \in \mathbb{R}^{k_0+m} \quad (2.5)$$

with  $w_{1t} = -J^{-1}Y_t = -2\sigma_0^{-2}\epsilon_t J^{-1}(\partial\epsilon_t/\partial\theta)$  and  $w_{2t} = (\epsilon_{t-1}, \dots, \epsilon_{t-m})'\epsilon_t$ . Note that, the existence of the matrix  $\Xi$  is a consequence of **(A3)** and of Davydov's inequality (see [11]).

The asymptotic distribution of  $\sqrt{n}\hat{\rho}_m$  will be obtained from the joint asymptotic distribution of

$$\sqrt{n}(\hat{\theta}'_n - \theta'_0, \gamma'_m)' = \frac{1}{\sqrt{n}} \sum_{t=1}^n w_t + o_{\mathbb{P}}(1) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Xi),$$

by applying the central limit theorem for mixing processes (see [20]).

Now, considering  $\hat{\gamma}(h)$  and  $\gamma(h)$  as values of the same function at the points  $\hat{\theta}_n$  and  $\theta_0$ , a Taylor expansion about  $\theta_0$  gives

$$\begin{aligned} \hat{\gamma}(h) &= \gamma(h) + \frac{1}{n} \sum_{t=h+1}^n \left\{ \epsilon_{t-h}(\theta) \frac{\partial \epsilon_t(\theta)}{\partial \theta'} + \frac{\partial \epsilon_{t-h}(\theta)}{\partial \theta'} \epsilon_t(\theta) \right\}_{\theta=\theta_n^*} (\hat{\theta}_n - \theta_0) + O_{\mathbb{P}}(1/n) \\ &= \gamma(h) + \mathbb{E} \left( \epsilon_{t-h}(\theta_0) \frac{\partial \epsilon_t(\theta_0)}{\partial \theta'} \right) (\hat{\theta}_n - \theta_0) + O_{\mathbb{P}}(1/n), \end{aligned}$$

where  $\theta_n^*$  is between  $\hat{\theta}_n$  and  $\theta_0$ . The last equality follows from the consistency of  $\hat{\theta}_n$  and the fact that  $(\partial \epsilon_{t-h}/\partial \theta')(\theta_0)$  is not correlated with  $\epsilon_t$  when  $h \geq 0$ . Then for  $h = 1, \dots, m$ , we have

$$\hat{\gamma}_m := (\hat{\gamma}(1), \dots, \hat{\gamma}(m))' = \gamma_m + \Phi_m(\hat{\theta}_n - \theta_0) + O_{\mathbb{P}}(1/n), \quad (2.6)$$

where

$$\Phi_m = \mathbb{E} \left\{ \begin{pmatrix} \epsilon_{t-1} \\ \vdots \\ \epsilon_{t-m} \end{pmatrix} \frac{\partial \epsilon_t}{\partial \theta'} \right\}. \quad (2.7)$$

The following Proposition, which is a generalization of Theorem 4 obtained by [6] gives the limiting distribution of the residual autocovariances and autocorrelations of SARMA models.

**Proposition 2.3.1** *When,  $p > 0$ ,  $P > 0$ ,  $q > 0$  and  $Q > 0$ , under the above assumptions, we have*

$$\sqrt{n}\hat{\gamma}_m \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Sigma_{\hat{\gamma}_m}) \quad \text{and} \quad \sqrt{n}\hat{\rho}_m \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Sigma_{\hat{\rho}_m}) \quad \text{where,}$$

$$\Sigma_{\hat{\gamma}_m} = \Sigma_{\gamma_m} + \Phi_m \Sigma \Phi_m' + \Phi_m \Sigma_{\hat{\theta}_n, \gamma_m} + \Sigma_{\hat{\theta}_n, \gamma_m}' \Phi_m' \quad \text{and} \quad \Sigma_{\hat{\rho}_m} = \frac{1}{\sigma_0^4} \Sigma_{\hat{\gamma}_m}.$$

The proof of this result is similar to that given by [6] for Theorem 4.

The asymptotic variance matrices  $\Sigma_{\hat{\gamma}_m}$  and  $\Sigma_{\hat{\rho}_m}$  depend on the unknown matrices  $\Xi$ ,  $\Phi_m$  and the scalar  $\sigma_0^2$ . Matrix  $\Phi_m$  and  $\sigma_0^2$  can be estimated by its empirical counterpart, respectively

$$\hat{\Phi}_m = \frac{1}{n} \sum_{t=1}^n \left\{ (\hat{\epsilon}_{t-1}, \dots, \hat{\epsilon}_{t-m})' \frac{\partial \hat{\epsilon}_t}{\partial \theta'} \right\} \quad \text{and} \quad \hat{\sigma}^2 = \hat{\gamma}(0) = \frac{1}{n} \sum_{t=1}^n \hat{\epsilon}_t^2.$$

Note that the matrix  $(2\pi)^{-1}\Xi$  is the spectral density at frequency zero of the process  $(w_t)$ , thus an estimator of  $\Xi$  is given in Theorem 6 of [7]. Other estimators of such long-run variances are available in the literature (see for instance [1], [2], [12], [25], for general references). For the numerical illustrations presented in this paper, we used a Vector AR (VAR) spectral estimator given in Theorem 6 of [7] consisting in: i) fitting VAR( $r$ ) models for  $r = 1, \dots, r_{\max}$  to the series  $\hat{w}_t$ ,  $t = 1, \dots, n$ , where  $\hat{w}_t$  is obtained by replacing  $\theta_0$  by  $\hat{\theta}_n$  in  $w_t$ ; ii) selecting the order  $r$  which minimizes an information criterion and approximating  $\Xi$  by  $(2\pi)$  times the spectral density at frequency zero of the estimated VAR( $r$ ) model. Hereafter, we used the AIC model selection criterion with  $r_{\max} = 5$ .

From Proposition 2.3.1 we can deduce the following result, which gives the exact limiting distribution of the standard portmanteau statistics (2.2) under general assumptions on the innovation process of the fitted SARMA( $p, q$ )( $P, Q$ ) $_s$  model.

**Theorem 2.3.2** *Under Assumptions in Proposition 2.3.1 and (H0), the statistics  $Q_m^{\text{LB}}$  and  $Q_m^{\text{BP}}$  converge in distribution, as  $n \rightarrow \infty$ , to*

$$Z_m(\xi_m) = \sum_{i=1}^m \xi_{i,m} Z_i^2$$

where  $\xi_m = (\xi_{1,m}, \dots, \xi_{m,m})'$  is the vector of the eigenvalues of the matrix  $\Sigma_{\hat{\rho}_m} = \sigma_0^{-4} \Sigma_{\hat{\gamma}_m}$  and  $Z_1, \dots, Z_m$  are independent  $\mathcal{N}(0, 1)$  variables.

As in [7], Theorem 2.3.2 shows that for the asymptotic distribution of  $Q_m^{\text{LB}}$  and  $Q_m^{\text{BP}}$ , the  $\chi_{m-k_0}^2$  approximation is no longer valid in the framework of weak SARMA( $p, q$ )( $P, Q$ ) $_s$  models. The true asymptotic distribution depends on nuisance parameters involving  $\sigma_0^2$ , the matrix  $\Phi_m$  and the elements of  $\Xi$ . Consequently, in order to obtain the asymptotic distribution of the portmanteau statistics (2.2) under weak assumptions on the noise, one needs a consistent estimator of the asymptotic covariance matrix  $\Sigma_{\hat{\rho}_m}$ . We let  $\hat{\Sigma}_{\hat{\rho}_m}$  the matrix obtained by replacing  $\Xi$  by  $\hat{\Xi}$ ,  $\Phi_m$  by  $\hat{\Phi}_m$  and  $\sigma_0^2$  by  $\hat{\sigma}^2$  in  $\Sigma_{\hat{\rho}_m}$ . Denote by  $\hat{\xi}_m = (\hat{\xi}_{1,m}, \dots, \hat{\xi}_{m,m})'$  the vector of the eigenvalues of  $\hat{\Sigma}_{\hat{\rho}_m}$ .

At the asymptotic level  $\alpha$ , the LB (Ljung-Box) test (resp. the BP (Box-Pierce) test) consists in rejecting the adequacy of the weak SARMA( $p, q$ )(P, Q) $_s$  model when

$$Q_m^{\text{LB}} > S_m(1 - \alpha) \quad (\text{resp. } Q_m^{\text{BP}} > S_m(1 - \alpha)),$$

where  $S_m(1 - \alpha)$  is such that  $\mathbb{P}\{Z_m(\hat{\xi}_m) > S_m(1 - \alpha)\} = \alpha$ . We emphasize the fact that the proposed modified versions of the Box-Pierce and Ljung-Box statistics are more difficult to implement because their critical values have to be computed from the data.

### 2.3.2 Self-normalized asymptotic distribution of the residual autocorrelations

The nonparametric kernel estimator (see [1, 25]), used to estimate the matrix  $\Xi$  causes serious difficulties regarding the choice of the sequence of weights. The parametric approach based on an autoregressive estimate of the spectral density of  $w_t$  studied for instance by [2, 3, 5, 7, 12] is also facing the problem of choosing the truncation parameter. So the choice of the order of truncation is often crucial and difficult. In this section, we propose as in [8] an alternative method where we do not estimate an asymptotic covariance matrix. It is based on a self-normalization based approach to construct a test-statistic which is asymptotically distribution-free under the null hypothesis (see [8], for a reference in the ARMA cases). The idea comes from [23] and has been already extended by [21, 27, 28, 29] to more general frameworks. See also [30] for a review on some recent developments on the inference of time series data using the self-normalized approach. In this case, the critical values are not computed from the data since they are tabulated. In some sense, this method is finally closer to the standard method in which the critical values are simply deduced from a  $\chi^2$ -table.

We denote  $\Lambda$  the matrix in  $\mathbb{R}^{m \times (k_0 + m)}$  defined in block formed by  $\Lambda = (\Phi_m | I_m)$ , where  $I_m$  is the identity matrix of order  $m$ . In view of (2.4) and (2.6), we deduce that

$$\sqrt{n} \hat{Y}_m = \frac{1}{\sqrt{n}} \sum_{t=1}^n \Lambda w_t + o_{\mathbb{P}}(1). \quad (2.8)$$

Contrarily to Subsection 2.3.1, we do not rely on the classical method that would consist in estimating the asymptotic covariance matrix of  $\Lambda w_t$ . We need to apply the functional central limit theorem holds for the process  $w = (w_t)_{t \geq 1}$  (see Lemma 1 in [23]).

Finally, we define the normalization matrix  $C_m \in \mathbb{R}^{m \times m}$  by

$$C_m = \frac{1}{n^2} \sum_{t=1}^n S_t S_t' \quad \text{where } S_t = \sum_{j=1}^t (\Lambda w_j - \Lambda \bar{w}) \quad \text{with } \bar{w} = \frac{1}{n} \sum_{t=1}^n w_t.$$

To ensure the invertibility of the normalization matrix  $C_m$  which is proved in Lemma 6 of [8], we need the following technical assumption on the distribution of  $\epsilon_t$ .

**(A4):** The process  $(\epsilon_t)_{t \in \mathbb{Z}}$  has a positive density on some neighborhood of zero.

Let  $(B_K(r))_{r \geq 0}$  be a  $K$ -dimensional Brownian motion starting from 0. For  $K \geq 1$ , we denote  $\mathcal{U}_K$  the random variable defined by

$$\mathcal{U}_K = B_K'(1) V_K^{-1} B_K(1) \quad \text{where } V_K = \int_0^1 (B_K(r) - r B_K(1)) (B_K(r) - r B_K(1))' dr. \quad (2.9)$$

The following theorem states the asymptotic distributions of the sample autocovariances and autocorrelations.

**Theorem 2.3.3** *We assume that  $p > 0$ ,  $q > 0$ ,  $P > 0$  or  $Q > 0$ . Under Assumptions of Proposition 2.3.1, (A4) and under the null hypothesis (H0), we have*

$$n \hat{\gamma}'_m C_m^{-1} \hat{\gamma}_m \xrightarrow[n \rightarrow \infty]{d} \mathcal{U}_m \text{ and } n \sigma_0^4 \hat{\rho}'_m C_m^{-1} \hat{\rho}_m \xrightarrow[n \rightarrow \infty]{d} \mathcal{U}_m$$

The proof of this result is postponed to Section 2.6.

Of course, the above theorem is useless for practical purpose, because it does not involve any observable quantities. In practice, one has to replace the matrix  $C_m$  and the variance of the noise  $\sigma_0^2$  by their empirical or observable counterparts. The matrix  $J$  can be easily estimated by his empirical counterpart

$$\hat{J} = \frac{2}{\hat{\sigma}^2} \frac{1}{n} \sum_{t=1}^n \frac{\partial e_t(\hat{\theta}_n)}{\partial \theta} \frac{\partial e_t(\hat{\theta}_n)}{\partial \theta'}.$$

Thus we define

$$\hat{\Lambda} = \left( \hat{\Phi}_m | I_m \right) \text{ and } \hat{w}_t = \left( -2 \hat{J}_n^{-1} \frac{\partial e_t(\hat{\theta}_n)}{\partial \theta'} \frac{1}{\hat{\sigma}^2} \hat{e}_t, \hat{e}_t \hat{e}_{t-1}, \dots, \hat{e}_t \hat{e}_{t-m} \right)'.$$

Finally we denote the normalization matrix  $\hat{C}_m \in \mathbb{R}^{m \times m}$  by

$$\hat{C}_m = \frac{1}{n^2} \sum_{t=1}^n \hat{S}_t \hat{S}_t' \text{ where } \hat{S}_t = \sum_{j=1}^t (\hat{\Lambda} \hat{w}_j - \hat{\Lambda} \bar{w}) \text{ with } \bar{w} = \frac{1}{n} \sum_{t=1}^n \hat{w}_t.$$

The following result is the applicable counterpart of Theorem 2.3.3.

**Theorem 2.3.4** *Assume that  $p > 0$ ,  $q > 0$ ,  $P > 0$  or  $Q > 0$ . Under Assumptions of Theorem 2.3.3, we have*

$$n \hat{\gamma}'_m \hat{C}_m^{-1} \hat{\gamma}_m \xrightarrow[n \rightarrow \infty]{d} \mathcal{U}_m \text{ and } Q_m^{\text{SN}} = n \hat{\sigma}^4 \hat{\rho}'_m \hat{C}_m^{-1} \hat{\rho}_m \xrightarrow[n \rightarrow \infty]{d} \mathcal{U}_m.$$

The proof of this result is postponed to Section 2.6.

Based on the above result, we propose a modified version of the Ljung-Box statistic when one uses the statistic

$$\tilde{Q}_m^{\text{SN}} = n \hat{\sigma}^4 \hat{\rho}'_m D_{n,m}^{1/2} \hat{C}_m^{-1} D_{n,m}^{1/2} \hat{\rho}_m, \quad (2.10)$$

where the matrix  $D_{n,m} \in \mathbb{R}^{m \times m}$  is diagonal with  $((n+2)/(n-1), \dots, (n+2)/(n-m))$  as diagonal terms.

## 2.4 Numerical illustrations

In this section, by means of Monte Carlo experiments, we investigate the finite sample properties of the modified version of the portmanteau tests that we introduced in this work. The numerical illustrations of this section are made with the open source statistical software R (see R Development Core Team, 2017) or (see <http://cran.r-project.org/>).

### 2.4.1 Simulated models

First of all, we introduce the models that we simulate and we indicate the conventions that we adopt in the discussion and in the tables:

- $\text{LB}_W$  and  $\text{BP}_W$  refer to modified LB and BP tests using  $Q_m^{\text{LB}}$  and  $Q_m^{\text{BP}}$  in Section 2.3.1
- $\text{LB}_S$  and  $\text{BP}_S$  refer to LB and BP tests using the standard statistics (2.2).

—  $\text{LB}_{\text{SN}}$  and  $\text{BP}_{\text{SN}}$  refer to modified tests using the self-normalized statistics in Section 2.3.2. To generate the strong and the weak SARMA models, we consider the following SARMA(0, 1)(0, 1)<sub>s</sub> model

$$X_t = \epsilon_t - b_{01}\epsilon_{t-1} - \mathbf{b}_{01}\epsilon_{t-s} + b_{01}\mathbf{b}_{01}\epsilon_{t-s-1}, \quad (2.11)$$

with  $\theta_0 = (b_{01}, \mathbf{b}_{01})' = (-0.6, -0.7)'$  and the innovation process  $(\epsilon_t)$  follows a strong or weak white noise.

The generalized autoregressive conditional heteroscedastic (GARCH) models is an important example of weak white noises in the univariate case (see [19]). So we first assume that in (2.11) the innovation process  $\epsilon$  is the following ARCH(1) model defined by

$$\begin{cases} \epsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = 1 + \alpha_1 \epsilon_{t-1}^2 \end{cases} \quad (2.12)$$

where  $(\eta_t)_{t \geq 1}$  is a sequence of iid standard Gaussian random variables. To generate the strong SARMA, we assume that in (2.11) the innovation process follows (2.12) with  $\alpha_1 = 0$ .

### 2.4.2 Empirical size

We first simulate  $N = 1,000$  independent trajectories of size  $n = 2,000$  of models (2.11). The same series is partitioned as two series of sizes  $n = 500$  and  $n = 2,000$ . For each of these  $N$  replications, we use the quasi-maximum likelihood estimation method to estimate the coefficient  $\theta_0$  and we apply portmanteau tests to the residuals for different values of  $m \in \{4, 8, 12, 15, 18, 20\}$ , where  $m$  is the number of autocorrelations used in the portmanteau test statistic. For the nominal level  $\alpha = 5\%$ , the empirical size over the  $N$  independent replications should vary between the significant limits 3.6% and 6.4% with probability 95% and belong to [3.2%, 6.9%] with a probability 99%. When the relative rejection frequencies are outside the 95% significant limits, they are displayed in bold type and they are underlined when they are outside the 99% significant limits in Tables 2.1 and 2.2.

For the standard Box-Pierce test, the model is therefore rejected when the statistic  $Q_m^{\text{BP}}$  or  $Q_m^{\text{LB}}$  is larger than  $\chi_{(m-2)}^2(0.95)$  in a SARMA(0, 1)(0, 1)<sub>s</sub> case (see [24]). Consequently the empirical size is not available (n.a.) for the statistic  $Q_m^{\text{BP}}$  or  $Q_m^{\text{LB}}$  because they are not applicable for  $m \leq 2$ . For the proposed self-normalized test  $\text{BP}_{\text{SN}}$  or  $\text{LB}_{\text{SN}}$ , the model is rejected when the statistic  $Q_m^{\text{SN}}$  or  $\tilde{Q}_m^{\text{SN}}$  is larger than  $\mathcal{U}_m(0.95)$ , where the critical values  $\mathcal{U}_K(0.95)$  (for  $K = 1, \dots, 20$ ) are tabulated in Lobato (see Table 1 in [23]).

Table 2.1 displays the relative rejection frequencies of the null hypothesis (**H0**) that the data generating process (DGP for short) follows a strong SARMA model (2.11)–(2.12) with  $\alpha_1 = 0$ , over the  $N$  independent replications. When the seasonal period is  $s = 4$ , for all tests, the percentages of rejection belong to the confident interval with probabilities 95% and 99%, except for  $\text{LB}_s$  and  $\text{BP}_s$  when  $m = 4$ . Consequently all these tests well control the error of first kind. In contrast, when  $s = 12$ , our proposed tests well also control the error of first kind (except for  $\text{LB}_w$  and  $\text{BP}_w$  when  $n = 500$ ) contrarily to the standard tests  $\text{LB}_s$ ,  $\text{BP}_s$  for all sizes. We draw the conclusion that, in this strong SARMA case, the proposed modified version may be clearly preferable to the standard ones.

Now, we repeat the same experiments on a weak SARMA models. As expected, Table 2.2 shows that the standard  $\text{LB}_s$  or  $\text{BP}_s$  test poorly performs in assessing the adequacy of this particular weak SARMA model. It can be seen that: 1) the observed relative rejection frequencies of  $\text{LB}_s$  and  $\text{BP}_s$  are definitely outside the significant limits, 2) the errors of the first kind are only globally well controlled by the proposed tests, for all  $s$  when  $n$  is large. We also tried the case where the ARCH(1) model (2.12) have infinite fourth moments. As showing in Figure 2.1, the results are qualitatively similar to what we observe here.



Figure 2.1 displays the residual autocorrelations of a realization of size  $n = 2,000$  for weak SARMA(0,1)(0,1)<sub>12</sub> model (2.11)-(2.12) with  $\alpha_1 = 1.3$  and their 5% significance limits under the strong SARMA(0,1)(0,1)<sub>12</sub> and weak SARMA(0,1)(0,1)<sub>12</sub> assumptions. This figure confirms clearly the conclusions drawn from Table 2.2. The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong SARMA assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak SARMA assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Proposition 2.3.1. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 2.3.4.

In these Monte Carlo experiments, we illustrate that the proposed test statistics have reasonable finite sample performance. Under nonindependent errors, it appears that the standard test statistics are generally non reliable, overrejecting severely, while the proposed tests statistics offer satisfactory levels. Even for independent errors, they seem preferable to the standard ones, when the number  $m$  of autocorrelations is small and when  $s = 12$ . Moreover, the error of first kind is well controlled. Contrarily to the standard tests based on  $BP_s$  or  $LB_s$ , the proposed tests can be used safely for  $m$  small (see for instance Figure 2.1). For all these above reasons, we think that the modified versions that we propose in this paper are preferable to the standard ones for diagnosing SARMA models under nonindependent errors.

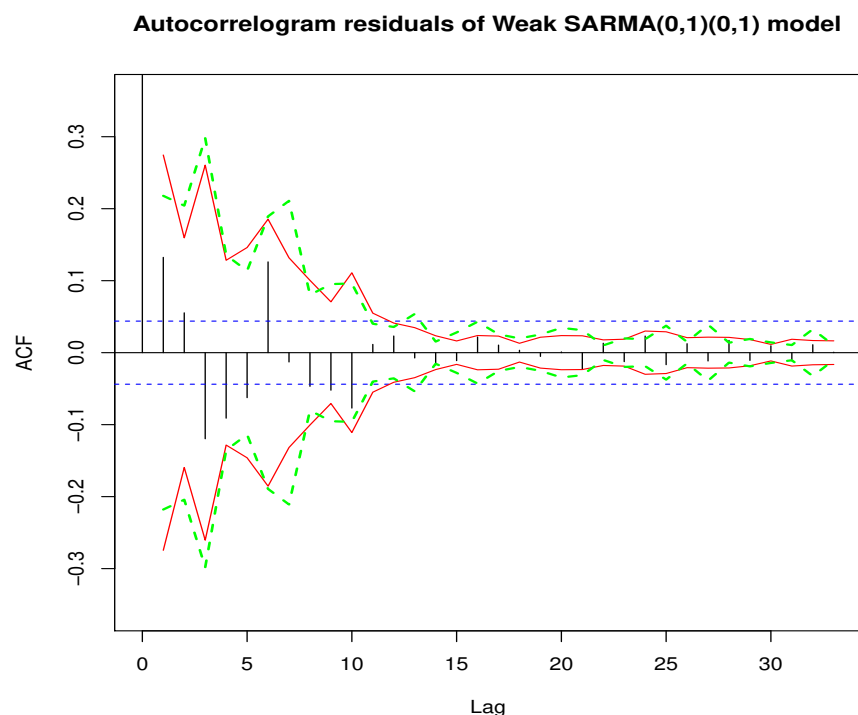


Figure 2.1 – Autocorrelation of a realization of size  $n = 2,000$  for weak SARMA(0,1)(0,1)<sub>12</sub> model (2.11)-(2.12) with  $\alpha_1 = 1.3$ . The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong SARMA assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak SARMA assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Proposition 2.3.1. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 2.3.4.

Table 2.1 – Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of SARMA(0, 1)(0, 1)<sub>s</sub> model (2.11)-(2.12) with  $\alpha_1 = 0$ . The nominal asymptotic level of the tests is  $\alpha = 5\%$ . The number of replications is  $N = 1,000$ .

$s$	Length $n$	Lag $m$	LB <sub>SN</sub>	BP <sub>SN</sub>	LB <sub>W</sub>	BP <sub>W</sub>	LB <sub>S</sub>	BP <sub>S</sub>
4	$n = 500$	4	4.3	4.0	<b>6.8</b>	<b>6.8</b>	<b>8.7</b>	<b>8.7</b>
		8	5.8	5.4	5.8	5.6	<b>7.2</b>	<b>6.7</b>
		12	4.9	4.8	6.0	5.4	<b>7.2</b>	6.3
		15	6.3	5.3	5.1	4.7	5.8	5.3
		18	5.7	5.1	5.4	4.3	5.7	4.9
		20	5.7	5.0	5.2	3.8	5.4	4.1
4	$n = 2,000$	4	4.1	4.1	5.7	5.6	<b>7.5</b>	<b>7.5</b>
		8	4.2	4.2	4.7	4.7	5.9	5.6
		12	4.9	4.8	3.9	3.9	4.6	4.4
		15	4.6	4.6	4.9	4.4	4.9	4.5
		18	4.0	3.8	4.4	4.2	4.6	4.1
		20	4.5	4.5	5.1	5.0	5.4	5.1
12	$n = 500$	4	4.7	4.5	<b>12.8</b>	<b>12.1</b>	<b>20.8</b>	<b>20.6</b>
		8	<b>6.5</b>	6.4	<b>12.1</b>	<b>11.4</b>	<b>15.1</b>	<b>14.6</b>
		12	5.4	5.1	<b>10.7</b>	<b>10.1</b>	<b>11.7</b>	<b>10.8</b>
		15	6.0	5.6	<b>8.7</b>	<b>8.4</b>	<b>9.8</b>	<b>9.4</b>
		18	5.1	4.5	<b>8.3</b>	<b>7.5</b>	<b>9.7</b>	<b>8.3</b>
		20	5.5	4.3	<b>8.6</b>	<b>7.6</b>	<b>9.8</b>	<b>8.4</b>
12	$n = 2,000$	4	4.4	4.3	<b>7.2</b>	<b>7.0</b>	<b>14.0</b>	<b>13.7</b>
		8	4.9	4.9	5.8	5.7	<b>9.9</b>	<b>9.8</b>
		12	4.5	4.4	6.0	5.6	<b>7.0</b>	<b>6.6</b>
		15	4.5	4.5	6.1	6.0	<b>7.0</b>	<b>6.9</b>
		18	5.4	5.3	6.0	5.7	<b>6.9</b>	<b>6.5</b>
		20	4.9	4.8	6.3	6.1	<b>7.5</b>	<b>7.1</b>

### 2.4.3 Empirical power

In this section we repeat the same experiments as in Section 2.4.1 to examine the power of the tests for the null hypothesis of a SARMA(0, 1)(0, 1)<sub>s</sub> against the following SARMA(1, 1)(0, 1)<sub>s</sub> alternative defined by

$$X_t = a_{01}X_{t-1} + \epsilon_t - b_{01}\epsilon_{t-1} - \mathbf{b}_{01}\epsilon_{t-s} + b_{01}\mathbf{b}_{01}\epsilon_{t-s-1}, \quad (2.13)$$

with  $\theta_0 = (a_{01}, b_{01}, \mathbf{b}_{01})' = (0.8, -0.6, -0.7)'$  and where the innovation process  $\epsilon$  follows a strong or weak white noise introduced in Section 2.4.1. For each of these  $N$  replications we fit a SARMA(0, 1)(0, 1)<sub>s</sub> models and perform standard and modified tests based on  $m = 4, 8, 12, 15, 18$  and  $20$  residual autocorrelations.

Tables 2.3 and 2.4 compare the empirical powers of Model (2.13)-(2.12) with  $\alpha_1 = 0$  and  $\alpha_1 = 0.45$  respectively over the  $N$  independent replications. For these particular strong and weak SARMA models, we notice that the standard BP<sub>S</sub> and LB<sub>S</sub> and our proposed tests have very similar powers except for BP<sub>SN</sub> and LB<sub>SN</sub> when  $n = 500$  in the weak case.



Table 2.2 – Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of SARMA(0, 1)(0, 1)<sub>s</sub> model (2.11)-(2.12) with  $\alpha_1 = 0.45$ . The nominal asymptotic level of the tests is  $\alpha = 5\%$ . The number of replications is  $N = 1,000$ .

$s$	Length $n$	Lag $m$	LB <sub>SN</sub>	BP <sub>SN</sub>	LB <sub>W</sub>	BP <sub>W</sub>	LB <sub>S</sub>	BP <sub>S</sub>
4	$n = 500$	4	<b>3.1</b>	<b>2.9</b>	4.5	4.3	<b>18.3</b>	<b>18.3</b>
		8	<b>2.8</b>	<b>2.6</b>	<b>3.3</b>	<b>3.2</b>	<b>11.7</b>	<b>11.3</b>
		12	<b>3.5</b>	<b>3.2</b>	<b>2.8</b>	<b>2.4</b>	<b>9.7</b>	<b>9.3</b>
		15	<b>2.8</b>	<b>2.3</b>	<b>2.9</b>	<b>2.7</b>	<b>10.0</b>	<b>9.0</b>
		18	<b>1.9</b>	<b>1.6</b>	<b>2.7</b>	<b>2.4</b>	<b>9.1</b>	<b>7.8</b>
		20	<b>1.7</b>	<b>1.5</b>	<b>2.6</b>	<b>2.3</b>	<b>9.0</b>	<b>7.9</b>
4	$n = 2,000$	4	5.5	5.4	4.2	4.2	<b>19.9</b>	<b>19.9</b>
		8	4.8	4.8	3.6	3.6	<b>13.9</b>	<b>13.7</b>
		12	4.5	4.3	3.8	3.6	<b>11.4</b>	<b>11.3</b>
		15	4.8	4.8	3.6	3.6	<b>10.7</b>	<b>10.5</b>
		18	3.7	3.7	<b>3.4</b>	<b>3.4</b>	<b>10.2</b>	<b>9.8</b>
		20	<b>3.2</b>	<b>3.1</b>	<b>3.5</b>	<b>3.5</b>	<b>9.4</b>	<b>9.1</b>
12	$n = 500$	4	3.6	<b>3.5</b>	<b>9.1</b>	<b>8.9</b>	<b>29.6</b>	<b>29.1</b>
		8	4.0	3.8	6.4	6.0	<b>20.4</b>	<b>19.7</b>
		12	4.2	<b>3.5</b>	5.5	5.2	<b>15.4</b>	<b>14.6</b>
		15	<b>2.2</b>	<b>1.9</b>	5.4	4.8	<b>14.6</b>	<b>13.5</b>
		18	<b>1.7</b>	<b>1.4</b>	4.3	4.2	<b>13.3</b>	<b>12.0</b>
		20	<b>1.8</b>	<b>1.6</b>	4.9	4.2	<b>12.9</b>	<b>11.8</b>
12	$n = 2,000$	4	4.4	4.4	4.5	4.5	<b>27.1</b>	<b>27.1</b>
		8	4.8	4.7	4.1	3.9	<b>20.3</b>	<b>20.1</b>
		12	3.8	3.8	4.6	4.4	<b>15.3</b>	<b>15.2</b>
		15	4.6	4.5	<b>3.5</b>	<b>3.4</b>	<b>13.9</b>	<b>13.7</b>
		18	3.8	3.7	3.6	<b>3.5</b>	<b>12.6</b>	<b>12.4</b>
		20	<b>3.5</b>	<b>3.3</b>	4.1	3.8	<b>11.9</b>	<b>11.8</b>

#### 2.4.4 Application to real data

We now consider an application to monthly mean total sunspot number obtained by taking a simple arithmetic mean of the daily total sunspot number over all days of each calendar month. The observations (sunspot) covered the period from January 01, 2010 to December 31, 2018 which correspond to  $n = 108$  observations. The series exhibit seasonal behavior ( $s = 12$ ). The data were obtained from the website of the World Data Center, Solar Influences Data Analysis Center, Royal Observatory of Belgium (<http://www.sidc.be/silso/datafiles>).

Let  $Z_t = \log(\text{sunspot}_t) - \log(\text{sunspot}_{t-1})$  and denoting by  $X_t = Z_t - \mathbb{E}(Z_t)$  the mean-corrected series. We adjust the particular SARMA(3, 1)(0, 1)<sub>12</sub> model of the form

$$X_t = a_{03}X_{t-3} + \epsilon_t - b_{01}\epsilon_{t-1} - \mathbf{b}_{01}\epsilon_{t-s} + b_{01}\mathbf{b}_{01}\epsilon_{t-s-1}.$$

The quasi-maximum likelihood estimators of  $\theta_0 = (a_{03}, \mathbf{b}_{01}, \mathbf{b}_{01})'$  were obtained as

$$\hat{\theta}_n = \begin{pmatrix} -0.1810 & [0.0253] & (0.0000) \\ 0.5438 & [0.0204] & (0.0000) \\ -0.1139 & [0.0237] & (0.0000) \end{pmatrix} \text{ and } \hat{\sigma}_\epsilon^2 = 0.1728,$$

Table 2.3 – Empirical power (in %) of the modified and standard versions of the LB and BP tests in the case of SARMA(0, 1)(0, 1)<sub>s</sub> model (2.13)-(2.12) with  $\alpha_1 = 0$ . The number of replications is  $N = 1,000$ .

$s$	Length $n$	Lag $m$	LB <sub>SN</sub>	BP <sub>SN</sub>	LB <sub>W</sub>	BP <sub>W</sub>	LB <sub>S</sub>	BP <sub>S</sub>
4	$n = 500$	4	98.9	98.9	100.0	100.0	100.0	100.0
		8	98.1	98.0	100.0	100.0	100.0	100.0
		12	96.3	96.3	100.0	100.0	100.0	100.0
		15	94.9	94.9	99.9	99.9	100.0	100.0
		18	93.7	93.5	99.8	99.7	100.0	100.0
	20	91.6	91.4	99.7	99.7	100.0	100.0	
	$n = 2,000$	4	100.0	100.0	100.0	100.0	100.0	100.0
		8	100.0	100.0	100.0	100.0	100.0	100.0
		12	100.0	100.0	100.0	100.0	100.0	100.0
		15	100.0	100.0	100.0	100.0	100.0	100.0
18		100.0	100.0	100.0	100.0	100.0	100.0	
12	$n = 500$	4	98.1	98.1	100.0	100.0	100.0	100.0
		8	97.7	97.7	99.9	99.9	100.0	100.0
		12	96.6	96.6	99.9	99.9	100.0	100.0
		15	95.5	95.4	99.9	99.9	100.0	100.0
		18	92.9	92.7	99.9	99.9	100.0	100.0
	20	90.9	90.5	99.9	99.9	100.0	100.0	
	$n = 2,000$	4	100.0	100.0	100.0	100.0	100.0	100.0
		8	100.0	100.0	100.0	100.0	100.0	100.0
		12	100.0	100.0	100.0	100.0	100.0	100.0
		15	100.0	100.0	100.0	100.0	100.0	100.0
18		100.0	100.0	100.0	100.0	100.0	100.0	
20	100.0	100.0	100.0	100.0	100.0	100.0		

where the estimated asymptotic standard errors obtained from (2.4) (respectively the  $p$ -values), of the estimated parameters (first column), are given into brackets (respectively in parentheses). We apply portmanteau tests to the residuals of this model. Figure 2.2 displays the residual autocorrelations and their 5% significance limits under the strong SARMA(3, 1)(0, 1)<sub>12</sub> and weak SARMA(3, 1)(0, 1)<sub>12</sub> assumptions. In view of Figure 2.2, the diagnostic checking of residuals does not indicate any inadequacy. All of the sample autocorrelations should lie between the bands (at 95%) shown as dashed lines (green color), solid lines (red color) and the horizontal dotted (blue color).

Table 2.4 – Empirical power (in %) of the modified and standard versions of the LB and BP tests in the case of ARMA(0, 1)(0, 1)<sub>s</sub> model (2.13)-(2.12) with  $\alpha_1 = 0.45$ . The number of replications is  $N = 1000$ .

$s$	Length $n$	Lag $m$	LB <sub>SN</sub>	BP <sub>SN</sub>	LB <sub>W</sub>	BP <sub>W</sub>	LB <sub>S</sub>	BP <sub>S</sub>
4	$n = 500$	4	89.2	89.1	98.5	98.5	100.0	100.0
		8	83.0	82.9	96.8	96.7	100.0	100.0
		12	71.5	71.5	96.3	96.3	100.0	100.0
		15	64.9	64.7	95.6	95.6	100.0	100.0
		18	56.2	55.7	95.1	95.0	100.0	100.0
	20	50.6	49.4	94.5	94.4	100.0	100.0	
	$n = 2,000$	4	99.5	99.5	100.0	100.0	100.0	100.0
		8	99.4	99.4	100.0	100.0	100.0	100.0
		12	99.1	99.1	100.0	100.0	100.0	100.0
		15	99.2	99.2	99.9	99.9	100.0	100.0
18		99.5	99.5	100.0	100.0	100.0	100.0	
12	$n = 500$	20	99.2	99.2	99.9	99.9	100.0	100.0
		4	90.1	90.1	98.4	98.3	100.0	100.0
		8	82.4	82.2	97.3	97.2	100.0	100.0
		12	74.3	74.0	97.0	96.9	100.0	100.0
		15	68.7	68.3	96.9	96.5	100.0	100.0
	$n = 2,000$	18	56.9	56.0	96.1	95.7	100.0	100.0
		20	48.9	47.9	96.0	95.7	100.0	100.0
		4	99.6	99.6	100.0	100.0	100.0	100.0
		8	99.6	99.6	99.9	99.9	100.0	100.0
		12	99.5	99.5	100.0	100.0	100.0	100.0
12	$n = 2,000$	15	99.4	99.4	100.0	100.0	100.0	100.0
		18	99.5	99.5	100.0	100.0	100.0	100.0
		20	99.0	99.0	100.0	99.9	100.0	100.0

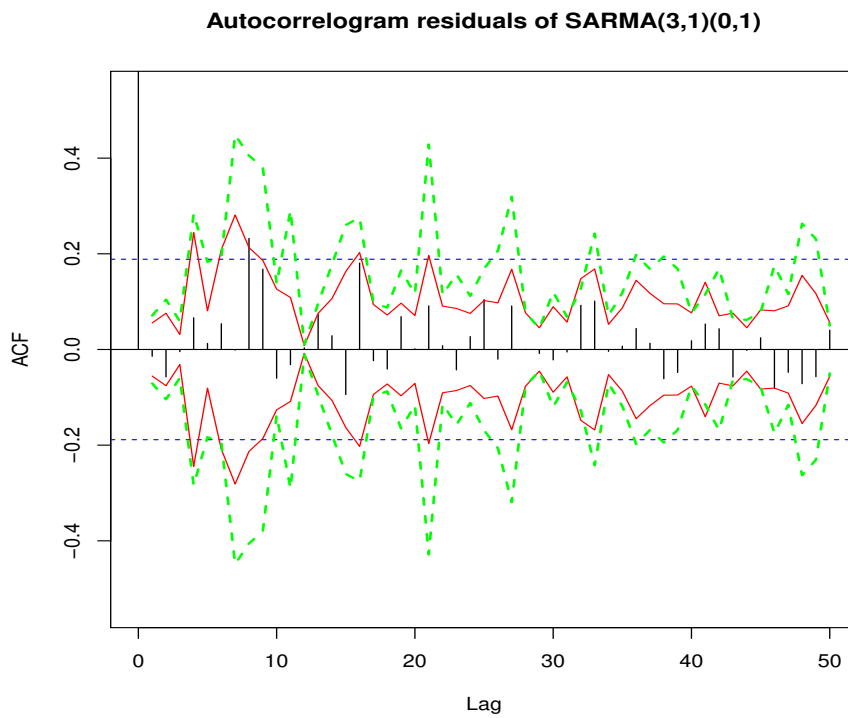


Figure 2.2 – Autocorrelation of the particular  $\text{SARMA}(3,1)(0,1)_{12}$  model residuals for the mean-logarithm-corrected of monthly mean total sunspot number. The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong SARMA assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak SARMA assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Proposition 2.3.1. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 2.3.4.

## 2.5 Conclusion

From these simulation experiments and from the asymptotic theory, we draw the conclusion that the standard methodology, based on the QMLE, allows to fit SARMA representations of a wide class of nonlinear time series. But it is often restrictive to consider that the innovation process is directly observed. In future works, we intent to study how the existing estimation (see [17, 19]) and diagnostic checking (see [31]) procedures should be adapted in the situation where the GARCH process used in these simulation experiments is not directly observed, but constitutes the innovation of an observed SARMA-(seasonal)GARCH process which will be able to extend considerably the range of applications.

## 2.6 Proofs

The proofs of Theorems 2.3.3 and 2.3.4 follow the same lines as in [8] and are similar. To have its own autonomy, the proofs will be rewrite and adapt.

### Proof of Theorem 2.3.3

We recall that the Skorokhod space  $\mathbb{D}^k[0,1]$  is the set of  $\mathbb{R}^k$ -valued functions defined on  $[0,1]$  which are right continuous and have left limits. It is endowed with the Skorokhod topology and the weak convergence on  $\mathbb{D}^k[0,1]$  is mentioned by  $\xrightarrow{\mathbb{D}^k}$ . We finally denote by  $\lfloor x \rfloor$  the integer part of the real  $x$ .

To prove the result we need to recall some following results of [4, 14, 15, 16, 24].

We denote by  $a_i^*$ ,  $b_i^*$ ,  $\mathbf{a}_i^*$  and  $\mathbf{b}_i^*$  the coefficients defined by

$$a_{\theta_0}^{-1}(z) = \sum_{i \geq 0} a_i^* z^i, \quad b_{\theta_0}^{-1}(z) = \sum_{i \geq 0} b_i^* z^i, \quad \mathbf{a}_{\theta_0}^{-1}(z) = \sum_{i \geq 0} \mathbf{a}_i^* z^i \quad \text{and} \quad \mathbf{b}_{\theta_0}^{-1}(z) = \sum_{i \geq 0} \mathbf{b}_i^* z^i, \quad |z| \leq 1.$$

Following [14] and [24] (see also [4]), the noise derivatives involving in the expression of  $J(\theta_0)$  and  $I(\theta_0)$  can be represented as

$$\begin{aligned} \frac{\partial \epsilon_t}{\partial \theta} &= \sum_{i \geq 1} \lambda_i \epsilon_{t-i}, \\ \lambda_i &= \left( -a_{i-1}^*, \dots, -a_{i-p}^*, b_{i-1}^*, \dots, b_{i-q}^*, -\mathbf{a}_{i-1s}^*, \dots, -\mathbf{a}_{i-ps}^*, \mathbf{b}_{i-1s}^*, \dots, \mathbf{b}_{i-qs}^* \right)' \in \mathbb{R}^{k_0}, \end{aligned} \quad (2.14)$$

with  $a_i^* = b_i^* = \mathbf{a}_i^* = \mathbf{b}_i^* = 0$  when  $i < 0$ .

For any  $\theta \in \Theta \subset \mathbb{R}^{k_0}$  and any  $(l, m) \in \{1, \dots, k_0\}^2$ , under the above Assumptions, there exists absolutely summable and deterministic sequences  $(c_i(\theta))_{i \geq 0}$ ,  $(\lambda_{i,l}(\theta))_{i \geq 1}$  and  $(\lambda_{i,l,m}(\theta))_{i \geq 1}$  such that, almost surely,

$$\epsilon_t(\theta) = \sum_{i=0}^{\infty} c_i(\theta) \epsilon_{t-i}, \quad \frac{\partial \epsilon_t(\theta)}{\partial \theta_l} = \sum_{i=1}^{\infty} \lambda_{i,l}(\theta) \epsilon_{t-i} \quad \text{and} \quad \frac{\partial^2 \epsilon_t(\theta)}{\partial \theta_l \partial \theta_m} = \sum_{i=2}^{\infty} \lambda_{i,l,m}(\theta) \epsilon_{t-i} \quad (2.15)$$

$$e_t(\theta) = \sum_{i=0}^{t-1} c_i(\theta) e_{t-i}, \quad \frac{\partial e_t(\theta)}{\partial \theta_l} = \sum_{i=1}^{t-1} \lambda_{i,l}(\theta) e_{t-i} \quad \text{and} \quad \frac{\partial^2 e_t(\theta)}{\partial \theta_l \partial \theta_m} = \sum_{i=2}^{t-1} \lambda_{i,l,m}(\theta) e_{t-i} \quad (2.16)$$

with  $c_0(\theta) = 1$ . A useful property of the above three sequences that they are asymptotically exponentially small. Indeed there exists  $\rho \in ]0,1[$  and a positive constant  $K$  such that, for all  $i \geq 1$ , we have

$$\sup_{\theta \in \Theta} \left( |c_i(\theta)| + |\lambda_{i,l}(\theta)| + |\lambda_{i,l,m}(\theta)| \right) \leq K \rho^i. \quad (2.17)$$

See Lemmas A.1. and A.2. of [16] for a more detailed treatment.

Now, in view of (2.8) it is clear that the asymptotic behaviour of  $\hat{\gamma}_m$  is related to the limit distribution of  $w_t = \left( -Y'_t J^{-1'}, \epsilon_t \epsilon_{t-1}, \dots, \epsilon_t \epsilon_{t-m} \right)'$ . First, we prove that  $\frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor nr \rfloor} \Lambda w_j$  converges on the Skorokhod space to a Brownian motion. More preciesly, we have to show that

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor nr \rfloor} \Lambda w_j \xrightarrow[n \rightarrow \infty]{\mathbb{D}^m} (\Psi \Psi')^{1/2} B_m(r) \quad (2.18)$$

where  $(B_m(r))_{r \geq 0}$  is a  $m$ -dimensional standard Brownian motion.

Using (2.14), the process  $w_t$  can be rewritten as

$$w_t = \left( -2\sigma_0^{-2} \left\{ \sum_{i=1}^{\infty} \lambda_{i,1}(\theta_0) \epsilon_t \epsilon_{t-i}, \dots, \sum_{i=1}^{\infty} \lambda_{i,k_0}(\theta_0) \epsilon_t \epsilon_{t-i} \right\}' J^{-1'}, \epsilon_t \epsilon_{t-1}, \dots, \epsilon_t \epsilon_{t-m} \right)'$$

and thus the non-correlation between  $\epsilon_t$  implies that  $w_t$  has zero expectation with values in  $\mathbb{R}^{k_0+m}$ . In order to apply the functional central limit theorem for strongly mixing process, we need to identify the asymptotic covariance matrix in the classical central limit theorem for the sequence  $(w_t)_{t \geq 1}$ . It is proved in Subsection 2.3.1 that

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n w_t \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Xi := 2\pi f_w(0)) \quad (2.19)$$

where  $f_w(0)$  is the spectral density of the stationary process  $(w_t)_{t \in \mathbb{Z}}$  evaluated at frequency 0. The main issue is to prove the existence of the matrix  $\Xi$  which is a consequence **(A3)** and Davydov's inequality [11]. For that sake, one has to introduce for any integer  $k$ , the random variables

$$w_t^k = \left( -2\sigma_0^{-2} \left\{ \sum_{i=1}^k \lambda_{i,1}(\theta_0) \epsilon_t \epsilon_{t-i}, \dots, \sum_{i=1}^k \lambda_{i,k_0}(\theta_0) \epsilon_t \epsilon_{t-i} \right\}' , \epsilon_t \epsilon_{t-1}, \dots, \epsilon_t \epsilon_{t-m} \right)'.$$

Since  $w^k$  depends on a finite number of values of the noise-process  $\epsilon$ , it also satisfies a mixing property (see Theorem 14.1 in [10], p. 210). Based on the Davydov inequality (see [11]), the arguments developed in the Lemma A.1 in [14] (see also [15]) imply that

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n w_t^k \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Xi_k) \quad (2.20)$$

where

$$\Xi_k := 2\pi f_{w^k}(0) = \sum_{h=-\infty}^{+\infty} \text{Cov}(w_t^k, w_{t-h}^k) = \sum_{h=-\infty}^{+\infty} \mathbb{E}(w_t^k w_{t-h}^{k'})$$

and thus (2.19) holds. Moreover we have that  $\lim_{k \rightarrow \infty} \Xi_k = \Xi$ .

Since the matrix  $\Xi$  is positive definite, it can be factorized as  $\Xi = \Delta \Delta'$  where the  $(k_0 + m) \times (k_0 + m)$  lower triangular matrix  $\Delta$  has nonnegative diagonal entries. Therefore, we have

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \Lambda w_t \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Lambda \Xi \Lambda'),$$

and the new variance matrix can also been factorized as  $\Lambda \Xi \Lambda' = (\Lambda \Delta)(\Lambda \Delta)' := \Psi \Psi'$ , where  $\Psi \in \mathbb{R}^{m \times (k_0+m)}$ . Thus,  $n^{-1/2} \sum_{t=1}^n (\Psi \Psi')^{-1/2} \Lambda w_t \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, I_m)$  where  $I_m$  is the identity matrix of order  $m$ . The above arguments also apply to matrix  $\Xi_k$  with some matrix  $\Psi_k$  which is defined analogously as  $\Psi$ . Consequently,

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \Lambda w_t^k \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Lambda \Xi_k \Lambda'),$$

and we also have  $n^{-1/2} \sum_{t=1}^n (\Psi_k \Psi_k')^{-1/2} \Lambda w_t^k \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, I_m)$

Now we are able to apply the functional central limit theorem for strongly mixing process of [20]. We have for any  $r \in (0, 1)$ ,

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor nr \rfloor} (\Psi_k \Psi_k')^{-1/2} \Lambda w_j^k \xrightarrow[n \rightarrow \infty]{\mathbb{D}^m} B_m(r).$$

For all  $j \in \{1, \dots, \lfloor nr \rfloor\}$ , we write

$$(\Psi \Psi')^{-1/2} \Lambda w_j^k = \left( (\Psi \Psi')^{-1/2} - (\Psi_k \Psi_k')^{-1/2} \right) \Lambda w_j^k + (\Psi_k \Psi_k')^{-1/2} \Lambda w_j^k$$



and we obtain that

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor nr \rfloor} (\Psi \Psi')^{-1/2} \Lambda w_j^k \xrightarrow[n \rightarrow \infty]{\mathbb{D}^m} B_m(r).$$

In order to conclude (2.18), it remains to observe that, uniformly with respect to  $n$ ,

$$Z_n^k(r) := \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor nr \rfloor} (\Psi \Psi')^{-1/2} \Lambda Y_j^k \xrightarrow[k \rightarrow \infty]{\mathbb{D}^m} 0, \quad (2.21)$$

where

$$Y_t^k = \left( -2\sigma_0^{-2} \left\{ \sum_{i=k+1}^{\infty} \lambda_{i,1}(\theta_0) \epsilon_t \epsilon_{t-i}, \dots, \sum_{i=k+1}^{\infty} \lambda_{i,k_0}(\theta_0) \epsilon_t \epsilon_{t-i} \right\}' J^{-1'}, \epsilon_t \epsilon_{t-1}, \dots, \epsilon_t \epsilon_{t-m} \right)'.$$

By Lemma 4 in [15], we have

$$\sup_n \text{Var} \left( \frac{1}{\sqrt{n}} \sum_{j=1}^n Y_j^k \right) \xrightarrow[k \rightarrow \infty]{} 0$$

and since  $\lfloor nr \rfloor \leq n$ ,

$$\sup_{0 \leq r \leq 1} \sup_n \left( \|Z_n^k(r)\| \right) \xrightarrow[k \rightarrow \infty]{} 0.$$

Thus (2.21) is true and the proof of the first step (2.18) is achieved.

The previous step ensures us that Assumption 1 in [23] is satisfied for the sequence  $(\Lambda w_t)_{t \geq 1}$ . We follow the arguments developed in Sections 2 and 3 in [23], the second step is to show that

$$C_m = \frac{1}{n^2} \sum_{t=1}^n S_t S_t' \xrightarrow[n \rightarrow \infty]{d} (\Psi \Psi')^{1/2} V_m (\Psi \Psi')^{1/2}, \quad (2.22)$$

by applying the continuous mapping theorem on the Skorokhod space and where the random variable  $V_m$  is defined in (2.9). The main issue is to obtain that

$$\begin{aligned} \frac{1}{\sqrt{n}} S_{\lfloor nr \rfloor} &= \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor nr \rfloor} (\Lambda w_j - \Lambda \bar{w}) = \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor nr \rfloor} \Lambda w_j - \frac{\lfloor nr \rfloor}{n} \left( \frac{1}{\sqrt{n}} \sum_{j=1}^n \Lambda w_j \right) \\ &\xrightarrow[n \rightarrow \infty]{\mathbb{D}^m} (\Psi \Psi')^{1/2} B_m(r) - r (\Psi \Psi')^{1/2} B_m(1), \end{aligned} \quad (2.23)$$

by using (2.18) and the fact that  $\lfloor nr \rfloor/n \rightarrow r$  as  $n \rightarrow \infty$ . In view of (2.23), it follows that by continuous mapping theorem

$$\begin{aligned} C_m &= \frac{1}{n^2} \sum_{t=1}^n S_t S_t' = \frac{1}{n} \sum_{t=1}^n \left( \frac{1}{\sqrt{n}} S_t \right) \left( \frac{1}{\sqrt{n}} S_t \right)' \\ &\xrightarrow[n \rightarrow \infty]{d} (\Psi \Psi')^{1/2} \left( \int_0^1 (B_m(r) - r B_m(1)) (B_m(r) - r B_m(1))' dr \right) (\Psi \Psi')^{1/2} \\ &= (\Psi \Psi')^{1/2} V_m (\Psi \Psi')^{1/2}, \end{aligned}$$

which prove (2.22). Since  $\sqrt{n} \hat{\gamma}_m = n^{-1/2} \sum_{t=1}^n \Lambda w_t + o_{\mathbb{P}}(1)$ , using (2.18) and (2.22) we obtain

$$\begin{aligned} n \hat{\gamma}_m' C_m^{-1} \hat{\gamma}_m &= \left( \frac{1}{\sqrt{n}} \sum_{t=1}^n \Lambda w_t + o_{\mathbb{P}}(1) \right)' C_m^{-1} \left( \frac{1}{\sqrt{n}} \sum_{t=1}^n \Lambda w_t + o_{\mathbb{P}}(1) \right) \\ &\xrightarrow[n \rightarrow \infty]{d} \left( [\Psi \Psi']^{1/2} B_m(1) \right)' \left( [\Psi \Psi']^{1/2} V_m \left\{ [\Psi \Psi']^{1/2} \right\}' \right)^{-1} \left( [\Psi \Psi']^{1/2} B_m(1) \right) \\ &= B_m'(1) V_m^{-1} B_m(1) =: \mathcal{U}_m. \end{aligned}$$

The proof of Theorem 2.3.3 is then complete.  $\square$

**Proof of Theorem 2.3.4:**

We write  $\hat{C}_m = C_m + \nabla_n$  where  $\nabla_n = n^{-2} \sum_{t=1}^n (S_t S_t' - \hat{S}_t \hat{S}_t')$ . There are three kinds of entries in the matrix  $\nabla_n$ . The first one is a sum composed of

$$v_t^{k,k'} = \epsilon_t^2(\theta_0) \epsilon_{t-k}(\theta_0) \epsilon_{t-k'}(\theta_0) - e_t^2(\hat{\theta}_n) e_{t-k}(\hat{\theta}_n) e_{t-k'}(\hat{\theta}_n)$$

for  $(k, k') \in \{1, \dots, m\}^2$ . Using (2.17) and the consistency of  $\hat{\theta}_n$ , we have  $v_t^{k,k'} = o(1)$  almost surely. The two last kinds of entries of  $\nabla_n$  come from the following quantities for  $i, j \in \{1, \dots, k_0\}$  and  $k \in \{1, \dots, m\}$

$$\begin{aligned} \tilde{v}_t^{k,i} &= \epsilon_t^2(\theta_0) \epsilon_{t-k}(\theta_0) \frac{\partial \epsilon_t(\theta_0)}{\partial \theta_i} - e_t^2(\hat{\theta}_n) e_{t-k}(\hat{\theta}_n) \frac{\partial e_t(\hat{\theta}_n)}{\partial \theta_i}, \\ \tilde{v}_t^{i,j} &= \epsilon_t^2(\theta_0) \frac{\partial \epsilon_t(\theta_0)}{\partial \theta_i} \frac{\partial \epsilon_t(\theta_0)}{\partial \theta_j} - e_t^2(\hat{\theta}_n) \frac{\partial e_t(\hat{\theta}_n)}{\partial \theta_i} \frac{\partial e_t(\hat{\theta}_n)}{\partial \theta_j} \end{aligned}$$

and they also satisfy  $\tilde{v}_t^{k,i} + \tilde{v}_t^{i,j} = o(1)$  almost surely by using (2.14) and (2.17). Consequently,  $\nabla_n = o(1)$  almost surely as  $n$  goes to infinity. Thus one may find a matrix  $\nabla_n^*$ , that tends to the null matrix almost surely, such that

$$n \hat{Y}_m' \hat{C}_m^{-1} \hat{Y}_m = n \hat{Y}_m' (C_m + \nabla_n)^{-1} \hat{Y}_m = n \hat{Y}_m' C_m^{-1} \hat{Y}_m + n \hat{Y}_m' \nabla_n^* \hat{Y}_m.$$

Thanks to the arguments developed in the proof of Theorem 2.3.3,  $n \hat{Y}_m' C_m^{-1} \hat{Y}_m$  converges in distribution. So  $n \hat{Y}_m' \nabla_n^* \hat{Y}_m$  tends to zero in distribution, hence in probability. Then  $n \hat{Y}_m' \hat{C}_m^{-1} \hat{Y}_m$  and  $n \hat{Y}_m' C_m^{-1} \hat{Y}_m$  have the same limit in distribution and the result is proved.  $\square$

## Bibliography

- [1] Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3):817–858. [58](#), [59](#)
- [2] Berk, K. N. (1974). Consistent autoregressive spectral estimates. *Ann. Statist.*, 2:489–502. Collection of articles dedicated to Jerzy Neyman on his 80th birthday. [58](#), [59](#)
- [3] Boubacar Mainassara, Y. (2011). Multivariate portmanteau test for structural VARMA models with uncorrelated but non-independent error terms. *J. Statist. Plann. Inference*, 141(8):2961–2975. [59](#)
- [4] Boubacar Mainassara, Y. (2014). Estimation of the variance of the quasi-maximum likelihood estimator of weak VARMA models. *Electron. J. Stat.*, 8(2):2701–2740. [69](#)
- [5] Boubacar Mainassara, Y. and Francq, C. (2011). Estimating structural VARMA models with uncorrelated but non-independent error terms. *J. Multivariate Anal.*, 102(3):496–505. [59](#)
- [6] Boubacar Mainassara, Y. and Ilmi Amir, A. (2018). Multivariate portmanteau tests for weak multiplicative seasonal varma models. *Stat. Pap.*, pages 1–32. [54](#), [55](#), [56](#), [58](#)
- [7] Boubacar Mainassara, Y. and Ilmi Amir, A. (2020). Multivariate portmanteau tests for weak multiplicative seasonal VARMA models. *Statist. Papers*, 61(6):2529–2560. [58](#), [59](#)
- [8] Boubacar Mainassara, Y. and Sausseureau, B. (2018). Diagnostic checking in multivariate arma models with dependent errors using normalized residual autocorrelations. *J. Amer. Statist. Assoc.*, 113(524):1813–1827. [54](#), [55](#), [59](#), [69](#)
- [9] Box, G. E. P. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J. Amer. Statist. Assoc.*, 65:1509–1526. [54](#)
- [10] Davidson, J. (1994). *Stochastic limit theory*. Advanced Texts in Econometrics. The Clarendon Press, Oxford University Press, New York. An introduction for econometricians. [70](#)
- [11] Davydov, J. A. (1968). The convergence of distributions which are generated by stationary random processes. *Teor. Veroyatnost. i Primenen.*, 13:730–737. [56](#), [57](#), [70](#)
- [12] den Haan, W. J. and Levin, A. T. (1997). A practitioner’s guide to robust covariance matrix estimation. In *Robust inference*, volume 15 of *Handbook of Statist.*, pages 299–342. North-Holland, Amsterdam. [58](#), [59](#)
- [13] Duchesne, P. (2007). On consistent testing for serial correlation in seasonal time series models. *Canad. J. Statist.*, 35(2):193–213. [54](#)
- [14] Francq, C., Roy, R., and Zakoïan, J.-M. (2005). Diagnostic checking in ARMA models with uncorrelated errors. *J. Amer. Statist. Assoc.*, 100(470):532–544. [54](#), [69](#), [70](#)
- [15] Francq, C. and Zakoïan, J.-M. (1998). Estimating linear representations of nonlinear processes. *J. Statist. Plann. Inference*, 68(1):145–165. [56](#), [69](#), [70](#), [71](#)
- [16] Francq, C. and Zakoïan, J.-M. (2000). Covariance matrix estimation for estimators of mixing weak ARMA models. *J. Statist. Plann. Inference*, 83(2):369–394. [69](#)
- [17] Francq, C. and Zakoïan, J.-M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10(4):605–637. [68](#)

- [18] Francq, C. and Zakoïan, J.-M. (2005). Recent results for linear time series models with non independent innovations. In *Statistical modeling and analysis for complex data problems*, volume 1 of *GERAD 25th Anniv. Ser.*, pages 241–265. Springer, New York. [54](#), [56](#)
- [19] Francq, C. and Zakoïan, J.-M. (2010). *GARCH Models: Structure, Statistical Inference and Financial Applications*. Wiley. [61](#), [68](#)
- [20] Herrndorf, N. (1984). A functional central limit theorem for weakly dependent sequences of random variables. *Ann. Probab.*, 12(1):141–153. [57](#), [70](#)
- [21] Kuan, C.-M. and Lee, W.-M. (2006). Robust M tests without consistent estimation of the asymptotic covariance matrix. *J. Amer. Statist. Assoc.*, 101(475):1264–1275. [59](#)
- [22] Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2):pp. 297–303. [54](#)
- [23] Lobato, I. N. (2001). Testing that a dependent process is uncorrelated. *J. Amer. Statist. Assoc.*, 96(455):1066–1076. [59](#), [61](#), [71](#)
- [24] McLeod, A. I. (1978). On the distribution of residual autocorrelations in Box-Jenkins models. *J. Roy. Statist. Soc. Ser. B*, 40(3):296–302. [54](#), [61](#), [69](#)
- [25] Newey, W. K. and West, K. D. (1987). A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708. [58](#), [59](#)
- [26] Romano, J. P. and Thombs, L. A. (1996). Inference for autocorrelations under weak assumptions. *J. Amer. Statist. Assoc.*, 91(434):590–600. [56](#)
- [27] Shao, X. (2010a). A self-normalized approach to confidence interval construction in time series. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 72(3):343–366. [59](#)
- [28] Shao, X. (2010b). Corrigendum: A self-normalized approach to confidence interval construction in time series. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 72(5):695–696. [59](#)
- [29] Shao, X. (2012). Parametric inference in stationary time series models with dependent errors. *Scand. J. Stat.*, 39(4):772–783. [59](#)
- [30] Shao, X. (2015). Self-normalization for time series: a review of recent developments. *J. Amer. Statist. Assoc.*, 110(512):1797–1817. [59](#)
- [31] Zhu, K. (2013). A mixed portmanteau test for ARMA-GARCH models by the quasi-maximum exponential likelihood estimation approach. *J. Time Series Anal.*, 34(2):230–237. [68](#)

# Multivariate portmanteau tests for weak multiplicative seasonal VARMA models

## Abstract

Numerous multivariate time series encountered in real applications display seasonal behavior. In this paper we consider portmanteau tests for testing the adequacy of structural multiplicative seasonal vector autoregressive moving-average (SVARMA) models under the assumption that the errors are uncorrelated but not necessarily independent (*i.e.* weak SVARMA). We study the asymptotic distributions of residual autocorrelations at seasonal lags of multiple of the length of the seasonal period under weak assumptions on the noise. We deduce the asymptotic distribution of the proposed multivariate portmanteau statistics, which can be quite different from the usual chi-squared approximation used under independent and identically distributed (iid) assumptions on the noise. A set of Monte Carlo experiments is presented.

**Keywords:** Goodness-of-fit test, quasi-maximum likelihood estimation, portmanteau tests, residual autocorrelation, weak SVARMA models

## 3.1 Introduction

Seasonality naturally occurs in many time series coming from various fields of study, such as meteorology, hydrology and economics, amongst others. In univariate time series, the multiplicative seasonal autoregressive moving average (SARMA), denoted  $\text{SARMA}(p, q)(p_s, q_s)$ , time series model is widely used in the modeling of seasonal time series (see Chapter 12 in [20]), where  $s$  denotes the length of the seasonal period. In econometric application, the univariate SARMA framework is very restrictive. Consequently the class of multivariate models are commonly used in time series analysis and econometrics. It describes the possible cross-relationships between the time series and not only the properties of the individual time series (see [29]).

Consider a  $d$ -dimensional stationary process  $X = (X_t)_{t \in \mathbb{Z}}$ , satisfying a multiplicative seasonal vectorial autoregressive moving average (SVARMA) representation of the form

$$\Lambda_{\theta_0}(L^s)A_{\theta_0}(L)X_t = \Psi_{\theta_0}(L^s)B_{\theta_0}(L)\epsilon_t, \quad \forall t \in \mathbb{Z}, \quad (3.1)$$

where the nonseasonal AR and MA operators are defined by  $A_{\theta_0}(L) = A_{00} - \sum_{i=1}^p A_{0i}L^i$  and  $B_{\theta_0}(L) = B_{00} - \sum_{i=1}^q B_{0i}L^i$ , respectively, while the seasonal AR and MA operators are given by  $\Lambda_{\theta_0}(L) = \Lambda_{00} - \sum_{i=1}^{p_s} \Lambda_{0i}L^{si}$  and  $\Psi_{\theta_0}(L) = \Psi_{00} - \sum_{i=1}^{q_s} \Psi_{0i}L^{si}$ , respectively and  $L$  stands for the backshift operator. The true value of the unknown parameter is denoted  $\theta_0$  and is defined in the sequel. Without loss of generality, assume that matrix  $A_{0p} \neq 0$  or  $B_{0q} \neq 0$  or  $\Lambda_{0p_s} \neq 0$  or  $\Psi_{0q_s} \neq 0$ . Note

that, in general, the matrix seasonal and nonseasonal AR and MA operators do not commute, and hence the order in which these operators are presented in the representation (3.1) will make a difference.

The representation (3.1), denoted  $\text{SVARMA}(p, q)(p_s, q_s)$ , is said to satisfy a weak SVARMA representation if  $\epsilon_t$  is a *weak white noise*, that is, if it satisfies

$$\mathbf{A0:} \quad \mathbb{E}(\epsilon_t) = 0, \text{Var}(\epsilon_t) = \Sigma_0 \text{ and } \text{Cov}(\epsilon_t, \epsilon_{t-h}) = 0 \text{ for all } t \in \mathbb{Z} \text{ and all } h \neq 0,$$

where the variance  $\Sigma_0$  is assumed to be non singular. It is customary to say that  $(X_t)$  is a strong SVARMA model if  $(\epsilon_t)$  is assumed to be a sequence of independent and identically distributed (iid for short) a sequence of random variables with mean 0 and common variance matrix (*i.e. strong white noise*). A strong white noise is obviously a weak white noise, because independence entails uncorrelatedness, but the reverse is not true.

Nonlinear models are becoming more and more employed because numerous real time series exhibit nonlinear dynamics, for instance conditional heteroscedasticity, which can not be generated by autoregressive moving-average (ARMA) models with iid noises. The main issue with nonlinear models is that they are generally hard to identify and implement. This is why it is interesting to consider weak VARMA models, that is VARMA models with weak white noises, such linear representations being universal approximations of the Wold decomposition. Linear and nonlinear multivariate processes also have exact weak VARMA representations because a same process may satisfy several models, and many important classes of nonlinear processes admit weak VARMA representations (see for instance [11, 15, 16, 17]).

Note that in the VARMA case, a variety of examples of nonlinear processes admitting weak VARMA representations in the weak sense

The class of SARMA models has been extensively studied in the literature. However, as noted by [33] (see p. 219), there has been much less investigation of seasonal modeling designed for multivariate data, by comparison with the univariate case. After identification and estimation of the SVARMA processes, the next important step in the SVARMA modeling consists in checking if the estimated model fits satisfactorily the data. This adequacy checking step allows to validate or invalidate the choice of the orders  $p$ ,  $q$ ,  $p_s$  and  $q_s$ . This step is particularly important because the number of parameters can be quickly increasing with  $p$ ,  $q$ ,  $p_s$  and  $q_s$ , which entails statistical difficulties.

Thus it is important to check the validity of a SVARMA model, for given orders  $p$ ,  $q$ ,  $p_s$  and  $q_s$ . This paper is devoted to the problem of the validation step of SVARMA representations.

Being inspired by the univariate portmanteau statistics introduced by [8, 28], [12] and [21] have introduced the following multivariate versions of the portmanteau statistics

$$\begin{aligned} Q_m^C &= n \sum_{h=1}^m \text{Tr}(\hat{\Gamma}'(h)\hat{\Gamma}^{-1}(0)\hat{\Gamma}(h)\hat{\Gamma}^{-1}(0)) \text{ and} \\ Q_m^H &= \sum_{h=1}^m \frac{n^2}{(n-h)} \text{Tr}(\hat{\Gamma}'(h)\hat{\Gamma}^{-1}(0)\hat{\Gamma}(h)\hat{\Gamma}^{-1}(0)), \end{aligned} \tag{3.2}$$

where  $n$  is the length of the series,  $m$  is a fixed integer and  $\hat{\Gamma}(h)$  is the residual autocovariances matrices function of the multivariate process  $\epsilon$ ,  $\text{Tr}(\cdot)$  denotes the trace of a matrix and  $A'$  the transpose of a matrix  $A$ . See also [22, 23] who gave several equivalent forms of the statistics (3.2). The statistics (3.2) are usually used to check the overall significance of the residual autocorrelations of a VARMA model (see [1, 27]).

In many situations, the statistics (3.2) are implemented to check the lack of fit of SVARMA models. However, as noted by [31] (see p. 301), the methods of [12] and [21] cannot be extended to the case of SVARMA models when  $s > 1$  because of the multiplicative constraints on the parameters. Using such tests for SVARMA models would be misleading and not enough as these tests consider the autocorrelations corresponding to the nonseasonal lags  $\leq m$  and ignore



the possibility of autocorrelations at seasonal lags of multiple period  $s$ . Motivated by the arguments given by [31] (see p. 301), we will propose a modification of the multivariate portmanteau tests (3.2) by using seasonal lags  $1s, 2s, 3s, \dots, ms$  to check the adequacy of seasonal component. Under the assumption that the noise sequence is iid, the standard tests procedure consist in rejecting the null hypothesis of a SVARMA( $p, q$ )( $p_s, q_s$ ) model if the statistics (3.2) are larger than a certain quantile of a chi-squared distribution with  $m - (p + q + p_s + q_s) > 0$  degrees of freedom (see, for instance, Section 4 of [35]). Consequently, these standard tests are not applicable for  $m \leq p + q + p_s + q_s$ . Contrarily to the standard tests (3.2), the proposed tests can be used safely for  $m$  small. The asymptotic theory of weak SVARMA model validation is mainly limited to the weak VARMA framework (for instance, see [5, 7, 26] for multivariate nonlinear case and [10, 15, 34] for univariate case). Other contribution is to improve the results concerning the statistical analysis of weak SVARMA models by considering the adequacy problem.

The works on the multivariate version of the portmanteau statistic are generally performed under the assumption that the errors  $\epsilon_t$  are independent. For instance, see [1, 27] and [35] who study the asymptotic distribution of the residual autocorrelations in strong VARMA and SVARMA cases, respectively. This independence assumption is often considered too restrictive by practitioners. It precludes conditional heteroscedasticity and/or other forms of nonlinearity. Relaxing this independence assumption allows to extend the range of application of the SVARMA models.

The main goal of the present paper is to complete the available results concerning the statistical analysis of weak SVARMA models by considering the adequacy problem under general error terms. We proceed to study the behaviour of the goodness-of fit portmanteau tests when the  $\epsilon_t$  are not independent. It is shown that the standard portmanteau tests can be quite misleading in the framework of non independent errors. Consequently, a modified version of these tests is proposed.

The article is organised as follows. In the next section, we briefly present the results on the quasi-maximum likelihood estimator (QMLE) of the model considered here. We give mild conditions ensuring the consistency and asymptotic normality of the QMLE. In Section 3.3, we modify the usual portmanteau test statistics defined by (3.2) to the SVARMA class. We study the asymptotic behaviour of the residuals autocovariances and autocorrelations under weak assumptions on the noise in Section 3.3. It is also shown how the proposed portmanteau tests (3.6) must be adapted in the case of multiplicative seasonal VARMA models with non-independent innovations. Section 3.4 proposes numerical illustrations. The technical proofs are relegated to the appendix.

## 3.2 Estimating weak SVARMA models

Let  $I_d$  be the  $d \times d$  identity matrix of order  $d$ . When  $A_{00} = B_{00} = \Lambda_{00} = \Psi_{00} = I_d$ , the SVARMA representation (3.1) is said to be in reduced form. Otherwise, it is said to be structural. The structural forms are mainly used in econometrics to introduce instantaneous relationships between economic variables. Of course, constraints are necessary for the identifiability of these representations.

Let  $[A_{00} \dots A_{0p} B_{00} \dots B_{0q} \Lambda_{00} \dots \Lambda_{0p_s} \Psi_{00} \dots \Psi_{0q_s}]$  be the  $d^2(p + q + p_s + q_s + 4)$  matrix of VAR and MA coefficients involved in the representation (3.1). The matrix  $\Sigma_0$  is considered as a nuisance parameter. The parameter of interest denoted  $\theta_0$  belongs to the parameter space  $\Theta \subset \mathbb{R}^{k_0}$ , and  $k_0$  is the number of unknown parameters, which is typically much smaller than  $d^2(p + q + p_s + q_s + 4)$ . The matrices  $A_{00}, \dots, A_{0p}, B_{00}, \dots, B_{0q}, \Lambda_{00}, \dots, \Lambda_{0p_s}, \Psi_{00}, \dots, \Psi_{0q_s}$  involved in (3.1) are specified by  $\theta_0$ . More precisely, we write  $A_{0i} = A_i(\theta_0)$ ,  $B_{0j} = B_j(\theta_0)$ ,  $\Lambda_{0k} = \Lambda_k(\theta_0)$  and  $\Psi_{0\ell} = \Psi_\ell(\theta_0)$  for  $i = 0, \dots, p$ ,  $j = 0, \dots, q$ ,  $k = 0, \dots, p_s$  and  $\ell = 0, \dots, q_s$ . Assume that the parametrization satisfies the following smoothness conditions.



**A1:** The applications  $\theta \mapsto A_i(\theta)$   $i = 0, \dots, p$ ,  $\theta \mapsto B_j(\theta)$   $j = 0, \dots, q$ ,  $\theta \mapsto \Lambda_k(\theta)$   $k = 0, \dots, p_s$  and  $\theta \mapsto \Psi_\ell(\theta)$   $\ell = 0, \dots, q_s$  admit continuous third order derivatives for all  $\theta \in \Theta$ .

For simplicity we now write  $A_i$ ,  $B_j$ ,  $\Lambda_k$  and  $\Psi_\ell$  instead of  $A_i(\theta)$ ,  $B_j(\theta)$ ,  $\Lambda_k(\theta)$  and  $\Psi_\ell(\theta)$ . Let  $A_\theta(z) = A_0 - \sum_{i=1}^p A_i z^i$ ,  $B_\theta(z) = B_0 - \sum_{i=1}^q B_i z^i$ ,  $\Lambda_\theta(z) = \Lambda_0 - \sum_{i=1}^{p_s} \Lambda_i z^{s_i}$  and  $\Psi_\theta(z) = \Psi_0 - \sum_{i=1}^{q_s} \Psi_i z^{s_i}$ . We assume that  $\Theta$  corresponds to stable and invertible representations, namely

**A2:** for all  $\theta \in \Theta$ , we have  $\det A_\theta(z) \det B_\theta(z) \det \Lambda_\theta(z) \det \Psi_\theta(z) \neq 0$  for all  $|z| \leq 1$ , where  $\det(C)$  stands for the determinant of the square matrix  $C$ .

### 3.2.1 Quasi-maximum likelihood estimation

For the estimation of weak SARMA and SVARMA models, the commonly used estimation method is the quasi-maximum likelihood estimation, which can be also viewed as a nonlinear least squares estimation (LSE). Let  $X_1, \dots, X_n$  be observations of a process satisfying the SVARMA representation (3.1). Note that from **A2** the matrices  $A_{00}$ ,  $B_{00}$ ,  $\Lambda_{00}$  and  $\Psi_{00}$  are invertible. Introducing the innovation process

$$e_t = A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{00} B_{00} \epsilon_t,$$

the structural representation (3.1) can be rewritten as the reduced SVARMA representation

$$\begin{aligned} X_t & - \sum_{i=1}^p A_{00}^{-1} A_{0i} X_{t-i} - \sum_{j=1}^{p_s} A_{00}^{-1} \Lambda_{00}^{-1} \Lambda_{0j} A_{00} X_{t-s_j} + \sum_{i=1}^p \sum_{j=1}^{p_s} A_{00}^{-1} \Lambda_{00}^{-1} \Lambda_{0j} A_{0i} X_{t-i-s_j} \\ & = e_t - \sum_{i=1}^q A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{00} B_{0i} B_{00}^{-1} \Psi_{00}^{-1} \Lambda_{00} A_{00} e_{t-i} - \sum_{j=1}^{q_s} A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{0j} \Psi_{00}^{-1} \Lambda_{00} A_{00} e_{t-s_j} \\ & + \sum_{i=1}^q \sum_{j=1}^{q_s} A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{0j} B_{0i} B_{00}^{-1} \Psi_{00}^{-1} \Lambda_{00} A_{00} e_{t-i-s_j}. \end{aligned} \quad (3.3)$$

For all  $\theta \in \Theta$ , we recursively define  $\tilde{e}_t(\theta)$  for  $t = 1, \dots, n$  by

$$\begin{aligned} \tilde{e}_t(\theta) & = X_t - \sum_{i=1}^p A_0^{-1} A_i X_{t-i} - \sum_{j=1}^{p_s} A_0^{-1} \Lambda_0^{-1} \Lambda_j A_0 X_{t-s_j} + \sum_{i=1}^p \sum_{j=1}^{p_s} A_0^{-1} \Lambda_0^{-1} \Lambda_j A_i X_{t-i-s_j} \\ & + \sum_{i=1}^q A_0^{-1} \Lambda_0^{-1} \Psi_0 B_i B_0^{-1} \Psi_0^{-1} \Lambda_0 A_0 \tilde{e}_{t-i}(\theta) + \sum_{j=1}^{q_s} A_0^{-1} \Lambda_0^{-1} \Psi_j \Psi_0^{-1} \Lambda_0 A_0 \tilde{e}_{t-s_j}(\theta) \\ & - \sum_{i=1}^q \sum_{j=1}^{q_s} A_0^{-1} \Lambda_0^{-1} \Psi_j B_i B_0^{-1} \Psi_0^{-1} \Lambda_0 A_0 \tilde{e}_{t-i-s_j}(\theta). \end{aligned} \quad (3.4)$$

with initial values  $\tilde{e}_0(\theta) = \tilde{e}_{-1}(\theta) = \dots = \tilde{e}_{-q-q_s+1}(\theta) = X_0 = X_{-1} = \dots = X_{-p-p_s+1} = 0$ . It will be shown that these initial values are asymptotically negligible and, in particular, that  $\tilde{e}_t(\theta) - e_t \rightarrow 0$  almost surely as  $t \rightarrow \infty$  (see [6]). The Gaussian quasi-likelihood is given by

$$\tilde{L}_n(\theta, \Sigma_e) = \prod_{t=1}^n \frac{1}{(2\pi)^{d/2} \sqrt{\det \Sigma_e}} \exp \left\{ -\frac{1}{2} \tilde{e}'_t(\theta) \Sigma_e^{-1} \tilde{e}_t(\theta) \right\},$$

where

$$\Sigma_e = \Sigma_e(\theta) = A_0^{-1} \Lambda_0^{-1} \Psi_0 B_0 \Sigma B_0' \Psi_0' \Lambda_0^{-1'} A_0^{-1'}.$$

Note that the variance of  $e_t$  is  $\Sigma_{e0} = \Sigma_e(\theta_0) = A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{00} B_{00} \Sigma_0 B_{00}' \Psi_{00}' \Lambda_{00}^{-1'} A_{00}^{-1'}$ . A quasi-maximum likelihood estimator (QMLE) of  $\theta$  and  $\Sigma_e$  are a measurable solution  $(\hat{\theta}_n, \hat{\Sigma}_e)$  of

$$(\hat{\theta}_n, \hat{\Sigma}_e) = \arg \max_{\theta \in \Theta, \Sigma_e} \tilde{L}_n(\theta, \Sigma_e) = \arg \min_{\theta \in \Theta, \Sigma_e} \tilde{\ell}_n(\theta, \Sigma_e), \quad \tilde{\ell}_n(\theta, \Sigma_e) = \frac{-2}{n} \log \tilde{L}_n(\theta, \Sigma_e).$$

### 3.2.2 Asymptotic properties

To establish the asymptotic properties of the QMLE,  $\hat{\theta}_n$ , we will use the following assumptions.

**A3:** We have  $\theta_0 \in \Theta$ , where  $\Theta$  is compact.

**A4:** The process  $(\epsilon_t)$  is stationary and ergodic.

**A5:** For all  $\theta \in \Theta$  such that  $\theta_0 \neq \theta$ , either the transfer functions

$$A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{00} B_{00} B_{00}^{-1}(z) \Psi_{\theta_0}^{-1}(z^s) \Lambda_{\theta_0}(z^s) A_{\theta_0}(z) \neq A_0^{-1} \Lambda_0^{-1} \Psi_0 B_0 B_0^{-1}(z) \Psi_{\theta}^{-1}(z^s) \Lambda_{\theta}(z^s) A_{\theta}(z)$$

$$\text{for some } z \in \mathbb{C}, \text{ or } A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{00} B_{00} \Sigma_0 B_{00}' \Psi_{00}' \Lambda_{00}^{-1'} A_{00}^{-1'} \neq A_0^{-1} \Lambda_0^{-1} \Psi_0 B_0 \Sigma_0 B_0' \Psi_0' \Lambda_0^{-1'} A_0^{-1'}.$$

**Remark 3.2.1** In the reduced SVARMA representation (3.3), the condition

$$A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{00} B_{00} \Sigma_0 B_{00}' \Psi_{00}' \Lambda_{00}^{-1'} A_{00}^{-1'} \neq A_0^{-1} \Lambda_0^{-1} \Psi_0 B_0 \Sigma_0 B_0' \Psi_0' \Lambda_0^{-1'} A_0^{-1'}$$

in **A5** can be dropped but may be important for structural SVARMA forms (3.1). In particular, **A5** is satisfied when we impose that  $A_0 = B_0 = \Lambda_0 = \Psi_0 = I_d$ , that **A2** holds, that the common left divisors of  $A_{\theta}(L)$ ,  $\Lambda_{\theta}(L^s)$ ,  $B_{\theta}(L)$  and  $\Psi_{\theta}(L^s)$  are unimodular (*i.e.* with nonzero constant determinant), and finally that the matrix  $[A_p : \Lambda_{p_s} : B_q : \Psi_{q_s}]$  is of full rank (see for instance [18] who given several procedures ensuring identifiability).

For the asymptotic normality of the QMLE, additional assumptions are required. It is necessary to assume that  $\theta_0$  is not on the boundary of the parameter space  $\Theta$ .

**A6:** We have  $\theta_0 \in \overset{\circ}{\Theta}$ , where  $\overset{\circ}{\Theta}$  denotes the interior of  $\Theta$ .

To control the serial dependence of the stationary process  $(\epsilon_t)$ , we now introduce (as in [6]) the strong mixing coefficients  $\alpha_{\epsilon}(h)$  defined by

$$\alpha_{\epsilon}(h) = \sup_{A \in \mathcal{F}_{-\infty}^t, B \in \mathcal{F}_{t+h}^{+\infty}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|,$$

where  $\mathcal{F}_{-\infty}^t = \sigma(\epsilon_u, u \leq t)$  and  $\mathcal{F}_{t+h}^{+\infty} = \sigma(\epsilon_u, u \geq t+h)$ . We use  $\|\cdot\|$  to denote the Euclidian norm  $\|Z\| = (\sum_{i=1}^{\tau} z_i^2)^{1/2}$  of a column vector  $Z = (z_1, \dots, z_{\tau})'$ . We will make an integrability assumption on the moment of the noise and a summability condition on the strong mixing coefficients  $(\alpha_{\epsilon}(k))_{k \geq 0}$ .

**A7:** We have  $\mathbb{E}\|\epsilon_t\|^{4+2\nu} < \infty$  and  $\sum_{k=0}^{\infty} \{\alpha_{\epsilon}(k)\}^{\frac{\nu}{2+\nu}} < \infty$  for some  $\nu > 0$ .

Note that assumption **A7** does not require independence of the noise, nor the fact that it is a martingale difference.

We use the matrix  $M_{\theta_0}$  of the coefficients of the reduced form (3.3), where

$$\begin{aligned} M_{\theta_0} = & [A_{00}^{-1} A_{01} : \dots : A_{00}^{-1} A_{0p} : A_{00}^{-1} \Lambda_{00}^{-1} \Lambda_{01} A_{00} : \dots : A_{00}^{-1} \Lambda_{00}^{-1} \Lambda_{0p_s} A_{00} : A_{00}^{-1} \Lambda_{00}^{-1} \Lambda_{01} A_{01} : \\ & \dots : A_{00}^{-1} \Lambda_{00}^{-1} \Lambda_{01} A_{0p} : \dots : A_{00}^{-1} \Lambda_{00}^{-1} \Lambda_{0p_s} A_{01} : \dots : A_{00}^{-1} \Lambda_{00}^{-1} \Lambda_{0p_s} A_{0p} : \\ & A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{00} B_{01} B_{00}^{-1} \Psi_{00}^{-1} \Lambda_{00} A_{00} : \dots : A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{00} B_{0q} B_{00}^{-1} \Psi_{00}^{-1} \Lambda_{00} A_{00} : \\ & A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{01} \Psi_{00}^{-1} \Lambda_{00} A_{00} : \dots : A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{0q_s} \Psi_{00}^{-1} \Lambda_{00} A_{00} : \\ & A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{01} B_{01} B_{00}^{-1} \Psi_{00}^{-1} \Lambda_{00} A_{00} : \dots : A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{01} B_{0q} B_{00}^{-1} \Psi_{00}^{-1} \Lambda_{00} A_{00} : \dots : \\ & A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{0q_s} B_{01} B_{00}^{-1} \Psi_{00}^{-1} \Lambda_{00} A_{00} : \dots : A_{00}^{-1} \Lambda_{00}^{-1} \Psi_{0q_s} B_{0q} B_{00}^{-1} \Psi_{00}^{-1} \Lambda_{00} A_{00}]. \end{aligned}$$

Now, we need a local identifiability assumption which completes **A5** and specifies how this matrix depends on the parameter  $\theta_0$ . We denote by  $A \otimes B$  the Kronecker product of two matrices  $A$  and  $B$ , and by  $\text{vec}(A)$  the vector obtained by stacking the columns of  $A$ . Let  $\dot{M}_{\theta_0}$  be the matrix  $\partial \text{vec}(M_{\theta}) / \partial \theta'$  evaluated at  $\theta_0$ .

**A8:** The matrix  $\dot{M}_{\theta_0}$  is of full rank  $k_0$ .

Verification of Assumptions **A5** and **A8** on the following example.

**Example 3.2.1** Considering the following structural SVARMA(1,0)(1,0)<sub>s</sub> equation

$$\begin{aligned} & \begin{pmatrix} 1 & 0 \\ -\lambda_{03} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -a_{03} & 1 \end{pmatrix} X_t = \epsilon_t + \begin{pmatrix} a_{01} & a_{02} \\ a_{04} & a_{05} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\lambda_{03} & 1 \end{pmatrix} X_{t-1} \\ & + \begin{pmatrix} \lambda_{01} & \lambda_{02} \\ \lambda_{04} & \lambda_{05} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -a_{03} & 1 \end{pmatrix} X_{t-s} - \begin{pmatrix} \lambda_{01} & \lambda_{02} \\ \lambda_{04} & \lambda_{05} \end{pmatrix} \begin{pmatrix} a_{01} & a_{02} \\ a_{04} & a_{05} \end{pmatrix} X_{t-1-s}. \end{aligned}$$

We assume that the two components of  $\epsilon_t$  are uncorrelated, with respective variances  $\sigma_{01}^2$  and  $\sigma_{02}^2$ . We thus have

$$\theta'_0 = (a_{01}, a_{02}, a_{03}, a_{04}, a_{05}, \lambda_{01}, \lambda_{02}, \lambda_{03}, \lambda_{04}, \lambda_{05}).$$

Note that the identifiability condition **A5** is satisfied because for all

$$\theta = (a_1, a_2, a_3, a_4, a_5, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5)' \neq \theta_0$$

we have

$$I_2 - D_1(\theta)z + D_2(\theta)z^s + D_3(\theta)z^{s+1} \neq I_2 - D_1(\theta_0)z + D_2(\theta_0)z^s + D_3(\theta_0)z^{s+1}$$

for some  $z \in \mathbb{C}$ , or

$$\begin{pmatrix} \sigma_1^2 & \sigma_1^2(a_3 + \lambda_3) \\ \sigma_1^2(a_3 + \lambda_3) & \sigma_1^2(a_3 + \lambda_3)^2 + \sigma_2^2 \end{pmatrix} \neq \begin{pmatrix} \sigma_{01}^2 & \sigma_{01}^2(a_{03} + \lambda_{03}) \\ \sigma_{01}^2(a_{03} + \lambda_{03}) & \sigma_{01}^2(a_{03} + \lambda_{03})^2 + \sigma_{02}^2 \end{pmatrix},$$

where

$$\begin{aligned} D_1(\theta) &= \begin{pmatrix} a_1 & a_2 \\ a_1 a_3 + a_4 & a_2 a_3 + a_5 \end{pmatrix} \\ D_2(\theta) &= \begin{pmatrix} a_3 \lambda_2 - \lambda_1 & -\lambda_2 \\ (a_3 + \lambda_3)(a_3 \lambda_2 - \lambda_1) - \lambda_4 + a_3 \lambda_5 & a_2 \lambda_4 - \lambda_5 - (a_3 + \lambda_3)(\lambda_2 - \lambda_1 a_2) \end{pmatrix} \\ D_3(\theta) &= \begin{pmatrix} a_1 \lambda_1 + a_4 \lambda_2 & a_2 \lambda_1 + a_5 \lambda_2 \\ (a_3 + \lambda_3)(a_1 \lambda_1 + a_4 \lambda_2) + a_1 \lambda_4 + a_4 \lambda_5 & a_5 [\lambda_5 + (a_3 + \lambda_3) \lambda_2] \end{pmatrix}. \end{aligned}$$

In this example, we have

$$M_{\theta_0} = [D_1(\theta_0) : D_2(\theta_0) : D_3(\theta_0)].$$

Thus

$$\dot{M}_{\theta_0} = \begin{pmatrix} 1 & a_{03} & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{01} & \dot{M}_{\theta_0}(1,10) & 0 & 0 \\ 0 & 0 & 1 & a_{03} & 0 & 0 & 0 & \dot{M}_{\theta_0}(2,8) & 0 & 0 & \lambda_{01} & 0 \\ 0 & a_{01} & 0 & a_{02} & \lambda_{02} & \dot{M}_{\theta_0}(3,6) & 0 & \dot{M}_{\theta_0}(3,8) & 0 & \dot{M}_{\theta_0}(3,10) & 0 & a_{05} \lambda_{02} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{02} & \dot{M}_{\theta_0}(4,10) & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{02} & \dot{M}_{\theta_0}(5,12) \\ 0 & 0 & 0 & 0 & 1 & \dot{M}_{\theta_0}(6,6) & 0 & \dot{M}_{\theta_0}(6,8) & a_{01} & \dot{M}_{\theta_0}(6,10) & a_{02} & 0 \\ 0 & 0 & 0 & 0 & a_{03} & \dot{M}_{\theta_0}(7,6) & -1 & \dot{M}_{\theta_0}(7,8) & a_{04} & \dot{M}_{\theta_0}(7,10) & a_{05} & \dot{M}_{\theta_0}(7,12) \\ 0 & 0 & 0 & 0 & 0 & \dot{M}_{\theta_0}(8,6) & 0 & \dot{M}_{\theta_0}(8,8) & 0 & \dot{M}_{\theta_0}(8,10) & 0 & \lambda_{02} a_{05} \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & a_{02} & 0 & a_{01} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_{03} & 0 & -1 & 0 & a_{04} & 0 & a_{05} \end{pmatrix}$$

is of full rank  $k_0 = 10$  (Assumption **A8** is satisfied), and with  $\dot{M}_{\theta_0}(1,10) = (a_{03} + \lambda_{03})\lambda_{01} + \lambda_{04}$ ,  $\dot{M}_{\theta_0}(2,8) = (a_{03} + \lambda_{03})\lambda_{01} + \lambda_{04}$ ,  $\dot{M}_{\theta_0}(3,6) = -\lambda_{01} + 2a_{03}\lambda_{02} + \lambda_{02}\lambda_{03} + \lambda_{05}$ ,  $\dot{M}_{\theta_0}(3,8) = -(\lambda_{02} - a_{02}\lambda_{01})$ ,  $\dot{M}_{\theta_0}(3,10) = (a_{01}\lambda_{01} + a_{04}\lambda_{02})$ ,  $\dot{M}_{\theta_0}(4,10) = (a_{03} + \lambda_{03})\lambda_{02} + \lambda_{05}$ ,  $\dot{M}_{\theta_0}(5,12) = (a_{03} + \lambda_{03})\lambda_{02} + \lambda_{05}$ ,  $\dot{M}_{\theta_0}(6,6) = -a_{03} - \lambda_{03}$ ,  $\dot{M}_{\theta_0}(6,8) = (a_{03} + \lambda_{03})a_{02}$ ,  $\dot{M}_{\theta_0}(6,10) = (a_{03} + \lambda_{03})a_{01}$ ,  $\dot{M}_{\theta_0}(7,6) = a_{03}(a_{03} + \lambda_{03})$ ,  $\dot{M}_{\theta_0}(7,8) = -a_{03} - \lambda_{03}$ ,  $\dot{M}_{\theta_0}(7,10) = (a_{03} + \lambda_{03})a_{04}$ ,  $\dot{M}_{\theta_0}(7,12) = (a_{03} + \lambda_{03})a_{05}$ ,  $\dot{M}_{\theta_0}(8,6) = a_{03}\lambda_{02} - \lambda_{01}$ ,  $\dot{M}_{\theta_0}(8,8) = -(\lambda_{02} - \lambda_{01}a_{02})$  and  $\dot{M}_{\theta_0}(8,10) = (a_{01}\lambda_{01} + \lambda_{02})a_{04}$ .

**Proposition 3.2.1** Let  $(X_t)$  be the causal solution of the SVARMA equation (3.1) satisfying **A0–A5** and let  $\hat{\theta}_n$  be a QMLE. Then  $\hat{\theta}_n \rightarrow \theta_0$  a.s. as  $n \rightarrow \infty$ . If in addition **A6–A8** hold true, we have

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathcal{N}(0, \Omega := J^{-1}IJ^{-1}),$$

where  $J = J(\theta_0, \Sigma_{e0})$  and  $I = I(\theta_0, \Sigma_{e0})$ , with

$$J(\theta, \Sigma_e) = \lim_{n \rightarrow \infty} \frac{-2}{n} \frac{\partial^2}{\partial \theta \partial \theta'} \log \tilde{L}_n(\theta, \Sigma_e) \quad \text{a.s. and}$$

$$I(\theta, \Sigma_e) = \lim_{n \rightarrow \infty} \text{Var} \frac{2}{\sqrt{n}} \frac{\partial}{\partial \theta} \log \tilde{L}_n(\theta, \Sigma_e).$$

The proof of this result is postponed to Section 3.5.

**Remark 3.2.2** In the standard strong SVARMA case, i.e. when **A4** is replaced by the assumption that  $(\epsilon_t)$  is an iid sequence, we have  $I = 2J$ , so that  $\Omega = 2J^{-1}$ . We also retrieve the result obtained by [35] (see Lemma 1, p. 2048) for strong SVARMA models.

### 3.3 Diagnostic checking in SVARMA models

For a time  $t \geq 1$ , let  $\hat{\epsilon}_t := \hat{e}_t = \tilde{e}_t(\hat{\theta}_n)$  be the quasi-maximum likelihood residuals when  $p > 0$ ,  $p_s > 0$ ,  $q > 0$  or  $q_s > 0$ . We use (3.4) to notice that we have  $\hat{e}_t = 0$  for  $t \leq 0$  and  $t > n$ . By (3.3), it holds that

$$\begin{aligned} \hat{e}_t &= \hat{X}_t - \sum_{i=1}^p \hat{A}_0^{-1} \hat{A}_i \hat{X}_{t-i} - \sum_{j=1}^{p_s} \hat{A}_0^{-1} \hat{\Lambda}_0^{-1} \hat{\Lambda}_j \hat{A}_0 \hat{X}_{t-sj} + \sum_{i=1}^p \sum_{j=1}^{p_s} \hat{A}_0^{-1} \hat{\Lambda}_0^{-1} \hat{\Lambda}_j \hat{A}_i \hat{X}_{t-i-sj} \\ &+ \sum_{i=1}^q \hat{A}_0^{-1} \hat{\Lambda}_0^{-1} \hat{\Psi}_0 \hat{B}_i \hat{B}_0^{-1} \hat{\Psi}_0^{-1} \hat{\Lambda}_0 \hat{A}_0 \hat{e}_{t-i} + \sum_{j=1}^{q_s} \hat{A}_0^{-1} \hat{\Lambda}_0^{-1} \hat{\Psi}_j \hat{\Psi}_0^{-1} \hat{\Lambda}_0 \hat{A}_0 \hat{e}_{t-sj} \\ &- \sum_{i=1}^q \sum_{j=1}^{q_s} \hat{A}_0^{-1} \hat{\Lambda}_0^{-1} \hat{\Psi}_j \hat{B}_i \hat{B}_0^{-1} \hat{\Psi}_0^{-1} \hat{\Lambda}_0 \hat{A}_0 \hat{e}_{t-i-sj}, \end{aligned}$$

for  $t = 1, \dots, n$ , with  $\hat{X}_t = 0$  for  $t \leq 0$  and  $\hat{X}_t = X_t$  for  $t \geq 1$ . Under the conditions indicated by [31] (see p. 301), we denote by

$$\gamma(sh) = \frac{1}{n} \sum_{t=1+sh}^n e_t e'_{t-sh} \quad \text{and} \quad \hat{\Gamma}_e(sh) = \frac{1}{n} \sum_{t=1+sh}^n \hat{e}_t \hat{e}'_{t-sh}$$

the white noise "empirical" autocovariances and the residual autocovariances. For a fixed integer  $m \geq 1$ , let

$$\begin{aligned} \Upsilon_m(s) &= \left( \{\text{vec} \gamma(1s)\}' , \dots , \{\text{vec} \gamma(ms)\}' \right)' \quad \text{and} \\ \hat{\Gamma}_m(s) &= \left( \{\text{vec} \hat{\Gamma}_e(1s)\}' , \dots , \{\text{vec} \hat{\Gamma}_e(ms)\}' \right)'. \end{aligned}$$

Denoting the diagonal matrices by

$$S_e = \text{Diag}(\sigma_e(1), \dots, \sigma_e(d)) \quad \text{and} \quad \hat{S}_e = \text{Diag}(\hat{\sigma}_e(1), \dots, \hat{\sigma}_e(d)),$$

where  $\sigma_e^2(i)$  is the variance of the  $i$ -th coordinate of  $e_t$  and  $\hat{\sigma}_e^2(i)$  is its sample estimate (i.e.  $\sigma_e^2(i) = \mathbb{E}e_{it}^2$  and  $\hat{\sigma}_e^2(i) = n^{-1} \sum_{t=1}^n \hat{e}_{it}^2$ ). The theoretical and sample autocorrelations at lag  $\ell$  are respectively defined by  $R_e(\ell) = S_e^{-1} \Gamma_e(\ell) S_e^{-1}$  and  $\hat{R}_e(\ell) = \hat{S}_e^{-1} \hat{\Gamma}_e(\ell) \hat{S}_e^{-1}$ , with  $\Gamma_e(\ell) := \mathbb{E}e_t e'_{t-\ell} = 0$  for all  $\ell \neq 0$ . Consider the vector of the first  $m$  sample autocorrelations

$$\hat{\rho}_m(s) = \left( \{\text{vec} \hat{R}_e(1s)\}' , \dots , \{\text{vec} \hat{R}_e(ms)\}' \right)'.$$

The portmanteau statistics (3.2) are generally used to test the null hypothesis

**(H0)** :  $(X_t)_{t \in \mathbb{Z}}$  satisfies a SVARMA( $p, q$ )( $p_s, q_s$ ) representation;  
against the alternative

**(H1)** :  $(X_t)_{t \in \mathbb{Z}}$  does not admit a SVARMA( $p, q$ )( $p_s, q_s$ ) representation or  $(X_t)_{t \in \mathbb{Z}}$  satisfies a SVARMA( $p, q$ )( $p'_s, q'_s$ ) representation with  $p'_s > p_s$  or  $q'_s > q_s$ .

### 3.3.1 The strong SVARMA case

In many situations, the statistics (3.2) are implemented to check the lack of fit of SVARMA models. However, as noted by [31] (see p. 301), the methods of [12] and [21] cannot be extended to the case of SVARMA models when  $s > 1$  because of the multiplicative constraints on the parameters. Using such tests for SVARMA models would be misleading and not enough as these tests consider the autocorrelations corresponding to the nonseasonal lags  $\leq m$  and ignore the possibility of autocorrelations at seasonal lags of multiple period  $s$ . Recently, [30] introduced the following versions of the portmanteau statistics for strong univariate SARMA models

$$Q_m^{\text{BP}}(s) = n \sum_{h=1}^m \hat{\rho}^2(sh) \text{ and } Q_m^{\text{LB}}(s) = n(n+2) \sum_{h=1}^m \frac{\hat{\rho}^2(sh)}{n-sh}. \quad (3.5)$$

Being inspired by (3.2) and the univariate portmanteau statistics (3.5), we introduced the following multivariate versions of the portmanteau statistics

$$\begin{aligned} Q_m^{\text{C}}(s) &= n \sum_{h=1}^m \text{Tr}(\hat{\Gamma}'(sh)\hat{\Gamma}^{-1}(0)\hat{\Gamma}(sh)\hat{\Gamma}^{-1}(0)) \text{ and} \\ Q_m^{\text{H}}(s) &= \sum_{h=1}^m \frac{n^2}{(n-sh)} \text{Tr}(\hat{\Gamma}'(sh)\hat{\Gamma}^{-1}(0)\hat{\Gamma}(sh)\hat{\Gamma}^{-1}(0)). \end{aligned} \quad (3.6)$$

This modification is justifiable under the conditions indicated by [31] (see p. 301). The limiting distribution of the resulting seasonal tests (3.6) are obtained by a straightforward extension of those obtained in [8, 12, 21, 28].

**Proposition 3.3.1** *Assume that the representation (3.1) satisfies a strong SVARMA model. Under the conditions indicated by [31] (see p. 301), namely: if  $p \ll s$  and  $q \ll s$  and the roots of  $\det A_\theta(z) \det B_\theta(z) = 0$  are not close to the unit circle and **H0**, the statistics defined in (3.6) follow a chi-squared distribution with  $d^2(m - p_s - q_s) > 0$  degrees of freedom for large  $n$  and  $ms \leq (n-1)$ .*

The proof of this result is postponed to Section 3.5.

In view of Proposition 3.3.1, the tests procedure consist in rejecting the null hypothesis of a SVARMA( $p, q$ )( $p_s, q_s$ ) model if the statistics (3.6) are larger than a  $\chi_{d^2(m-p_s-q_s)}^2$ . Consequently, these tests also are not applicable for  $m \leq p_s + q_s$ .

### 3.3.2 The weak SVARMA case

The works on the portmanteau statistic are generally performed under the assumption that the errors  $\epsilon_t$  are independent. This independence assumption is often considered too restrictive by practitioners. It precludes conditional heteroscedasticity and/or other forms of nonlinearity (see [17], for a review on weak univariate ARMA models). In this framework, we relax the standard independence assumption on the error term in order to be able to cover SVARMA representations of general nonlinear models.

This paper is devoted to the problem of the validation step of weak SVARMA representations. For such models, we show that the asymptotic distributions of the statistics defined in (3.6) are no longer chi-square distributions but a mixture of chi-squared distributions, weighted by eigenvalues of the asymptotic covariance matrix of the vector of autocorrelations. Contrarily to the standard tests (3.2) and (3.6), the proposed tests can be used safely for  $m$  small.

### Asymptotic distribution of the residual autocorrelations

The asymptotic distribution of  $\sqrt{n}\hat{\rho}_m(s)$  will be obtained from the joint asymptotic distribution of  $\sqrt{n}(\hat{\theta}'_n - \theta'_0, \gamma'_m(s))$ . The mixing assumptions **A7** entail the asymptotic normality of the empirical autocovariances  $\gamma_m(s)$ . Let

$$\Gamma(s\ell, s\ell') = \sum_{h=-\infty}^{\infty} \mathbb{E}(\{e_{t-s\ell} \otimes e_t\} \{e_{t-h-s\ell'} \otimes e_{t-h}\}''), \quad \text{for } (\ell, \ell') \neq (0, 0).$$

We are now able to state the following Theorem, which is an extension of a result given in [5].

**Theorem 3.3.2** *Assume that  $p > 0$  and  $p_s > 0$  or  $q > 0$  and  $q_s > 0$ . Under Assumptions **A0–A8**, as  $n \rightarrow \infty$ ,*

$$\sqrt{n}(\hat{\theta}'_n - \theta'_0, \gamma'_m(s))' = \frac{1}{\sqrt{n}} \sum_{t=1}^n w_t + o_{\mathbb{P}}(1) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Xi),$$

where

$$\Xi = \begin{pmatrix} \Omega & \Sigma_{\hat{\theta}_n, \gamma_m(s)} \\ \Sigma'_{\hat{\theta}_n, \gamma_m(s)} & \Sigma_{\gamma_m(s)} \end{pmatrix} = \sum_{h=-\infty}^{+\infty} \mathbb{E} w_t w'_{t-h}, \quad \text{with } w_t = \begin{pmatrix} w_{1t} \\ w_{2t} \end{pmatrix} \quad (3.7)$$

where  $w_{1t} = -J^{-1}\Upsilon_t = -2J^{-1}(\partial e'_t / \partial \theta) \Sigma_{e0}^{-1} e_t$  and  $w_{2t} = (\epsilon_{t-1s}, \dots, \epsilon_{t-ms})' \epsilon_t$  and where

$$\Sigma_{\gamma_m(s)} = \{\Gamma(s\ell, s\ell')\}_{1 \leq \ell, \ell' \leq m}, \quad \Sigma'_{\gamma_m(s), \hat{\theta}_n} = \lim_{n \rightarrow \infty} \text{Cov}(n^{-1/2} \sum_{t=1}^n J^{-1}\Upsilon_t, \sqrt{n}\gamma_m(s))$$

and  $\Omega = J^{-1}IJ^{-1}$ . The matrices  $I$  and  $J$  are defined in Proposition 3.2.1.

The proof of this result is postponed to Section 3.5.

Define the matrix

$$\Phi_m(s) = \mathbb{E} \left\{ \begin{pmatrix} e_{t-1s} \\ \vdots \\ e_{t-ms} \end{pmatrix} \otimes \frac{\partial e_t}{\partial \theta'} \right\}. \quad (3.8)$$

The following Theorem gives the limiting distribution of the residual autocovariances and autocorrelations.

**Theorem 3.3.3** *Under the none Assumptions as in Theorem 3.3.2, we have*

$$\sqrt{n}\hat{\Gamma}_m(s) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Sigma_{\hat{\Gamma}_m(s)}) \quad \text{and} \quad \sqrt{n}\hat{\rho}_m(s) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Sigma_{\hat{\rho}_m(s)}) \quad \text{where,}$$

$$\Sigma_{\hat{\Gamma}_m(s)} = \Sigma_{\gamma_m(s)} + \Phi_m(s)\Omega\Phi'_m(s) + \Phi_m(s)\Sigma_{\hat{\theta}_n, \gamma_m(s)} + \Sigma'_{\hat{\theta}_n, \gamma_m(s)}\Phi'_m(s) \quad (3.9)$$

$$\Sigma_{\hat{\rho}_m(s)} = \{I_m \otimes (S_e \otimes S_e)^{-1}\} \Sigma_{\hat{\Gamma}_m(s)} \{I_m \otimes (S_e \otimes S_e)^{-1}\} \quad (3.10)$$

and  $\Phi_m(s)$  is given by (3.8).

The proof of this result is postponed to Section 3.5.

**Remark 3.3.1** When  $s = 1$ , we retrieve the result obtained by [5].

In the standard strong SVARMA case, *i.e.* when **A4** is replaced by the assumption that  $(\epsilon_t)$  is an iid sequence and when  $s = 1$ , we have  $\Sigma_{\gamma_m(1)} = I_m \otimes \Sigma_{e0} \otimes \Sigma_{e0}$ ,  $\Omega = 2J^{-1}$  (see Remark 3.2.2) and  $\Sigma'_{\hat{\theta}_n, \gamma_m(s)} = -\Phi_m(1)\Omega$ . Thus  $\Sigma_{\hat{\Gamma}_m(1)} = I_m \otimes \Sigma_{e0} \otimes \Sigma_{e0} - \Phi_m(1)\Omega\Phi'_m(1)$ , the result obtained by [35] is retrieved.

**Limiting distribution of the portmanteau statistics (3.6)**

Several equivalent forms of the statistics (3.2) are given by [21]. By using the elementary identities  $\text{vec}(AB) = (I \otimes A)\text{vec}B$ ,  $(A \otimes B)(C \otimes D) = AC \otimes BD$  and  $\text{Tr}(ABC) = \text{vec}(A')'(C' \otimes I)\text{vec}B$ , analogs basic forms can be found for the statistics (3.6)

$$\begin{aligned} Q_m^C(s) &= n \sum_{h=1}^m \text{Tr}(\hat{\Gamma}'_e(sh)\hat{\Gamma}_e^{-1}(0)\hat{\Gamma}_e(sh)\hat{\Gamma}_e^{-1}(0)) \\ &= n\hat{\rho}'_m(s) (I_m \otimes \{\hat{R}_e^{-1}(0) \otimes \hat{R}_e^{-1}(0)\})\hat{\rho}_m(s). \end{aligned}$$

The same equivalent forms hold for  $Q_m^H(s)$ . From Theorem 3.3.3 we deduce the following result, in the case of weak SVARMA models, which gives the exact asymptotic distribution of the standard portmanteau statistics (3.6). We will see that the distribution may be very different from the  $\chi^2_{d^2(m-p_s-q_s)}$  in the case of strong SVARMA models.

**Theorem 3.3.4** *Under Assumptions in Theorem 3.3.3 and H0, the statistics  $Q_m^C$  and  $Q_m^H$  defined in (3.6) converge in distribution, as  $n \rightarrow \infty$ , to*

$$Z_m(\xi_m) = \sum_{i=1}^{d^2m} \xi_{i,d^2m} Z_i^2$$

where  $\xi_m = (\xi_{1,d^2m}, \dots, \xi_{d^2m,d^2m})'$  is the vector of the eigenvalues of the matrix

$$\Omega_m = (I_m \otimes \Sigma_e^{-1/2} \otimes \Sigma_e^{-1/2}) \Sigma_{\hat{\Gamma}_m} (I_m \otimes \Sigma_e^{-1/2} \otimes \Sigma_e^{-1/2}),$$

and  $Z_1, \dots, Z_{d^2m}$  are independent  $\mathcal{N}(0, 1)$  variables.

In view of Theorem 3.3.3, the asymptotic variance matrices  $\Sigma_{\hat{\Gamma}_m(s)}$  and  $\Sigma_{\hat{\rho}_m(s)}$  depend on the unknown matrices  $\Xi$ ,  $\Phi_m(s)$  and on the nuisance parameters involving in  $\Sigma_e$ . The matrices  $\Phi_m(s)$  and  $\Sigma_e$  can be estimated by its empirical counterpart, respectively

$$\hat{\Phi}_m(s) = \frac{1}{n} \sum_{t=1}^n \left\{ (\hat{e}_{t-1s}, \dots, \hat{e}_{t-ms})' \frac{\partial \hat{e}_t}{\partial \theta'} \right\} \text{ and } \hat{\Sigma}_e = \hat{\Gamma}_e(0) = \frac{1}{n} \sum_{t=1}^n \hat{e}_t \hat{e}_t'.$$

In the econometric literature the nonparametric kernel estimator, also called heteroskedastic autocorrelation consistent (HAC) estimator (see [32], or [2]), is widely used to estimate covariance matrices of the form  $\Xi$ . Interpreting  $(2\pi)^{-1}\Xi$  as the spectral density of the stationary process  $(w_t) = (w'_{1t}, w'_{2t})'$  evaluated at frequency 0 (see [9], p. 459), an alternative method consists in using a parametric AR estimate of the spectral density of  $w_t$ . This approach, which has been studied by [4] (see also [14]), rests on the expression

$$\Xi = \Phi^{-1}(1) \Sigma_u \Phi'^{-1}(1)$$

when  $(w_t)$  satisfies an AR( $\infty$ ) representation of the form

$$\Phi(L)w_t := w_t + \sum_{i=1}^{\infty} \Phi_i w_{t-i} = u_t, \quad (3.11)$$

where  $u_t$  is a weak white noise with variance matrix  $\Sigma_u$ . Since  $w_t$  is not observable, let  $\hat{w}_t$  be the vector obtained by replacing  $\theta_0$  by  $\hat{\theta}_n$  in  $w_t$ . Let  $\hat{\Phi}_r(z) = I_{k_0+m} + \sum_{i=1}^r \hat{\Phi}_{r,i} z^i$ , where  $\hat{\Phi}_{r,1}, \dots, \hat{\Phi}_{r,r}$  denote the coefficients of the least squares regression of  $\hat{w}_t$  on  $\hat{w}_{t-1}, \dots, \hat{w}_{t-r}$ . Let  $\hat{u}_{r,t}$  be the residuals of this regression, and let  $\hat{\Sigma}_{\hat{u}_r}$  be the empirical variance of  $\hat{u}_{r,1}, \dots, \hat{u}_{r,n}$ .

We are now able to state the following Theorem, which is an extension of a result given in [15].



**Theorem 3.3.5** *In addition to the assumptions of Theorem 3.3.4, assume that the process  $(Y_t)$  admits an AR( $\infty$ ) representation (3.11) in which the roots of  $\det \Phi(z) = 0$  are outside the unit disk,  $\|\Phi_i\| = o(i^{-2})$ , and  $\Sigma_u = \text{Var}(u_t)$  is non-singular. Moreover we assume that  $\mathbb{E} \|\epsilon_t\|^{8+4\nu} < \infty$  and  $\sum_{k=0}^{\infty} \{\alpha_{X,\epsilon}(k)\}^{\nu/(2+\nu)} < \infty$  for some  $\nu > 0$ , where  $\{\alpha_{X,\epsilon}(k)\}_{k \geq 0}$  denotes the sequence of the strong mixing coefficients of the process  $(X'_t, \epsilon'_t)'$ . Then, the spectral estimator of  $\Xi$*

$$\hat{\Xi}^{\text{SP}} := \hat{\Phi}_r^{-1}(1) \hat{\Sigma}_{\hat{u}_r} \hat{\Phi}_r'^{-1}(1) \rightarrow \Xi$$

in probability when  $r = r(n) \rightarrow \infty$  and  $r^3/n \rightarrow 0$  as  $n \rightarrow \infty$ .

Let  $\hat{\Omega}_m$  be the matrix obtained by replacing  $\Xi$  by  $\hat{\Xi}$  and  $\Sigma_e$  by  $\hat{\Sigma}_e$  in  $\Omega_m$ . Denote by  $\hat{\xi}_m = (\hat{\xi}_{1,d^2m}, \dots, \hat{\xi}_{d^2m,d^2m})'$  the vector of the eigenvalues of  $\hat{\Omega}_m$ . At the asymptotic level  $\alpha$ , the test consists in rejecting the adequacy of the weak SVARMA model when

$$Q_m^{\text{C}}(s) > S_m(1 - \alpha) \quad (\text{resp.} \quad Q_m^{\text{H}}(s) > S_m(1 - \alpha))$$

where  $S_m(1 - \alpha)$  is such that  $\mathbb{P}\{Z_m(\hat{\xi}_m) > S_m(1 - \alpha)\} = \alpha$ .

## 3.4 Numerical illustrations

In this section, by means of Monte Carlo experiments, we investigate the finite sample properties of the modified version of the portmanteau tests that we introduced in this work. The numerical illustrations of this section are made with the open source statistical software R (see R Development Core Team, 2017) or (see <http://cran.r-project.org/>).

The VAR order  $r$ , in Theorem 3.3.5, can be fixed in a strong SVARMA case, for instance  $r = 1$ . By contrast, in the weak SVARMA cases the order  $r$  is set as  $r = 1, \dots, 5$  and is automatically selected by Akaike Information Criterion (AIC) using the function `VARselect()` of the `vars` R package. The  $p$ -values

$$\mathbb{P}\{Z_m(\hat{\xi}_m) > Q_m^{\text{C}}(s)\} \text{ and } \mathbb{P}\{Z_m(\hat{\xi}_m) > Q_m^{\text{H}}(s)\}, \quad Z_m(\hat{\xi}_m) = \sum_{i=1}^m \hat{\xi}_{i,m} Z_i^2,$$

are evaluated by the Imhof algorithm (see [24]) and by using the function `imhof()` of the R package `CompQuadForm`.

### 3.4.1 Simulated models

First of all, we introduce the models that we simulate.

We indicate the conventions that we adopt in the discussion and in the tables:

- $\text{LB}_W(s)$  and  $\text{BP}_W(s)$  refer to modified LB and BP tests using  $Q_m^{\text{H}}(s)$  and  $Q_m^{\text{C}}(s)$  in Section 3.3.2
- $\text{LB}_M(s)$  and  $\text{BP}_M(s)$  refer to modified LB and BP tests using  $Q_m^{\text{H}}(s)$  and  $Q_m^{\text{C}}(s)$  in Section 3.3.1
- $\text{LB}_{\text{UD}}$  and  $\text{BP}_{\text{UD}}$  refer to LB and BP tests with  $Q_m^{\text{H}}$  and  $Q_m^{\text{C}}$  as in [35].

To generate the strong and the weak SVARMA models, we consider the following bivariate SVARMA(1, 0)(0, 1) $_s$  model

$$\begin{pmatrix} X_{1,t} \\ X_{2,t} \end{pmatrix} = \begin{pmatrix} A_{11,1} & A_{12,1} \\ A_{21,1} & A_{22,1} \end{pmatrix} \begin{pmatrix} X_{1,t-1} \\ X_{2,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix} - \begin{pmatrix} \Psi_{11,1} & \Psi_{12,1} \\ \Psi_{21,1} & \Psi_{22,1} \end{pmatrix} \begin{pmatrix} \epsilon_{1,t-s} \\ \epsilon_{2,t-s} \end{pmatrix} \quad (3.12)$$

with

$$\begin{aligned}\theta_0 &= (A_{11,1}, A_{21,1}, A_{12,1}, A_{22,1}, \Psi_{11,1}, \Psi_{21,1}, \Psi_{12,1}, \Psi_{22,1})' \\ &= (1.2, 0.6, -0.5, 0.3, -0.6, 0.3, 0.3, 0.6)'\end{aligned}$$

and  $\epsilon_t = (\epsilon_{1,t}, \epsilon_{2,t})'$  that follows a strong or weak white noise. Note that

$$\det \left\{ I_2 - z_1 \begin{pmatrix} 1.2 & -0.5 \\ 0.6 & 0.3 \end{pmatrix} \right\} = 0 \text{ and } \det \left\{ I_2 - z_2 \begin{pmatrix} -0.6 & 0.3 \\ 0.3 & 0.6 \end{pmatrix} \right\} = 0$$

for  $z_1 = 1.136 \pm 0.473i$  (hence  $|z_1| = 1.23$ ) and for  $z_2 = \pm(0.45)^{-1/2}$  (and hence  $|z_2| = 1.49$ ).

We first consider the strong SVARMA(1, 0)(0, 1)<sub>s</sub> case by assuming that the innovation process  $\epsilon = ((\epsilon_{1,t}, \epsilon_{2,t})')_{t \geq 1}$  in (3.12) is defined by an iid sequence such that

$$\begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix} \stackrel{\text{law}}{=} \mathcal{N}(0, I_2). \quad (3.13)$$

The GARCH models have numerous extensions to the multivariate framework (see [3] for a review). Jeantheau (see [25]) has proposed a simple extension of the multivariate GARCH( $p, q$ ) with conditional constant correlation. For simplicity we consider the following bivariate ARCH(1) model proposed in [25] and defined by

$$\begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix} = \begin{pmatrix} h_{11,t} & 0 \\ 0 & h_{22,t} \end{pmatrix} \begin{pmatrix} \eta_{1,t} \\ \eta_{2,t} \end{pmatrix} \quad (3.14)$$

where

$$\begin{pmatrix} h_{11,t}^2 \\ h_{22,t}^2 \end{pmatrix} = \begin{pmatrix} 0.3 \\ 0.2 \end{pmatrix} + \begin{pmatrix} 0.45 & 0.00 \\ 0.40 & 0.25 \end{pmatrix} \begin{pmatrix} \epsilon_{1,t-1}^2 \\ \epsilon_{2,t-1}^2 \end{pmatrix}.$$

Finally, for the weak SVARMA(1, 0)(0, 1)<sub>s</sub> case, we assume that in (3.12) the innovation process  $\epsilon$  is an ARCH(1) model defined in (3.14).

### 3.4.2 Empirical size

We first simulate  $N = 1,000$  independent trajectories of size  $n = 2,000$  of models (3.12). The same series is partitioned as two series of sizes  $n = 500$  and  $n = 2,000$ . For each of these  $N$  replications, we use the quasi-maximum likelihood estimation method to estimate the coefficient  $\theta_0$  and we apply portmanteau tests to the residuals for different values of  $m \in \{1, 2, 3, 6, 12\}$  in the proposed tests cases and  $m \in \{1s, 2s, 3s, 6s, 12s\}$  in the standard case, where  $m$  is the number of autocorrelations used in the portmanteau test statistic. We examined the empirical frequencies of rejection of the null hypothesis of adequacy at different nominal levels (1%, 5% and 10%). For each nominal level, we obtained from the  $N$  realizations the empirical frequencies of rejection of the null hypothesis of adequacy. For the nominal level  $\alpha = 5\%$ , the empirical size over the  $N$  independent replications should vary between the significant limits 3.6% and 6.4% with probability 95%. For the nominal level  $\alpha = 1\%$ , the significant limits are 0.3% and 1.7%, and for the nominal level  $\alpha = 10\%$ , they are 8.1% and 11.9%. When the relative rejection frequencies are outside the significant limits, they are displayed in bold type in Tables 3.1, 3.2 and 3.3.

The model (3.12) is therefore rejected when the statistic  $Q_m^C$  or  $Q_m^H$  (resp.  $Q_m^C(s)$  or  $Q_m^H(s)$ ) is larger than  $\chi_{4(m-2)}^2(1 - \alpha)$  (resp. than  $\chi_{4(m-1)}^2(1 - \alpha)$ ). Consequently the empirical size is not available (n.a.) for the statistics  $Q_m^C$  or  $Q_m^H$  (resp.  $Q_m^C(s)$  or  $Q_m^H(s)$ ) because they are not applicable when for  $m \leq 2$  (resp. for  $m = 1$ ). The  $BP_W(s)$  test (resp. the  $LB_W(s)$  test) rejects the adequacy of the weak SVARMA(1, 0)(0, 1)<sub>s</sub> model when the first (resp. the second)  $p$ -value is less than the asymptotic level  $\alpha$ .

Tables 3.1, 3.2 and 3.3 display the relative rejection frequencies of the null hypothesis (**H0**) that the data generating process (DGP for short) follows a strong (Model I) or a weak (Model II) bivariate SVARMA(1,0)(0,1)<sub>s</sub> model, over the N independent replications. For the strong SVARMA model I and for each nominal level, all the relative rejection frequencies are globally inside the significant limits when the seasonal period is  $s = 4$  contrarily to the case  $s = 12$  when  $n = 500$  for  $\alpha = 5\%$  and  $\alpha = 10\%$ .

As expected, for the weak SVARMA model II, Tables 3.1, 3.2 and 3.3 show that the standard  $LB_{UD}$  or  $BP_{UD}$  test poorly performs in assessing the adequacy of these weak SVARMA models. For each nominal level, it can be seen that: 1) the observed relative rejection frequencies of  $LB_{UD}$  and  $BP_{UD}$  are definitely too high, 2)  $BP_W$  or  $BP_M$  works except when  $m \geq 3$ , 3) the errors of the first kind of the proposed tests are globally well controlled when  $n$  is large.

Figure 3.1 plots the autocorrelations for Weak SVARMA(1,0)(0,1)<sub>12</sub> model (3.12)-(3.14). Approximately 95% of the sample autocorrelations of a realization of size  $n = 2,000$  should lie between the bands shown as solid lines (red color). The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 3.3.3. The bands shown as dotted lines (blue color) correspond to the standard (Strong) SVARMA formula.

We draw the conclusion that, even in the strong SVARMA case, the proposed modified version may be clearly preferable to the tests based on  $BP_{UD}$  or  $LB_{UD}$ , when the number  $m$  or  $ms$  of autocorrelations used is small. In all case, the proposed tests well control the error of first kind, contrarily to the  $LB_{UD}$  and  $BP_{UD}$  tests.

### 3.4.3 Application to real data

We now consider an application to monthly U.S. housing starts (HOUST) (total: new privately owned housing units started) and housing sold (HSN1F) (new one family houses sold), in thousands for seasonally adjusted annual rate. The observations covered the period from January 01, 1963 to July 31, 2018 which correspond to  $n = 667$  observations. Figure 3.2 plots the original series of monthly U.S. housing starts (serie 1=HOUST) and housing sold (serie 2=HSN1F), in Thousands. The series exhibit seasonal behavior ( $s = 12$ ). The data were obtain from the website of the U.S. Bureau of the Census (<https://fred.stlouisfed.org/series/HOUST> and <https://fred.stlouisfed.org/series/HSN1F>).

Denoting by

$$(X_t) = \{\log(HOUST_t) - \mathbb{E}(\log(HOUST_t)), \log(HSN1F_t) - \mathbb{E}(\log(HSN1F_t))\}'$$

the mean-corrected series. Hence, the model of the form

$$(I_2 - AL)X_t = (I_2 - \Psi L^{12})(I_2 - BL)\epsilon_t$$

was estimated. The quasi-maximum likelihood estimators were obtained as

$$\hat{\theta}_n = \begin{pmatrix} 0.9955 & [0.0103] & (0.0000) \\ 0.0072 & [0.0098] & (0.4162) \\ -0.0071 & [0.0087] & (0.4649) \\ 0.9858 & [0.0078] & (0.0000) \\ 0.4851 & [0.0507] & (0.0000) \\ -0.1271 & [0.0386] & (0.0000) \\ -0.4061 & [0.0466] & (0.0065) \\ 0.3133 & [0.0410] & (0.0000) \\ 0.1712 & [0.0491] & (0.0005) \\ 0.0639 & [0.0397] & (0.8076) \\ -0.0097 & [0.0338] & (0.0592) \\ 0.1349 & [0.0403] & (0.0009) \end{pmatrix} \text{ and } \hat{\Sigma}_\epsilon = \begin{pmatrix} 0.0049 & 0.0016 \\ 0.0016 & 0.0047 \end{pmatrix},$$

where the estimated asymptotic standard errors obtained from Proposition 3.2.1 (respectively the  $p$ -values), of the estimated parameters (first column), are given into brackets (respectively in parentheses). We then deduce that

$$\hat{A} = \begin{pmatrix} 0.9955 & -0.0071 \\ 0.0072 & 0.9858 \end{pmatrix}, \hat{B} = \begin{pmatrix} 0.4851 & -0.4061 \\ -0.1271 & 0.3133 \end{pmatrix} \text{ and } \hat{\Psi} = \begin{pmatrix} 0.1712 & -0.0097 \\ 0.0639 & 0.1349 \end{pmatrix}.$$

The off-diagonal elements of  $\hat{A}$  and  $\hat{\Psi}$  are not significant and could reasonably be set to zero. We apply portmanteau tests to the residuals of this model. In view of Figure 3.3, the diagnostic checking of residuals does not indicate any inadequacy, approximately 90% of the sample autocorrelations should lie between the bands shown as solid lines (red color). Note that, one can also try several SVARMA models on this data.

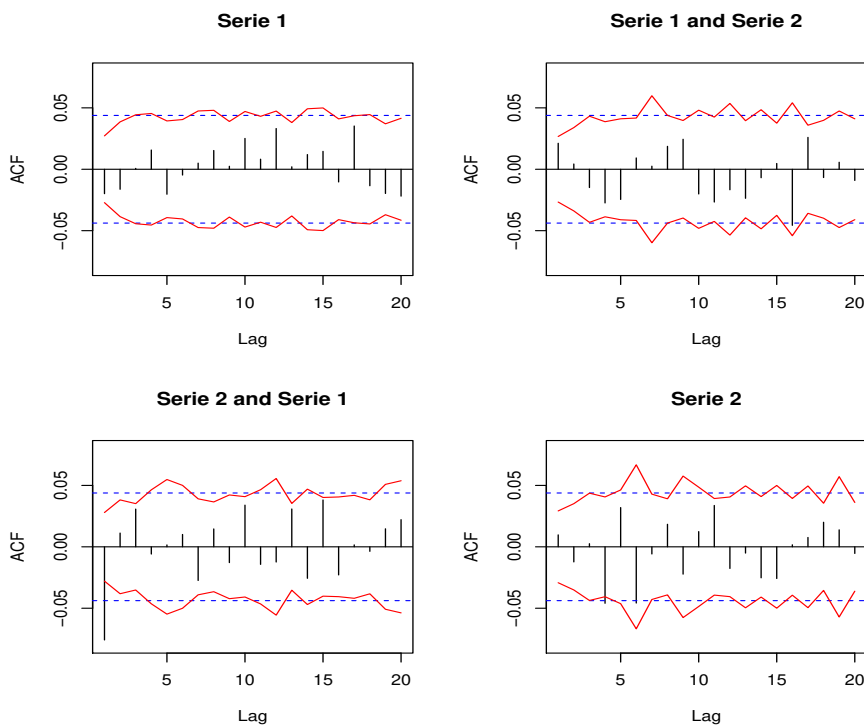


Figure 3.1 – Autocorrelations for Weak SVARMA(1,0)(0,1)<sub>12</sub> model (3.12)-(3.14). Approximately 95% of the sample autocorrelations of a realization of size  $n = 2,000$  should lie between the bands shown as solid lines (red color). The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 3.3.3. The bands shown as dotted lines (blue color) correspond to the standard (Strong) SVARMA formula. The Lag printed is the real lag (in standard case) divided by the season  $s = 12$  (the real standard lags used in the simulation are the lags printed multiplied by the season  $s = 12$ ).

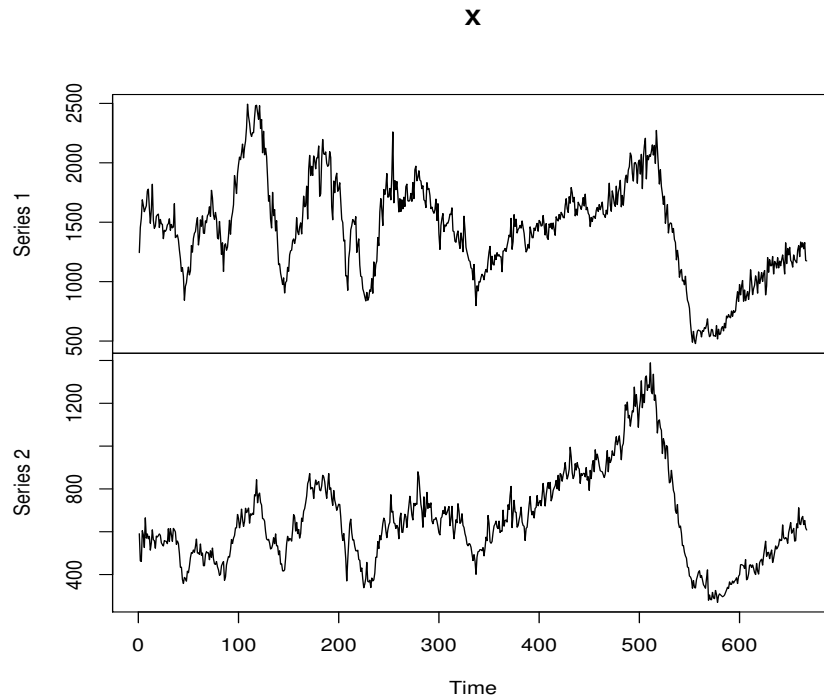


Figure 3.2 – The monthly U.S. housing starts (series 1 corresponds to HOUST) and housing sold (series 2 corresponds to HSN1F), in thousands. The observations covered the period from January 01, 1963 to July 31, 2018 which correspond to  $n = 667$  observations.

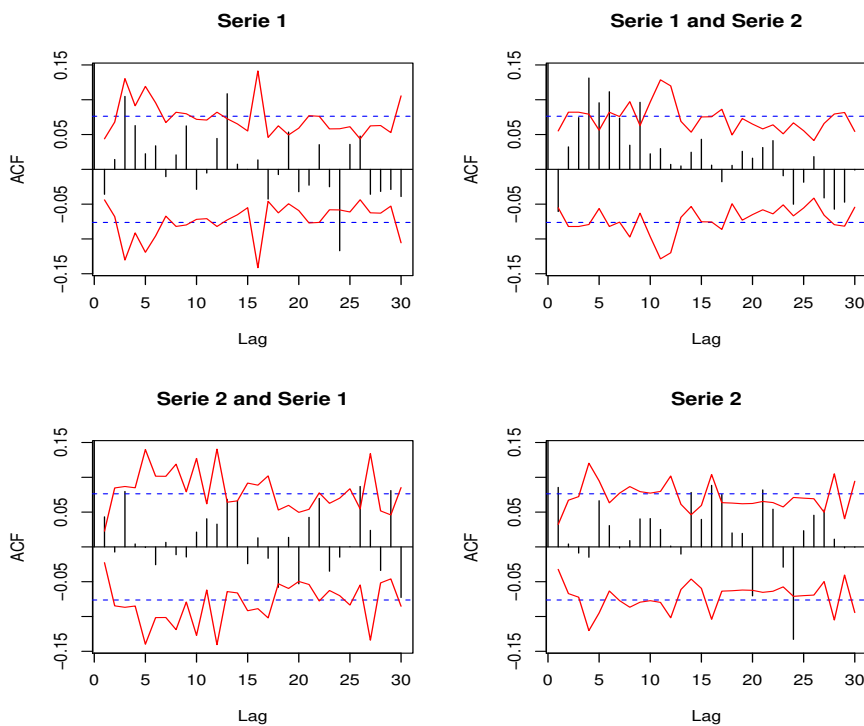


Figure 3.3 – Autocorrelations of the SVARMA(1,0)(1,1)<sub>12</sub> residuals for the mean-logarithm-corrected of monthly U.S. housing starts (serie 1 corresponds to HOUST) and housing sold (serie 2 corresponds to HSN1F). The real standard lags used are the lags printed multiplied by the season  $s = 12$ .

Table 3.1 – Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of SVARMA(1,0)(0,1)<sub>s</sub> model. The nominal asymptotic level of the tests is  $\alpha = 1\%$ .

Model	$s$	Length $n$	Lag $m$	LB <sub>W</sub> ( $s$ )	BP <sub>W</sub> ( $s$ )	LB <sub>M</sub> ( $s$ )	BP <sub>M</sub> ( $s$ )	Lag $ms$	LB <sub>UD</sub>	BP <sub>UD</sub>
I	4	$n = 500$	1	0.9	0.9	n.a.	n.a.	4	<b>6.5</b>	<b>6.5</b>
			2	0.8	0.8	1.0	0.9	8	<b>2.0</b>	<b>1.9</b>
			3	0.6	0.4	0.6	0.5	12	1.3	1.3
			6	0.4	0.3	0.5	0.4	24	1.5	0.9
			12	0.9	0.3	1.4	0.6	48	1.3	0.3
		$n = 2,000$	1	1.2	1.2	n.a.	n.a.	4	<b>5.2</b>	<b>5.2</b>
			2	1.5	1.5	1.6	1.6	8	<b>1.9</b>	<b>1.9</b>
			3	1.0	0.9	1.0	1.0	12	1.6	1.6
			6	1.3	1.3	1.4	1.3	24	1.0	0.8
			12	1.1	1.1	1.2	1.1	48	0.7	0.4
	12	$n = 500$	1	0.6	0.5	n.a.	n.a.	12	<b>3.1</b>	<b>2.7</b>
			2	1.2	0.9	1.5	1.0	24	<b>1.9</b>	1.2
			3	1.4	0.9	1.6	1.2	36	1.5	0.5
			6	<b>2.6</b>	0.9	<b>2.2</b>	0.7	72	<b>1.8</b>	<b>0.2</b>
			12	<b>4.0</b>	0.3	1.6	0.2	144	<b>2.2</b>	<b>0.0</b>
		$n = 2,000$	1	0.4	0.4	n.a.	n.a.	12	<b>3.2</b>	<b>3.1</b>
			2	0.4	0.3	0.6	0.5	24	1.5	1.3
			3	0.7	0.7	0.7	0.7	36	1.2	1.1
			6	1.0	0.7	1.0	0.8	72	1.7	1.1
			12	1.4	1.0	1.2	0.9	144	<b>2.1</b>	0.3
II	4	$n = 500$	1	1.5	1.5	n.a.	n.a.	4	<b>28.0</b>	<b>27.2</b>
			2	0.6	0.5	1.3	1.2	8	<b>12.4</b>	<b>12.0</b>
			3	0.4	0.3	0.6	0.6	12	<b>8.1</b>	<b>7.6</b>
			6	0.4	<b>0.2</b>	1.1	0.4	24	<b>3.5</b>	<b>3.0</b>
			12	<b>0.0</b>	<b>0.0</b>	0.3	<b>0.1</b>	48	<b>3.1</b>	0.9
		$n = 2,000$	1	1.1	1.1	n.a.	n.a.	4	<b>35.6</b>	<b>35.4</b>
			2	1.1	1.1	<b>1.9</b>	<b>1.9</b>	8	<b>18.0</b>	<b>17.9</b>
			3	0.6	0.6	1.3	1.2	12	<b>11.7</b>	<b>11.7</b>
			6	<b>0.1</b>	<b>0.1</b>	0.6	0.6	24	<b>7.6</b>	<b>7.3</b>
			12	0.3	<b>0.2</b>	0.8	0.8	48	<b>5.7</b>	<b>4.4</b>
	12	$n = 500$	1	0.8	0.8	n.a.	n.a.	12	<b>13.2</b>	<b>12.0</b>
			2	0.8	0.5	1.5	1.0	24	<b>5.9</b>	<b>4.3</b>
			3	0.5	<b>0.2</b>	0.5	<b>0.2</b>	36	<b>4.8</b>	<b>2.7</b>
			6	0.8	<b>0.2</b>	0.9	<b>0.2</b>	72	<b>3.2</b>	0.7
			12	0.4	<b>0.0</b>	0.6	<b>0.1</b>	144	<b>4.6</b>	<b>0.0</b>
		$n = 2,000$	1	1.2	1.0	n.a.	n.a.	12	<b>17.3</b>	<b>17.2</b>
			2	0.9	0.8	1.5	1.5	24	<b>9.1</b>	<b>8.9</b>
			3	0.8	0.8	1.0	0.9	36	<b>6.9</b>	<b>6.4</b>
			6	0.9	0.8	1.0	1.0	72	<b>4.7</b>	<b>3.0</b>
			12	1.4	1.1	1.6	0.9	144	<b>3.3</b>	1.4

I: Strong SVARMA(1,0)(0,1)<sub>s</sub> model (3.12)-(3.13).II: Weak SVARMA(1,0)(0,1)<sub>s</sub> model (3.12)-(3.14).

Table 3.2 – Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of SVARMA(1,0)(0,1)<sub>s</sub> model. The nominal asymptotic level of the tests is  $\alpha = 5\%$ .

Model	$s$	Length $n$	Lag $m$	LB <sub>W</sub> ( $s$ )	BP <sub>W</sub> ( $s$ )	LB <sub>M</sub> ( $s$ )	BP <sub>M</sub> ( $s$ )	Lag $m_s$	LB <sub>UD</sub>	BP <sub>UD</sub>
I	4	$n = 500$	1	4.1	4.0	n.a.	n.a.	4	<b>22.1</b>	<b>21.7</b>
			2	4.7	4.3	6.2	5.9	8	<b>7.8</b>	<b>7.3</b>
			3	4.6	3.8	4.2	3.6	12	6.1	5.0
			6	4.3	<b>3.0</b>	4.0	<b>3.3</b>	24	5.6	<b>3.0</b>
			12	5.1	<b>2.6</b>	4.7	<b>2.3</b>	48	5.6	<b>2.0</b>
		$n = 2,000$	1	5.2	5.2	n.a.	n.a.	4	<b>23.6</b>	<b>23.4</b>
			2	5.9	5.8	<b>7.1</b>	<b>7.1</b>	8	<b>8.0</b>	<b>7.7</b>
			3	5.6	5.4	5.6	5.5	12	5.6	5.5
			6	5.3	5.1	5.1	4.8	24	4.7	4.1
			12	6.2	5.7	6.3	5.4	48	5.2	4.2
	12	$n = 500$	1	4.4	4.1	n.a.	n.a.	12	<b>12.8</b>	<b>11.8</b>
			2	5.6	4.4	<b>6.8</b>	5.5	24	<b>7.9</b>	6.2
			3	<b>6.8</b>	5.3	6.4	4.9	36	<b>7.0</b>	3.6
			6	<b>8.0</b>	4.4	6.3	<b>3.4</b>	72	<b>7.2</b>	<b>0.8</b>
			12	<b>15.2</b>	3.7	6.2	<b>0.8</b>	144	<b>7.1</b>	<b>0.0</b>
		$n = 2,000$	1	4.6	4.6	n.a.	n.a.	12	<b>11.8</b>	<b>11.5</b>
			2	4.5	4.5	6.3	5.8	24	<b>6.7</b>	<b>6.6</b>
			3	4.7	4.1	5.2	4.8	36	5.4	4.2
			6	5.1	3.9	4.5	3.6	72	4.9	<b>3.5</b>
			12	5.6	4.2	4.6	<b>3.3</b>	144	6.4	<b>2.0</b>
II	4	$n = 500$	1	5.6	5.3	n.a.	n.a.	4	<b>53.2</b>	<b>52.9</b>
			2	4.4	4.2	<b>6.9</b>	6.4	8	<b>25.3</b>	<b>24.2</b>
			3	4.1	<b>3.4</b>	4.9	4.6	12	<b>19.1</b>	<b>18.1</b>
			6	<b>3.0</b>	<b>2.4</b>	4.0	<b>3.4</b>	24	<b>12.8</b>	<b>9.7</b>
			12	<b>1.7</b>	<b>0.6</b>	3.7	<b>1.5</b>	48	<b>8.6</b>	4.1
		$n = 2,000$	1	4.7	4.7	n.a.	n.a.	4	<b>58.9</b>	<b>58.9</b>
			2	5.0	5.0	<b>7.0</b>	<b>7.0</b>	8	<b>32.8</b>	<b>32.5</b>
			3	<b>3.5</b>	<b>3.5</b>	4.8	4.6	12	<b>24.9</b>	<b>24.4</b>
			6	3.6	<b>3.5</b>	5.0	4.9	24	<b>16.7</b>	<b>15.9</b>
			12	4.1	<b>3.1</b>	5.0	4.1	48	<b>11.9</b>	<b>10.4</b>
	12	$n = 500$	1	4.5	4.3	n.a.	n.a.	12	<b>28.7</b>	<b>27.1</b>
			2	6.1	4.5	<b>6.9</b>	6.4	24	<b>16.3</b>	<b>13.6</b>
			3	4.6	<b>3.5</b>	4.6	<b>3.1</b>	36	<b>13.2</b>	<b>8.6</b>
			6	5.3	<b>2.2</b>	4.7	<b>1.5</b>	72	<b>10.2</b>	<b>2.2</b>
			12	<b>1.4</b>	<b>0.3</b>	3.9	<b>0.3</b>	144	<b>10.2</b>	<b>0.1</b>
		$n = 2,000$	1	4.9	4.8	n.a.	n.a.	12	<b>34.6</b>	<b>34.1</b>
			2	4.9	4.8	<b>7.1</b>	<b>6.6</b>	24	<b>20.4</b>	<b>19.4</b>
			3	4.8	4.3	5.5	5.2	36	<b>16.1</b>	<b>14.6</b>
			6	3.7	<b>3.1</b>	4.0	<b>3.4</b>	72	<b>12.8</b>	<b>9.9</b>
			12	5.3	3.8	5.2	<b>3.2</b>	144	<b>9.0</b>	3.8

I: Strong SVARMA(1,0)(0,1)<sub>s</sub> model (3.12)-(3.13).II: Weak SVARMA(1,0)(0,1)<sub>s</sub> model (3.12)-(3.14).



Table 3.3 – Empirical size (in %) of the modified and standard versions of the LB and BP tests in the case of SVARMA(1,0)(0,1)<sub>s</sub> model. The nominal asymptotic level of the tests is  $\alpha = 10\%$ .

Model	$s$	Length $n$	Lag $m$	LB <sub>W</sub> ( $s$ )	BP <sub>W</sub> ( $s$ )	LB <sub>M</sub> ( $s$ )	BP <sub>M</sub> ( $s$ )	Lag $ms$	LB <sub>UD</sub>	BP <sub>UD</sub>
I	4	$n = 500$	1	8.9	8.7	n.a.	n.a.	4	<b>35.4</b>	<b>35.5</b>
			2	9.9	9.4	<b>13.0</b>	<b>12.6</b>	8	<b>14.5</b>	<b>13.9</b>
			3	9.7	8.3	9.9	9.2	12	10.7	9.4
			6	9.3	8.4	8.8	<b>7.5</b>	24	11.1	8.6
			12	10.6	<b>6.4</b>	9.1	<b>5.5</b>	48	9.6	<b>3.6</b>
	4	$n = 2,000$	1	9.1	9.1	n.a.	n.a.	4	<b>36.4</b>	<b>36.1</b>
			2	10.6	10.5	13.6	13.5	8	<b>14.7</b>	<b>14.3</b>
			3	10.7	10.5	10.8	10.7	12	10.8	10.6
			6	10.9	10.6	9.9	9.5	24	10.1	9.6
			12	<b>12.3</b>	10.5	11.7	10.3	48	9.6	8.3
	12	$n = 500$	1	10.9	10.4	n.a.	n.a.	12	<b>22.2</b>	<b>20.8</b>
			2	<b>12.1</b>	10.3	<b>14.6</b>	<b>12.8</b>	24	<b>15.7</b>	11.3
			3	11.8	9.0	10.7	8.8	36	<b>12.5</b>	<b>7.3</b>
			6	<b>15.9</b>	9.0	11.4	<b>6.1</b>	72	<b>13.5</b>	<b>2.2</b>
			12	<b>26.7</b>	<b>7.3</b>	11.0	<b>2.0</b>	144	11.9	<b>0.1</b>
	12	$n = 2,000$	1	9.3	9.1	n.a.	n.a.	12	<b>20.9</b>	<b>20.7</b>
			2	9.7	9.3	<b>13.4</b>	<b>13.3</b>	24	<b>13.2</b>	<b>12.7</b>
			3	10.4	10.0	10.7	10.0	36	11.3	10.0
			6	11.2	9.9	10.3	8.5	72	10.4	<b>6.5</b>
			12	11.9	<b>7.4</b>	8.9	<b>5.9</b>	144	<b>12.2</b>	<b>3.5</b>
II	4	$n = 500$	1	10.3	10.1	n.a.	n.a.	4	<b>66.5</b>	<b>66.2</b>
			2	9.5	8.9	<b>14.5</b>	<b>14.1</b>	8	<b>36.9</b>	<b>36.2</b>
			3	8.3	<b>7.8</b>	10.2	9.6	12	<b>28.9</b>	<b>27.2</b>
			6	<b>7.5</b>	<b>5.9</b>	8.6	<b>7.1</b>	24	<b>20.4</b>	<b>17.0</b>
			12	<b>3.6</b>	<b>2.1</b>	<b>7.3</b>	<b>4.1</b>	48	<b>15.4</b>	<b>6.8</b>
	4	$n = 2,000$	1	10.2	10.2	n.a.	n.a.	4	<b>71.3</b>	<b>71.2</b>
			2	8.2	8.1	12.7	12.7	8	<b>44.3</b>	<b>44.1</b>
			3	<b>6.6</b>	<b>6.5</b>	10.5	9.9	12	<b>33.1</b>	<b>32.6</b>
			6	8.4	<b>7.5</b>	9.5	9.2	24	<b>24.2</b>	<b>23.2</b>
			12	8.3	<b>7.0</b>	9.3	8.5	48	<b>19.2</b>	<b>16.1</b>
	12	$n = 500$	1	9.1	8.3	n.a.	n.a.	12	<b>41.0</b>	<b>38.4</b>
			2	11.1	10.1	<b>13.5</b>	11.2	24	<b>24.8</b>	<b>21.3</b>
			3	9.8	<b>7.2</b>	9.0	<b>7.3</b>	36	<b>20.2</b>	<b>14.0</b>
			6	11.9	<b>6.0</b>	9.0	<b>4.9</b>	72	<b>16.6</b>	<b>4.3</b>
			12	<b>2.6</b>	<b>0.6</b>	9.1	<b>0.7</b>	144	<b>16.5</b>	<b>0.2</b>
	12	$n = 2,000$	1	9.6	9.2	n.a.	n.a.	12	<b>46.1</b>	<b>45.6</b>
			2	9.2	9.1	<b>12.1</b>	<b>12.1</b>	24	<b>29.4</b>	<b>28.2</b>
			3	10.2	9.6	11.2	10.3	36	<b>23.1</b>	<b>21.9</b>
			6	8.1	<b>7.5</b>	<b>7.8</b>	<b>6.9</b>	72	<b>19.2</b>	<b>14.8</b>
			12	10.7	8.1	9.5	<b>6.6</b>	144	<b>15.9</b>	<b>5.5</b>

I: Strong SVARMA(1,0)(0,1)<sub>s</sub> model (3.12)-(3.13).II: Weak SVARMA(1,0)(0,1)<sub>s</sub> model (3.12)-(3.14).

## 3.5 Proofs

### Proof of Proposition 3.2.1

The proof of Proposition 3.2.1 is similar to that given by [6] for weak VARMA models.  $\square$

**Proof of Proposition 3.3.1.** The proof is a straightforward extension of those obtained in [8, 12, 21, 28].  $\square$

**Proof of Theorem 3.3.2.** Let  $\tilde{\ell}_n(\theta, \Sigma_e) = -2n^{-1} \log \tilde{L}_n(\theta, \Sigma_e)$ . As in [6], it can be shown that  $\ell_n(\theta, \Sigma_e) = \tilde{\ell}_n(\theta, \Sigma_e) + o(1)$  a.s, where

$$\ell_n(\theta, \Sigma_e) := -\frac{2}{n} \log L_n(\theta, \Sigma_e) = \frac{1}{n} \sum_{t=1}^n \{d \log(2\pi) + \log \det \Sigma_e + e_t'(\theta) \Sigma_e^{-1} e_t(\theta)\},$$

and where  $(e_t(\theta))$  is given by

$$\begin{aligned} e_t(\theta) &= X_t - \sum_{i=1}^p A_0^{-1} A_i X_{t-i} - \sum_{j=1}^{p_s} A_0^{-1} \Lambda_0^{-1} \Lambda_j A_0 X_{t-sj} + \sum_{i=1}^p \sum_{j=1}^{p_s} A_0^{-1} \Lambda_0^{-1} \Lambda_j A_i X_{t-i-sj} \\ &+ \sum_{i=1}^q A_0^{-1} \Lambda_0^{-1} \Psi_0 B_i B_0^{-1} \Psi_0^{-1} \Lambda_0 A_0 e_{t-i}(\theta) + \sum_{j=1}^{q_s} A_0^{-1} \Lambda_0^{-1} \Psi_j \Psi_0^{-1} \Lambda_0 A_0 e_{t-sj}(\theta) \\ &- \sum_{i=1}^q \sum_{j=1}^{q_s} A_0^{-1} \Lambda_0^{-1} \Psi_j B_i B_0^{-1} \Psi_0^{-1} \Lambda_0 A_0 e_{t-i-sj}(\theta). \end{aligned} \quad (3.15)$$

It can be also shown uniformly in  $\theta \in \Theta$  that

$$\frac{\partial \ell_n(\theta, \Sigma_e)}{\partial \theta} = \frac{\partial \tilde{\ell}_n(\theta, \Sigma_e)}{\partial \theta} + o(1) \quad a.s.$$

The same equality holds for the second-order derivatives of  $\tilde{\ell}_n(\theta, \Sigma_e)$ . Under **A6**, we have almost surely  $\hat{\theta}_n \rightarrow \theta_0 \in \overset{\circ}{\Theta}$ . Thus  $\partial \tilde{\ell}_n(\hat{\theta}_n, \hat{\Sigma}_e) / \partial \theta = 0$  for sufficiently large  $n$ , and a standard Taylor expansion of the derivative of  $\tilde{\ell}_n$  about  $(\theta_0, \Sigma_{e0})$ , taken at  $(\hat{\theta}_n, \hat{\Sigma}_e)$ , yields

$$\begin{aligned} 0 &= \sqrt{n} \frac{\partial \tilde{\ell}_n(\hat{\theta}_n, \hat{\Sigma}_e)}{\partial \theta} = \sqrt{n} \frac{\partial \tilde{\ell}_n(\theta_0, \Sigma_{e0})}{\partial \theta} + \frac{\partial^2 \tilde{\ell}_n(\theta^*, \Sigma_e^*)}{\partial \theta \partial \theta'} \sqrt{n} (\hat{\theta}_n - \theta_0) \\ &= \sqrt{n} \frac{\partial \ell_n(\theta_0, \Sigma_{e0})}{\partial \theta} + \frac{\partial^2 \ell_n(\theta_0, \Sigma_{e0})}{\partial \theta \partial \theta'} \sqrt{n} (\hat{\theta}_n - \theta_0) + o_{\mathbb{P}}(1), \end{aligned} \quad (3.16)$$

where  $\theta^*$  is between  $\theta_0$  and  $\hat{\theta}_n$ , and  $\Sigma_e^*$  is between  $\Sigma_{e0}$  and  $\hat{\Sigma}_e$ , with  $\hat{\Sigma}_e = n^{-1} \sum_{t=1}^n \tilde{e}_t(\hat{\theta}_n) \tilde{e}_t'(\hat{\theta}_n)$ . Thus, by standard arguments, we have from (3.16):

$$\begin{aligned} \sqrt{n} (\hat{\theta}_n - \theta_0) &= -J^{-1} \sqrt{n} \frac{\partial \ell_n(\theta_0, \Sigma_{e0})}{\partial \theta} + o_{\mathbb{P}}(1) \\ &= J^{-1} \sqrt{n} Y_n + o_{\mathbb{P}}(1) \end{aligned}$$

where

$$Y_n = -\frac{\partial \ell_n(\theta_0, \Sigma_{e0})}{\partial \theta} = -\frac{1}{n} \sum_{t=1}^n \frac{\partial}{\partial \theta} \{d \log(2\pi) + \log \det \Sigma_{e0} + e_t'(\theta_0) \Sigma_{e0}^{-1} e_t(\theta_0)\}. \quad (3.17)$$

Using well-known results on matrix derivatives (see (5) of Appendix A.13 in [29]), we have

$$Y_n = -\frac{2}{n} \sum_{t=1}^n \frac{\partial e_t'(\theta_0)}{\partial \theta} \Sigma_{e0}^{-1} e_t(\theta_0).$$

Now, using the elementary relation  $\text{vec}(ABC) = (C' \otimes A) \text{vec}(B)$  (see (4) of Appendix A.12 in [29]), we have  $\text{vec} \gamma(s\ell) = n^{-1} \sum_{t=s\ell+1}^n e_{t-s\ell} \otimes e_t$ . It is easily shown that for  $\ell, \ell' \geq 1$ ,

$$\begin{aligned} \text{Cov}(\sqrt{n} \text{vec} \gamma(s\ell), \sqrt{n} \text{vec} \gamma(s\ell')) &= \frac{1}{n} \sum_{t=s\ell+1}^n \sum_{t'=s\ell'+1}^n \mathbb{E}(\{e_{t-s\ell} \otimes e_t\} \{e_{t'-s\ell'} \otimes e_{t'}\}') \\ &\rightarrow \Gamma(s\ell, s\ell') \quad \text{as } n \rightarrow \infty. \end{aligned}$$

Then, we have

$$\Sigma_{\gamma_m(s)} = \{\Gamma(s\ell, s\ell')\}_{1 \leq \ell, \ell' \leq m}$$

By stationarity of  $(e_t)$  and  $(Y_t)$  and the dominated convergence Theorem, we have

$$\begin{aligned} \text{Cov}(\sqrt{n} J^{-1} Y_n, \sqrt{n} \text{vec} \gamma(s\ell)) &= -\frac{2}{n} \sum_{t=1}^n \sum_{t'=s\ell+1}^n J^{-1} \text{Cov} \left( \frac{\partial e'_t(\theta_0)}{\partial \theta} \Sigma_{e_0}^{-1} e_t, e_{t-s\ell} \otimes e_t \right) \\ &= -\frac{2}{n} \sum_{h=-n+1}^{n-1} (n-|h|) J^{-1} \text{Cov} \left( \frac{\partial e'_t(\theta_0)}{\partial \theta} \Sigma_{e_0}^{-1} e_t, e_{t-h-s\ell} \otimes e_{t-h} \right) \\ &\rightarrow -\sum_{h=-\infty}^{+\infty} 2J^{-1} \mathbb{E} \left( \frac{\partial e'_t(\theta_0)}{\partial \theta} \Sigma_{e_0}^{-1} e_t \{e_{t-s\ell-h} \otimes e_{t-h}\}' \right). \end{aligned}$$

Then we have

$$\Sigma'_{\gamma_m(s), \hat{\theta}_n} = -2J^{-1} \sum_{h=-\infty}^{+\infty} \mathbb{E} \left( \frac{\partial e'_t(\theta_0)}{\partial \theta} \Sigma_{e_0}^{-1} e_t \left\{ \begin{pmatrix} e_{t-1s-h} \\ \vdots \\ e_{t-ms-h} \end{pmatrix} \otimes e_{t-h} \right\}' \right).$$

Applying the central limit Theorem (CLT) for mixing processes (see [19]) we directly obtain

$$\lim_{n \rightarrow \infty} \text{Var}(\sqrt{n} J^{-1} Y_n) = J^{-1} \mathbb{I} J^{-1}$$

which gives the asymptotic covariance matrix given in (3.7). The existence of these matrices is ensured by the Davydov inequality (see [13]) and **A7**. The proof is then complete.  $\square$

**Proof of Theorem 3.3.3.** Considering  $\hat{\Gamma}(sh)$  and  $\gamma(sh)$  as values of the same function at the points  $\hat{\theta}_n$  and  $\theta_0$ , a Taylor expansion about  $\theta_0$  gives

$$\begin{aligned} \text{vec} \hat{\Gamma}_e(sh) &= \text{vec} \gamma(sh) + \frac{1}{n} \sum_{t=sh+1}^n \left\{ e_{t-sh}(\theta) \otimes \frac{\partial e_t(\theta)}{\partial \theta'} \right. \\ &\quad \left. + \frac{\partial e_{t-sh}(\theta)}{\partial \theta'} \otimes e_t(\theta) \right\}_{\theta=\theta_n^*} (\hat{\theta}_n - \theta_0) + O_{\mathbb{P}}(1/n) \\ &= \text{vec} \gamma(sh) + \mathbb{E} \left( e_{t-sh}(\theta_0) \otimes \frac{\partial e_t(\theta_0)}{\partial \theta'} \right) (\hat{\theta}_n - \theta_0) + O_{\mathbb{P}}(1/n), \end{aligned}$$

where  $\theta_n^*$  is between  $\hat{\theta}_n$  and  $\theta_0$ . The last equality follows from the consistency of  $\hat{\theta}_n$  and the fact that  $(\partial e_{t-sh} / \partial \theta')(\theta_0)$  is not correlated with  $e_t$  when  $h \geq 0$ . Then for  $h = 1, \dots, m$ ,

$$\hat{\Gamma}_m(s) := \left( \{\text{vec} \hat{\Gamma}_e(1s)\}' , \dots , \{\text{vec} \hat{\Gamma}_e(ms)\}' \right)' = \gamma_m(s) + \Phi_m(s) (\hat{\theta}_n - \theta_0) + O_{\mathbb{P}}(1/n), \quad (3.18)$$

where  $\Phi_m(s)$  is defined in (3.8). From Theorem 3.3.2, we have obtained the asymptotic joint distribution of  $\gamma_m(s)$  and  $\hat{\theta}_n - \theta_0$ . Using (3.18) and the TCL of [19], we obtain that the asymptotic

distribution of  $\sqrt{n}\hat{\Gamma}_m(s)$ , is normal, with mean zero and covariance matrix

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Var}(\sqrt{n}\hat{\Gamma}_m(s)) &= \lim_{n \rightarrow \infty} \text{Var}(\sqrt{n}\gamma_m(s)) + \Phi_m(s) \lim_{n \rightarrow \infty} \text{Var}(\sqrt{n}(\hat{\theta}_n - \theta_0))\Phi'_m(s) \\ &\quad + \Phi_m(s) \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}(\hat{\theta}_n - \theta_0), \sqrt{n}\gamma_m(s)) \\ &\quad + \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\gamma_m(s), \sqrt{n}(\hat{\theta}_n - \theta_0))\Phi'_m(s) \\ &= \Sigma_{\gamma_m(s)} + \Phi_m(s)\Omega\Phi'_m(s) + \Phi_m(s)\Sigma_{\hat{\theta}_n, \gamma_m(s)} + \Sigma'_{\hat{\theta}_n, \gamma_m(s)}\Phi'_m(s). \end{aligned}$$

From a Taylor expansion about  $\theta_0$  of  $\text{vec}\hat{\Gamma}_e(0)$  we have,  $\text{vec}\hat{\Gamma}_e(0) = \text{vec}\gamma(0) + O_{\mathbb{P}}(n^{-1/2})$ . Moreover,  $\sqrt{n}(\text{vec}\gamma(0) - \mathbb{E}\text{vec}\gamma(0)) = O_{\mathbb{P}}(1)$  by the CLT for mixing processes (see [19]). Thus  $\sqrt{n}(\hat{S}_e \otimes \hat{S}_e - S_e \otimes S_e) = O_{\mathbb{P}}(1)$  and, using (3.9) and the ergodic Theorem, we obtain

$$n \{ \text{vec}(\hat{S}_e^{-1}\hat{\Gamma}_e(sh)\hat{S}_e^{-1}) - \text{vec}(S_e^{-1}\hat{\Gamma}_e(sh)S_e^{-1}) \} = O_{\mathbb{P}}(1).$$

In the previous equalities, we also use  $\text{vec}(ABC) = (C' \otimes A) \text{vec}(B)$  and  $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$  when A and B are invertible. It follows that

$$\begin{aligned} \hat{\rho}_m(s) &= \left( \{ \text{vec}\hat{R}_e(1s) \}' , \dots , \{ \text{vec}\hat{R}_e(ms) \}' \right)' \\ &= \left( \{ (\hat{S}_e \otimes \hat{S}_e)^{-1} \text{vec}\hat{\Gamma}_e(1s) \}' , \dots , \{ (\hat{S}_e \otimes \hat{S}_e)^{-1} \text{vec}\hat{\Gamma}_e(ms) \}' \right)' \\ &= \{ I_m \otimes (\hat{S}_e \otimes \hat{S}_e)^{-1} \} \hat{\Gamma}_m(s) = \{ I_m \otimes (S_e \otimes S_e)^{-1} \} \hat{\Gamma}_m(s) + O_{\mathbb{P}}(n^{-1}). \end{aligned}$$

We now obtain (3.10) from (3.9). Hence, we have

$$\text{Var}(\sqrt{n}\hat{\rho}_m(s)) = \{ I_m \otimes (S_e \otimes S_e)^{-1} \} \Sigma_{\hat{\Gamma}_m(s)} \{ I_m \otimes (S_e \otimes S_e)^{-1} \}'.$$

This completes the proof.  $\square$

**Proof of Theorem 3.3.5.** The proof is similar to that given by [15] for Theorem 3.  $\square$

## Bibliography

- [1] S. K. Ahn. Distribution for residual autocovariances in multivariate autoregressive models with structured parameterization. *Biometrika*, 75(3):590–593, 1988. [76](#), [77](#)
- [2] Donald W. K. Andrews. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3):817–858, 1991. [84](#)
- [3] Luc Bauwens, Sébastien Laurent, and Jeroen V. K. Rombouts. Multivariate GARCH models: a survey. *J. Appl. Econometrics*, 21(1):79–109, 2006. [86](#)
- [4] Kenneth N. Berk. Consistent autoregressive spectral estimates. *Ann. Statist.*, 2:489–502, 1974. Collection of articles dedicated to Jerzy Neyman on his 80th birthday. [84](#)
- [5] Y. Boubacar Mainassara. Multivariate portmanteau test for structural VARMA models with uncorrelated but non-independent error terms. *J. Statist. Plann. Inference*, 141(8):2961–2975, 2011. [77](#), [83](#)
- [6] Y. Boubacar Mainassara and C. Francq. Estimating structural VARMA models with uncorrelated but non-independent error terms. *J. Multivariate Anal.*, 102(3):496–505, 2011. [78](#), [79](#), [93](#)

- [7] Y. Boubacar Maïnassara and B. Sausserau. Diagnostic checking in multivariate ARMA models with dependent errors using normalized residual autocorrelations. *J. Amer. Statist. Assoc.*, To appear, 2017. [77](#)
- [8] G. E. P. Box and David A. Pierce. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J. Amer. Statist. Assoc.*, 65:1509–1526, 1970. [76](#), [82](#), [93](#)
- [9] Peter J. Brockwell and Richard A. Davis. *Time series: theory and methods*. Springer Series in Statistics. Springer-Verlag, New York, second edition, 1991. [84](#)
- [10] Chun-Zheng Cao, Jin-Guan Lin, and Li-Xing Zhu. Heteroscedasticity and/or autocorrelation diagnostics in nonlinear models with AR(1) and symmetrical errors. *Statist. Papers*, 51(4):813–836, 2010. [77](#)
- [11] Maddalena Cavicchioli. Weak VARMA representations of regime-switching state-space models. *Statist. Papers*, 57(3):705–720, 2016. [76](#)
- [12] Ratnam V. Chitturi. Distribution of residual autocorrelations in multiple autoregressive schemes. *J. Amer. Statist. Assoc.*, 69:928–934, 1974. [76](#), [82](#), [93](#)
- [13] Ju. A. Davydov. The convergence of distributions which are generated by stationary random processes. *Teor. Veroyatnost. i Primenen.*, 13:730–737, 1968. [94](#)
- [14] Wouter J. den Haan and Andrew T. Levin. A practitioner’s guide to robust covariance matrix estimation. In *Robust inference*, volume 15 of *Handbook of Statist.*, pages 299–342. North-Holland, Amsterdam, 1997. [84](#)
- [15] Christian Francq, Roch Roy, and Jean-Michel Zakoïan. Diagnostic checking in ARMA models with uncorrelated errors. *J. Amer. Statist. Assoc.*, 100(470):532–544, 2005. [76](#), [77](#), [84](#), [95](#)
- [16] Christian Francq and Jean-Michel Zakoïan. Stationarity of multivariate markov-switching arma models. *J. Econometrics*, 102(2):339–364, 2001. [76](#)
- [17] Christian Francq and Jean-Michel Zakoïan. Recent results for linear time series models with non independent innovations. In *Statistical modeling and analysis for complex data problems*, volume 1 of *GERAD 25th Anniv. Ser.*, pages 241–265. Springer, New York, 2005. [76](#), [82](#)
- [18] E. J. Hannan. The identification and parametrization of ARMAX and state space forms. *Econometrica*, 44(4):713–723, 1976. [79](#)
- [19] Norbert Herrndorf. A functional central limit theorem for weakly dependent sequences of random variables. *Ann. Probab.*, 12(1):141–153, 1984. [94](#), [95](#)
- [20] Keith Hipel and A. Ian McLeod. *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier, Amsterdam, 1994. [75](#)
- [21] J. R. M. Hosking. The multivariate portmanteau statistic. *J. Amer. Statist. Assoc.*, 75(371):602–608, 1980. [76](#), [82](#), [84](#), [93](#)
- [22] J. R. M. Hosking. Equivalent forms of the multivariate portmanteau statistic. *J. Roy. Statist. Soc. Ser. B*, 43(2):261–262, 1981. [76](#)

- [23] J. R. M. Hosking. Corrigendum: “Equivalent forms of the multivariate portmanteau statistic” [J. Roy. Statist. Soc. Ser. B **43** (1981), no. 2, 261–262; MR0626774 (82h:62153)]. *J. Roy. Statist. Soc. Ser. B*, 51(2):303, 1989. [76](#)
- [24] J. P. Imhof. Computing the distribution of quadratic forms in normal variables. *Biometrika*, 48:419–426, 1961. [85](#)
- [25] Thierry Jeantheau. Strong consistency of estimators for multivariate ARCH models. *Econometric Theory*, 14(1):70–86, 1998. [86](#)
- [26] Naoya Katayama. Chi-squared portmanteau tests for structural VARMA models with uncorrelated errors. *J. Time Series Anal.*, 33(6):863–872, 2012. [77](#)
- [27] W. K. Li and A. I. McLeod. Distribution of the residual autocorrelations in multivariate ARMA time series models. *J. Roy. Statist. Soc. Ser. B*, 43(2):231–239, 1981. [76](#), [77](#)
- [28] G. M. Ljung and G. E. P. Box. On a measure of lack of fit in time series models. *Biometrika*, 65(2):pp. 297–303, 1978. [76](#), [82](#), [93](#)
- [29] Helmut Lütkepohl. *New introduction to multiple time series analysis*. Springer-Verlag, Berlin, 2005. [75](#), [93](#), [94](#)
- [30] Esam Mahdi. Portmanteau test statistics for seasonal serial correlation in time series models. *SpringerPlus*, 5(1):1485, 2016. [82](#)
- [31] A. I. McLeod. On the distribution of residual autocorrelations in Box-Jenkins models. *J. Roy. Statist. Soc. Ser. B*, 40(3):296–302, 1978. [76](#), [77](#), [81](#), [82](#)
- [32] Whitney K. Newey and Kenneth D. West. A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708, 1987. [84](#)
- [33] Gregory C. Reinsel. *Elements of multivariate time series analysis*. Springer Series in Statistics. Springer-Verlag, New York, second edition, 1997. [76](#)
- [34] Carlos Eduardo M. Relvas and Gilberto A. Paula. Partially linear models with first-order autoregressive symmetric errors. *Statist. Papers*, 57(3):795–825, 2016. [77](#)
- [35] Eugen Ursu and Pierre Duchesne. On multiplicative seasonal modelling for vector time series. *Statist. Probab. Lett.*, 79(19):2045–2052, 2009. [77](#), [81](#), [83](#), [85](#)

# Portmanteau tests for Periodic ARMA models with uncorrelated but non-independent error terms

## Abstract

In this paper we derive the asymptotic distributions of residual and normalized residual empirical autocovariances and autocorrelations of periodic autoregressive moving-average (PARMA) models under the assumption that the errors are uncorrelated but not necessarily independent. We then deduce the modified portmanteau statistics. We establish the asymptotic behaviour of the proposed statistics. It is shown that the asymptotic distribution of the portmanteau tests is that of a weighted sum of independent chi-squared random variables, which can be quite different from the usual chi-squared approximation used under independent and identically distributed (iid) assumptions on the noise. We also propose another test based on a self-normalization approach to check the adequacy of PARMA models. A set of Monte Carlo experiments and an application to financial data are presented.

**Keywords:** Goodness-of-fit test, portmanteau test statistics, residual autocorrelations, seasonality, self-normalization, weak PARMA models, weighted least squares

## 4.1 Introduction

Time series with periodic means and autocovariances frequently arise from various fields of study, such as meteorology, hydrology and economics, amongst others (see for instance [28], [25, Chapter 14] and [26]). Modeling methods for periodic series focus on the important class of periodic autoregressive moving-average (PARMA) models as the PARMA second moment structure is indeed periodic. PARMA models are an extension of autoregressive moving average (ARMA) models in the sense that they allow the model parameters to vary with respect to time.

A series  $X = (X_t)_{t \in \mathbb{Z}}$  with mean  $\mu_t = \mu_{t+T}$  and finite second moments is said to be a PARMA series with period  $T$  ( $T$  is assumed known throughout) if it is a solution to the periodic linear difference equation

$$(X_t - \mu_t) - \sum_{i=1}^p \phi_i(t)(X_{t-i} - \mu_{t-i}) = \epsilon_t - \sum_{j=1}^q \theta_j(t)\epsilon_{t-j}, \quad (4.1)$$

where the coefficients satisfy  $\phi_i(t+T) = \phi_i(t)$  for  $1 \leq i \leq p$  and  $\theta_j(t+T) = \theta_j(t)$  for  $1 \leq j \leq q$ . One can allow  $p$  and  $q$  to vary periodically if needed (see [20, 34]).



A notation that emphasizes seasonality uses  $X_{nT+\nu}$  as the data point during the  $\nu$ -th season of the  $n$ -th cycle of data. Here  $\nu$  is a seasonal suffix that satisfies  $1 \leq \nu \leq T$ ; we allow a 0-th cycle of data so that  $X_1$  denotes the first observation (season 1 of cycle 0). The PARMA model in (4.1) with  $p$  and  $q$  vary periodically is equivalently written as

$$(X_{nT+\nu} - \mu_\nu) - \sum_{k=1}^{p_\nu} \phi_k(\nu)(X_{nT+\nu-k} - \mu_{\nu-k}) = \epsilon_{nT+\nu} - \sum_{l=1}^{q_\nu} \theta_l(\nu)\epsilon_{nT+\nu-l} \quad (4.2)$$

and is denoted by  $\text{PARMA}_T(p_1, \dots, p_\nu, \dots, p_T; q_1, \dots, q_\nu, \dots, q_T)$ . The linear periodic innovation process  $(\epsilon_t) = (\epsilon_{nT+\nu})_{n \in \mathbb{Z}}$  is assumed to be a stationary sequence satisfies

$$\text{(A0):} \quad \mathbb{E}[\epsilon_t] = 0, \text{Var}(\epsilon_t) = \sigma_\nu^2 > 0 \text{ and } \mathbb{E}(\epsilon_t \epsilon_{t'}) = 0 \quad \forall t \neq t'.$$

Under the above assumptions the process  $(\epsilon_{nT+\nu})_{n \in \mathbb{Z}}$  is called a weak periodic white noise. An example of weak periodic white noise is the periodic generalized autoregressive conditional heteroscedastic (PGARCH) model (see [7]). Many others examples can also be find in [20]. It is customary to say that  $(X_{nT+\nu})_{n \in \mathbb{Z}}$  is a strong  $\text{PARMA}_T(p_1, \dots, p_\nu, \dots, p_T; q_1, \dots, q_\nu, \dots, q_T)$  representation and we will do this henceforth if in (4.1) or (4.2)  $(\epsilon_{nT+\nu})_{n \in \mathbb{Z}}$  is a strong periodic white noise, namely an independent and identically distributed (iid for short) sequence of random variables with mean 0 and finite variance. A strong white noise is obviously a weak white noise because independence entails uncorrelatedness. Of course the converse is not true. In contrast with this previous definition, the representation (4.1) or (4.2) is said to be a weak  $\text{PARMA}_T(p_1, \dots, p_\nu, \dots, p_T; q_1, \dots, q_\nu, \dots, q_T)$  if no additional assumption is made on  $(\epsilon_{nT+\nu})_{n \in \mathbb{Z}}$ , that is if  $(\epsilon_{nT+\nu})_{n \in \mathbb{Z}}$  is only a weak periodic white noise (not necessarily iid).

When the order of both the autoregressive and moving average components are not allowed to vary with season, *i.e.*, when  $p_1 = \dots = p_\nu = \dots = p_T = p$  and  $q_1 = \dots = q_\nu = \dots = q_T = q$  we simply write  $\text{PARMA}_T(p; q)$  instead of  $\text{PARMA}_T(p, \dots, p; q, \dots, q)$ . In the sequel we suppose that the process  $(X_{nT+\nu})_{n \in \mathbb{Z}}$  is centered, that is  $\mu_1 = \dots = \mu_T = 0$ . Thus Equation (4.2) becomes

$$X_{nT+\nu} - \sum_{k=1}^{p_\nu} \phi_k(\nu)X_{nT+\nu-k} = \epsilon_{nT+\nu} - \sum_{l=1}^{q_\nu} \theta_l(\nu)\epsilon_{nT+\nu-l} \quad (4.3)$$

After identification and estimation of the PARMA processes, the next important step in the modeling consists in checking if the estimated model fits the data satisfactorily. Thus, under the null hypothesis that the model has been correctly identified the residuals,  $\hat{\epsilon}_{nT+\nu}$ , are approximately a periodic white noise. This adequacy checking step validates or invalidates the choice of the orders  $(p_1, \dots, p_\nu, \dots, p_T; q_1, \dots, q_\nu, \dots, q_T)$ .

Thus it is important to check the validity of a  $\text{PARMA}_T(p_1, \dots, p_\nu, \dots, p_T; q_1, \dots, q_\nu, \dots, q_T)$  model, for given orders  $(p_1, \dots, p_\nu, \dots, p_T; q_1, \dots, q_\nu, \dots, q_T)$ . Based on the residual empirical autocorrelations, [14] derived a goodness-of-fit test, the portmanteau test, for univariate strong ARMA models (*i.e.* under the assumption that the error term is iid). [32] proposed a modified portmanteau test which is nowadays one of the most popular diagnostic checking tool in ARMA modeling of time series. Since the article by [32], portmanteau tests have been important tools in time series analysis, in particular for testing the adequacy of an estimated ARMA model. See also [31] for a reference book on the portmanteau tests.

This paper is devoted to the problem of the validation step of the PARMA representations (4.3) processes. The works on the portmanteau statistic of PARMA models are generally performed under the assumption that the errors  $\epsilon_{nT+\nu}$  are independent (see [37, 38] or [19, 47]). This independence assumption is often considered too restrictive by practitioners. It precludes PGARCH and/or other forms of periodic nonlinearity (see [20]). All these above results have been obtained under the iid assumption on the periodic noise and they may be invalid when the periodic series is uncorrelated but dependent. However the standard portmanteau test needs to be adapted to take into account the possible lack of independence of the errors terms. In

this framework we relax the standard independence assumption on the periodic error term in order to be able to cover PARMA representations of general nonlinear models and to extend the range of application of the PARMA models. For such models we show that the asymptotic distributions of the proposed statistics are no longer chi-squared distributions but a mixture of chi-squared distributions, weighted by eigenvalues of the asymptotic covariance matrix of the vector of autocorrelations as in [21]. We also proposed another modified statistics based on a self-normalization approach which are asymptotically distribution-free under the null hypothesis (see [13] for a reference in a weak ARMA case). Contrary to the standard tests the proposed tests can be used safely for  $m$  small, where  $m$  is the number of autocorrelations used in the portmanteau test statistic. Other contribution is to improve the results concerning the statistical analysis of weak PARMA models by considering the adequacy problem.

The article is organized as follows. In Section 4.2 we recall the results on the least squares estimator (LSE) asymptotic distribution of a weak PARMA model obtained by [20] under ergodic and mixing assumptions. Section 4.3 presents our main aim. Then we study the asymptotic behaviour of the residuals autocovariances and autocorrelations of a weak PARMA model in Section 4.3.1. In Section 4.3.2 we derive the asymptotic distribution of residuals autocovariances and autocorrelations using self-normalization approach and we establish the asymptotic behaviour of the proposed statistics. Section 4.4 proposes numerical illustrations. The empirical power are also investigated. Section 4.5 illustrates the portmanteau test for PARMA( $p, q$ ) models applied to stock market indices. The technical proofs are relegated to the appendix.

## 4.2 Preliminaries

We recall in this section some technical results which we are need on weak PARMA and are contained in [20].

### 4.2.1 Causality and invertibility of PARMA models

There is no loss of generality in taking that  $p_1 = \dots = p_\nu = \dots = p_T = p$  and  $q_1 = \dots = q_\nu = \dots = q_T = q$  in (4.3) by adding coefficients equal to zero (see [34] for more details). Thus the PARMA $_T(p; q)$  process  $(X_{nT+\nu})_{n \in \mathbb{Z}}$  satisfies the following difference equations

$$X_{nT+\nu} - \sum_{k=1}^p \phi_k(\nu) X_{nT+\nu-k} = \epsilon_{nT+\nu} - \sum_{l=1}^q \theta_l(\nu) \epsilon_{nT+\nu-l}. \quad (4.4)$$

We denote by  $\lfloor a \rfloor$  the integer part of the real  $a$ . Let  $p^* = \lfloor (p-1)/T \rfloor + 1$  and  $q^* = \lfloor (q-1)/T \rfloor + 1$ . As in [48] the difference equations (4.4) can be written in the following T-dimensional vector ARMA (VARMA) form:

$$\Phi_0 \mathbf{X}_n - \sum_{k=1}^{p^*} \Phi_k \mathbf{X}_{n-k} = \Theta_0 \boldsymbol{\epsilon}_n - \sum_{l=1}^{q^*} \Theta_l \boldsymbol{\epsilon}_{n-l} \quad (4.5)$$

where  $\mathbf{X}_n = (X_{nT+1}, \dots, X_{nT+T})'$  and  $\boldsymbol{\epsilon}_n = (\epsilon_{nT+1}, \dots, \epsilon_{nT+T})'$ . The  $(T \times T)$ -matrices coefficients  $\Phi_k$ ,  $k = 0, \dots, p^*$  and  $\Theta_l$ ,  $l = 0, \dots, q^*$ , are defined by

$$(\Phi_0)_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i < j \\ -\phi_{i-j}(i) & \text{if } i > j \end{cases}, \quad (\Theta_0)_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i < j \\ -\theta_{i-j}(i) & \text{if } i > j \end{cases},$$

and  $(\Phi_k)_{i,j} = \phi_{kT+i-j}(i)$ ,  $k = 1, \dots, p^*$  and  $(\Theta_l)_{i,j} = \theta_{lT+i-j}(i)$ ,  $l = 1, \dots, q^*$ . Here it is implicit that  $\phi_h(\nu) = 0$ , for  $h \notin \{1, \dots, p\}$  and  $\theta_h(\nu) = 0$ , for  $h \notin \{1, \dots, q\}$ . The covariance matrix of the T-dimensional white noise  $\boldsymbol{\epsilon}_n$  is  $\Sigma_{\boldsymbol{\epsilon}} = \text{Diag}(\sigma_1^2, \dots, \sigma_T^2) > 0$ . We denote by L the back-shift operator

such that  $L^k \boldsymbol{\epsilon}_n = \boldsymbol{\epsilon}_{n-k}$ . It is important to note that the lag operator  $L$  operates on the cycle index  $n$ . When it acts on the time index  $t = nT + v$  of the periodic process  $(\epsilon_t)$ , it gives  $L^k \epsilon_{nT+v} = \epsilon_{(n-k)T+v}$ . Equation (4.5) can be written more compactly as

$$\Phi(L)\mathbf{X}_n = \Theta(L)\boldsymbol{\epsilon}_n \quad (4.6)$$

where  $\Phi(z) = \Phi_0 - \sum_{i=1}^{p^*} \Phi_i z^i$  and  $\Theta(z) = \Theta_0 - \sum_{i=1}^{q^*} \Theta_i z^i$ . From (4.5) we can in principle deduce the properties of weak PARMA parameters estimation, identification and validation from existing results on parameters estimation, identification and validation of the weak VARMA models (see for instance [8, 9, 11, 12, 13, 29]). Therefore we have preferred to work in the univariate PARMA setting for various reasons. In particular the results obtained directly in terms of the univariate PARMA representation are more directly usable because fewer parameters are involved and their estimation is easier (see [20] for more details).

We denote by  $\det(A)$  the determinant of a matrix  $A$ . We assume that the  $\text{PARMA}_T(p; q)$  process  $(X_{nT+v})_{n \in \mathbb{Z}}$  corresponds to stable and invertible representations, namely

**(A1):** we have  $\det \Phi(z) \det \Theta(z) \neq 0$  for all  $|z| \leq 1$ . Furthermore as in [11] we assume that the weak VARMA model (4.6) is identifiable (see also [15, 35, 41]).

### 4.2.2 Estimating weak PARMA models

In this section we recall the results on the weighted least squares (WLS) asymptotic distribution obtained by [20] when  $(\epsilon_{nT+v})$  satisfies  $\alpha$ -mixing assumptions. For notation, let  $\phi(v) = (\phi_1(v), \dots, \phi_p(v))'$  and  $\theta(v) = (\theta_1(v), \dots, \theta_q(v))'$  respectively denote the vectors of autoregressive and moving average parameters for a specified season  $v$ . The  $T(p+q)$ -dimensional collection of all PARMA parameters is denoted by

$$\alpha = (\phi(1)', \dots, \phi(T)', \theta(1)', \dots, \theta(T))' \in \mathbb{R}^{(p+q)T}.$$

The periodic white noise variances  $\sigma^2 = (\sigma_1^2, \dots, \sigma_T^2)'$  will be treated as a nuisance parameters.

Let  $X_1, \dots, X_{NT}$  be a data sample from the causal and invertible PARMA model (4.4) with the true parameter value  $\alpha = \alpha_0$  and  $\sigma^2 = \sigma_0^2$ . The sample contains  $N$  full periods of data which are indexed from 0 to  $N-1$ . Indeed when  $0 \leq n \leq N-1$  and  $1 \leq v \leq T$ ,  $nT+v$  goes from 1 to  $NT$ . The unknown parameter of interest  $\alpha_0$  is supposed to belong to the parameter space

$$\Delta = \left\{ \alpha = (\phi(1)', \dots, \phi(T)', \theta(1)', \dots, \theta(T))' \text{ such that (A1) is verified} \right\}.$$

Without loss of generality, we may assume that  $\Delta$  is compact. For  $\alpha \in \Delta$ , let  $\epsilon_{nT+v}(\alpha)$  be the periodically second-order stationary solution of

$$\epsilon_{nT+v}(\alpha) = X_{nT+v} - \sum_{k=1}^p \phi_k(v) X_{nT+v-k} + \sum_{l=1}^q \theta_l(v) \epsilon_{nT+v-l}(\alpha). \quad (4.7)$$

The variable  $\epsilon_{nT+v}(\alpha)$  can be approximated, for  $0 < nT+v \leq NT$ , by  $e_{nT+v}(\alpha)$  defined recursively by

$$e_{nT+v}(\alpha) = X_{nT+v} - \sum_{k=1}^p \phi_k(v) X_{nT+v-k} + \sum_{l=1}^q \theta_l(v) e_{nT+v-l}(\alpha), \quad (4.8)$$

where the unknown initial values are set to zero:  $e_0(\alpha) = \dots = e_{1-q}(\alpha) = X_0 = \dots = X_{1-p} = 0$ . As showed by [20, Lemma 7] these initial values are asymptotically negligible uniformly in  $\alpha$ .

The random variable  $\hat{\alpha}_{\text{OLS}}$  is called the ordinary least squares (OLS) estimator of  $\alpha$  if it satisfies, almost surely,

$$S_N(\hat{\alpha}_{\text{OLS}}) = \min_{\alpha \in \Delta} S_N(\alpha) \quad \text{where} \quad S_N(\alpha) = \frac{1}{2N} \sum_{n=0}^{N-1} \sum_{v=1}^T e_{nT+v}^2(\alpha).$$

Because of the presence of heteroscedastic innovations, the OLS estimator might be inefficient. Given some vectors of weights  $\omega^2 = (\omega_1^2, \dots, \omega_T^2)'$ , [20] and [5] showed that the OLS estimator is asymptotically outperformed by the weighted least squares (WLS) estimator  $\hat{\alpha}_{\text{WLS}} = \hat{\alpha}_{\text{WLS}}^{\omega^2}$  defined by

$$Q_N^{\omega^2}(\hat{\alpha}_{\text{WLS}}) = \min_{\alpha \in \Delta} Q_N^{\omega^2}(\alpha) \quad \text{where} \quad Q_N^{\omega^2}(\alpha) = \frac{1}{2N} \sum_{n=0}^{N-1} \sum_{v=1}^T \omega_v^{-2} e_{nT+v}^2(\alpha).$$

[20] and [5] showed that an optimal WLS estimator is the generalized least squares (GLS) estimator defined by

$$\hat{\alpha}_{\text{GLS}} = \hat{\alpha}_{\text{WLS}}^{\sigma^2}, \quad \text{where} \quad \sigma^2 = (\sigma_1^2, \dots, \sigma_T^2)'. \quad (4.9)$$

As mentioned by [20] the GLS estimator assumes that  $\sigma^2$  is known. In practice this parameter has also to be estimated. Given any consistent estimator  $\hat{\sigma}^2 = (\hat{\sigma}_1^2, \dots, \hat{\sigma}_T^2)'$  of  $\sigma^2$ , a quasi-generalized least squares (QLS) estimator of  $\alpha_0$  is defined by

$$\hat{\alpha}_{\text{QLS}} = \hat{\alpha}_{\text{WLS}}^{\hat{\sigma}^2}. \quad (4.10)$$

One possible consistent estimator of  $\sigma_v^2$  can be obtained by

$$\hat{\sigma}_v^2 = \frac{1}{N} \sum_{n=0}^{N-1} e_{nT+v}^2(\hat{\alpha}_{\text{OLS}}).$$

To establish the consistency of the least squares estimators (LSE), an additional assumption is needed.

**(A2):** The T-dimensional white noise  $\{\epsilon_n = (\epsilon_{nT+1}, \dots, \epsilon_{nT+T})', n \in \mathbb{Z}\}$  is ergodic and strictly stationary.

For the asymptotic normality of the LSE, additional assumptions are also required. It is necessary to assume that  $\alpha_0$  is not on the boundary of the parameter space  $\Delta$ .

**(A3):** We have  $\alpha_0 \in \overset{\circ}{\Delta}$ , where  $\overset{\circ}{\Delta}$  denotes the interior of  $\Delta$ .

To control the serial dependence of the stationary process  $(\epsilon_n)$ , we introduce the strong mixing coefficients  $\alpha_\epsilon(h)$  defined by

$$\alpha_\epsilon(h) = \sup_{A \in \mathcal{F}_{-\infty}^n, B \in \mathcal{F}_{n+h}^{+\infty}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|,$$

where  $\mathcal{F}_{-\infty}^n = \sigma(\epsilon_u, u \leq n)$  and  $\mathcal{F}_{n+h}^{+\infty} = \sigma(\epsilon_u, u \geq n+h)$ . We use  $\|\cdot\|$  to denote the Euclidian norm of a vector. We will make an integrability assumption on the moment of the noise and a summability condition on the strong mixing coefficients  $(\alpha_\epsilon(k))_{k \geq 0}$ .

**(A4):** We have  $\mathbb{E}\|\epsilon_n\|^{4+2\kappa} < \infty$  and  $\sum_{k=0}^{\infty} \{\alpha_\epsilon(k)\}^{\frac{\kappa}{2+\kappa}} < \infty$  for some  $\kappa > 0$ .

In all the sequel we denote by  $\xrightarrow{d}$  the convergence in distribution. The symbol  $o_{\mathbb{P}}(1)$  is used for a sequence of random variables that converges to zero in probability. Under the above assumptions, [20] showed that  $\hat{\alpha}_{\text{OLS}} \rightarrow \alpha_0$ ,  $\hat{\alpha}_{\text{WLS}} \rightarrow \alpha_0$ ,  $\hat{\alpha}_{\text{GLS}} \rightarrow \alpha_0$  and  $\hat{\alpha}_{\text{QLS}} \rightarrow \alpha_0$  *a.s.* as  $N \rightarrow \infty$  and

$$\sqrt{N}(\hat{\alpha}_{\text{LS}} - \alpha_0) \xrightarrow[N \rightarrow \infty]{d} \mathcal{N}(0, \Omega^{\text{LS}}), \quad (4.11)$$

where the exponent LS stands for OLS, WLS, GLS and QLS with

$$\Omega^{\text{OLS}} = \Omega(\alpha_0, (1, \dots, 1)'), \quad \Omega^{\text{WLS}} = \Omega(\alpha_0, \omega^2), \quad \Omega^{\text{GLS}} = \Omega(\alpha_0, \sigma^2), \quad \Omega^{\text{QLS}} = \Omega(\alpha_0, \hat{\sigma}^2)$$

and

$$\Omega(\alpha_0, \omega^2) = (J(\alpha_0, \omega^2))^{-1} I(\alpha_0, \omega^2) (J(\alpha_0, \omega^2))^{-1}.$$

The matrices  $I(\alpha_0, \omega^2)$  and  $J(\alpha_0, \omega^2)$  being defined as

$$J(\alpha_0, \omega^2) = \sum_{v=1}^T \omega_v^{-2} \mathbb{E} \left[ \left( \frac{\partial \epsilon_{nT+v}(\alpha)}{\partial \alpha} \right) \left( \frac{\partial \epsilon_{nT+v}(\alpha)}{\partial \alpha} \right)' \right]_{\alpha=\alpha_0}, \quad (4.12)$$

$$I(\alpha_0, \omega^2) = \sum_{v=1}^T \sum_{v'=1}^T \omega_v^{-2} \omega_{v'}^{-2} \sum_{k=-\infty}^{\infty} \mathbb{E} \left[ \epsilon_v(\alpha) \epsilon_{kT+v'}(\alpha) \left( \frac{\partial \epsilon_v(\alpha)}{\partial \alpha} \right) \left( \frac{\partial \epsilon_{kT+v'}(\alpha)}{\partial \alpha} \right)' \right]_{\alpha=\alpha_0}. \quad (4.13)$$

### 4.3 Diagnostic checking in weak PARMA models

After the estimation phase, the next important step consists in checking if the estimated model fits satisfactorily the data. In this section we derive the limiting distribution of the residual autocorrelations and that of the portmanteau statistics based on the standard and the self-normalized approaches in the framework of weak PARMA models (4.3). Note that the results stated in this section extend directly for models (4.2) with constants.

Let  $\hat{\epsilon}_{nT+v} = \hat{e}_{nT+v} = e_{nT+v}(\hat{\alpha}_{\text{LS}})$  be the least squares periodic residuals when  $p+q > 0$ . We use (4.8) to notice that we have  $\hat{e}_t = 0$  for  $t := nT+v \leq 0$  and  $t > NT$ . By (4.4), for  $n = 0, 1, \dots, N-1$ , it holds that

$$\hat{e}_{nT+v} = X_{nT+v} - \sum_{k=1}^p \hat{\phi}_k(v) X_{nT+v-k} + \sum_{l=1}^q \hat{\theta}_l(v) \hat{e}_{nT+v-l},$$

for  $1 \leq t \leq NT$ , with  $\hat{X}_t = 0$  for  $t \leq 0$  and  $\hat{X}_t = X_t$  for  $t \geq 1$ . We denote by

$$\Upsilon(v, l) = \frac{1}{N} \sum_{n=l+1}^N \epsilon_{nT+v} \epsilon_{nT+v-l} \text{ and } \hat{\Upsilon}(v, l) = \frac{1}{N} \sum_{n=l+1}^N \hat{e}_{nT+v} \hat{e}_{nT+v-l} \quad (4.14)$$

the periodic white noise "empirical" autocovariances and residuals autocovariances. It should be noted that  $\Upsilon(v, l)$  is not a computable statistic because it depends on the unobserved innovations  $\epsilon_{nT+v} = \epsilon_{nT+v}(\alpha_0)$ . For a fixed integer  $m \geq 1$  we let

$$\Upsilon_m(v) := (\Upsilon(v, 1), \dots, \Upsilon(v, m))' \text{ and } \hat{\Upsilon}_m(v) := (\hat{\Upsilon}(v, 1), \dots, \hat{\Upsilon}(v, m))',$$

be the vectors of "empirical" autocovariances and residuals autocovariances for a specified season  $v$ . Similarly the global "empirical" autocovariances and residuals autocovariances across all the seasons are denoted by

$$\Upsilon_{mT} := (\Upsilon_m(1), \dots, \Upsilon_m(T))' \text{ and } \hat{\Upsilon}_{mT} := (\hat{\Upsilon}_m(1), \dots, \hat{\Upsilon}_m(T))'.$$

Note that the mixing assumption **(A4)** will entail the asymptotic normality of  $\Upsilon_m(v)$  and  $\Upsilon_m(T)$ .

The theoretical and sample autocorrelations at lag  $l$  are respectively defined by

$$\rho(v, l) = \frac{\Gamma(v, l)}{\sqrt{\Gamma(v, 0)\Gamma(v-l, 0)}} \text{ and } \hat{\rho}(v, l) = \frac{\hat{\Upsilon}(v, l)}{\sqrt{\hat{\Upsilon}(v, 0)\hat{\Upsilon}(v-l, 0)}} \text{ with } \hat{\Upsilon}(v-l, 0) = \frac{1}{N} \sum_{n=1}^N \hat{e}_{nT+v-l}^2,$$

and where  $\Gamma(v, l) := \text{Cov}(\epsilon_{nT+v}, \epsilon_{nT+v-l})$  with  $\Gamma(v, 0) := \sigma_v^2$ . See for instance [25, 42] for some properties of  $\Gamma(v, l)$ . In the sequel we will also need these expressions:

$$\hat{\rho}_m(v) = (\hat{\rho}(v, 1), \dots, \hat{\rho}(v, m))' \text{ and } \hat{\rho}_{mT} = (\hat{\rho}_m(1), \dots, \hat{\rho}_m(T))',$$

where  $\hat{\rho}_m(\nu)$  denotes the vector of the first  $m$  sample autocorrelations for a specified season  $\nu$  and  $\hat{\rho}_m(T)$  is the vector of the global sample autocorrelations across all the seasons.

Based on the residual empirical autocorrelations, [14] have proposed a goodness-of-fit test, the so-called portmanteau test (BP), for strong ARMA models. A modification of their test has been proposed by [32] which is nowadays one of the most popular diagnostic checking tools in ARMA modeling of time series. To test simultaneously whether all residual autocorrelations at lags  $1, \dots, m$  of a PARMA model are equal to zero for a specified period  $\nu$ , the portmanteau of [14, 32] can be adapted as proposed by [37, 38] (see also [25]). Based on these results [37, 38] suggested the following portmanteau statistics defined by

$$Q_m^{\text{BP}}(\nu) = N \sum_{h=1}^m \hat{\rho}^2(\nu, h) \text{ and } Q_m^{\text{LBM}}(\nu) = N \sum_{h=1}^m \varrho(h, \nu, N, T) \hat{\rho}^2(\nu, h), \quad (4.15)$$

where  $\varrho(h, \nu, N, T)$  is called the Ljung-Box-McLeod (LBM) correction and is defined by

$$\varrho(h, \nu, N, T) = \begin{cases} (N+2)/(N-h/T) & \text{if } h \equiv 0 \pmod{T} \\ N/(N - \lfloor (h - \nu + T)/T \rfloor) & \text{otherwise.} \end{cases}$$

The statistic  $Q_m^{\text{LBM}}(\nu)$  has the same asymptotic chi-squared distribution as  $Q_m^{\text{BP}}(\nu)$  and has the reputation of doing better for small or medium sized sample (see [37, 38]). If the strong  $\text{PARMA}_T(p; q)$  model is adequate for a specified period  $\nu$ ,  $Q_m^{\text{BP}}(\nu)$  or  $Q_m^{\text{LBM}}(\nu)$  will be asymptotically distributed as a chi-square variable with  $m - (p + q)$  (and  $m - (p_\nu + q_\nu)$  for strong PARMA (4.3)) degrees of freedom. Note that the lag  $m$  used in (4.15) could be chosen to be different across the seasons but in most applications it is reasonable to use the same value of  $m$  for all seasons (see for instance [25]).

Across seasons the Portmanteau test statistics are asymptotically independent for  $\nu = 1, \dots, T$ . Consequently for the case of the Portmanteau test statistics in (4.15) an overall check to test if the residuals across all the seasons are periodic white noise is given by

$$Q_m^{\text{BP}} = \sum_{\nu=1}^T Q_m^{\text{BP}}(\nu) \text{ and } Q_m^{\text{LBM}} = \sum_{\nu=1}^T Q_m^{\text{LBM}}(\nu). \quad (4.16)$$

Thus if the strong  $\text{PARMA}_T(p; q)$  model is adequate the global statistics (4.16) will also be asymptotically distributed as a chi-squared variable with  $mT - T(p + q)$  (and  $mT - \sum_{\nu=1}^T (p_\nu + q_\nu)$  for strong PARMA (4.3)) degrees of freedom (see [25]).

The statistics (4.16) (resp. (4.15)) are usually used to test the following null hypothesis

**(H0)** :  $(X_{nT+\nu})_{n \in \mathbb{Z}}$  satisfies a  $\text{PARMA}_T(p; q)$  representation (resp. for a specified period  $\nu$ );

against the alternative

**(H1)** :  $(X_{nT+\nu})_{n \in \mathbb{Z}}$  does not admit a  $\text{PARMA}_T(p; q)$  representation (resp. for a specified period  $\nu$ ) or  $(X_{nT+\nu})_{n \in \mathbb{Z}}$  satisfies a  $\text{PARMA}_T(p'; q')$  representation (resp. for a specified period  $\nu$ ) with  $p' > p$  or  $q' > q$ .

For weak  $\text{PARMA}_T(p; q)$  models we will shown that the asymptotic distributions of the statistics defined in (4.15) and (4.16) are no longer chi-squared distributions but a mixture of chi-squared distributions weighted by eigenvalues of the asymptotic covariance matrix of the vector of autocorrelations.

### 4.3.1 Asymptotic distribution of the residual autocorrelations

By a Taylor expansion of  $\sqrt{N}\hat{\gamma}(\nu, h)$  we have

$$\sqrt{N}\hat{\gamma}(\nu, h) = \sqrt{N}\gamma(\nu, h) + \left( \mathbb{E} \left[ \epsilon_{nT+\nu-h} \frac{\partial}{\partial \alpha} \epsilon_{nT+\nu}(\alpha_0) \right] \right) \sqrt{N}(\hat{\alpha}_{\text{LS}} - \alpha_0) + o_{\mathbb{P}}(1). \quad (4.17)$$



We now remark that in Equation (4.17),  $\mathbb{E}[\epsilon_{nT+v-h}(\partial\epsilon_{nT+v}(\alpha_0)/\partial\alpha')]$  is the line  $h$  of the following matrix  $\Psi_m(\nu) \in \mathbb{R}^{m \times (p+q)T}$  defined by

$$\Psi_m(\nu) = \mathbb{E} \left\{ \left( \begin{array}{c} \epsilon_{nT+\nu-1} \\ \vdots \\ \epsilon_{nT+\nu-m} \end{array} \right) \frac{\partial \epsilon_{nT+\nu}}{\partial \alpha'} \right\}. \quad (4.18)$$

So for  $h = 1, \dots, m$ , Equation (4.17) becomes

$$\sqrt{N}\hat{\gamma}_m(\nu) = \left( \sqrt{N}\hat{\gamma}(\nu, 1), \dots, \sqrt{N}\hat{\gamma}(\nu, m) \right)' = \sqrt{N}\gamma_m(\nu) + \Psi_m(\nu)\sqrt{N}(\hat{\alpha}_{LS} - \alpha_0) + o_{\mathbb{P}}(1). \quad (4.19)$$

From (4.19) it is clear that the asymptotic distribution of the residual autocovariances  $\sqrt{N}\hat{\gamma}_m(\nu)$  is related to the asymptotic behavior of  $\sqrt{N}(\hat{\alpha}'_{LS} - \alpha'_0, \gamma'_m(\nu))'$ .

In view of [20, Theorem 1] and (A3) we have almost surely  $\hat{\alpha}_{LS} \rightarrow \alpha_0 \in \overset{\circ}{\Theta}$ . Thus  $\partial Q_N^{\omega^2}(\hat{\alpha}_{LS})/\partial\alpha = 0$  for sufficiently large  $N$  and a Taylor expansion gives

$$\sqrt{N}\frac{\partial}{\partial\alpha}O_N^{\omega^2}(\alpha_0) + J(\alpha_0, \omega^2)\sqrt{N}(\hat{\alpha}_{LS} - \alpha_0) = o_{\mathbb{P}}(1), \text{ where } O_N^{\omega^2}(\alpha) = \frac{1}{2N} \sum_{n=0}^{N-1} \sum_{\nu=1}^T \omega_{\nu}^{-2} \epsilon_{nT+\nu}^2(\alpha). \quad (4.20)$$

Consequently under the assumption that  $J(\alpha_0, \omega^2)$  is invertible and from (4.20) we deduce that

$$\sqrt{N}(\hat{\alpha}_{LS} - \alpha_0) = -J(\alpha_0, \omega^2)^{-1} \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \sum_{\nu=1}^T \omega_{\nu}^{-2} \epsilon_{nT+\nu} \frac{\partial \epsilon_{nT+\nu}(\alpha_0)}{\partial \alpha} + o_{\mathbb{P}}(1). \quad (4.21)$$

**Proposition 4.3.1** *Assume that  $p+q > 0$ . Under the assumptions (A0), (A1), (A2), (A3) and (A4), the random vector*

$$\sqrt{N} \left( (\hat{\alpha}_{LS} - \alpha_0)', \gamma'_m(\nu) \right)'$$

*has a limiting centered normal distribution with covariance matrix*

$$\Xi(\nu) = \begin{pmatrix} \Omega^{LS} & \Sigma_{\hat{\alpha}, \gamma_m(\nu)} \\ \Sigma'_{\hat{\alpha}, \gamma_m(\nu)} & \Sigma_{\gamma_m(\nu)} \end{pmatrix} = \sum_{h=-\infty}^{\infty} \mathbb{E} \left[ W_{nT+\nu} W'_{(n-h)T+\nu} \right],$$

where from (4.14) and (4.21) we have

$$W_{nT+\nu} = \begin{pmatrix} W_{1, nT+\nu} \\ W_{2, nT+\nu} \end{pmatrix} = \begin{pmatrix} -J(\alpha_0, \omega^2)^{-1} \sum_{\nu=1}^T \omega_{\nu}^{-2} \epsilon_{nT+\nu} \frac{\partial \epsilon_{nT+\nu}(\alpha_0)}{\partial \alpha} \\ (\epsilon_{nT+\nu-1}, \dots, \epsilon_{nT+\nu-m})' \epsilon_{nT+\nu} \end{pmatrix}. \quad (4.22)$$

The proof of the proposition is given in Subsection 4.6.1 of the appendix.

The following theorem which is an extension of the results given in [21] provides the limit distribution of the residual autocovariances and autocorrelations of weak PARMA models for a specified season  $\nu$ . We also provide the limit distribution of the global residual autocovariances and autocorrelations of weak PARMA models across all the seasons. Let

$$D_m(\nu) = \text{Diag}(\sigma_{\nu}\sigma_{\nu-1}, \dots, \sigma_{\nu}\sigma_{\nu-m})$$

and

$$D_{mT} = \text{Diag}(\sigma_1\sigma_{1-1}, \dots, \sigma_1\sigma_{1-m}, \dots, \sigma_T\sigma_{T-1}, \dots, \sigma_T\sigma_{T-m}).$$



**Theorem 4.3.1** *Under the assumptions of Proposition 4.3.1 we have*

$$\sqrt{N}\hat{\gamma}_m(\mathbf{v}) \xrightarrow[N \rightarrow \infty]{d} \mathcal{N}(0, \nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v})) \quad (4.23)$$

where

$$\begin{aligned} \nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v}) &= \Sigma_{\gamma_m(\mathbf{v})} + \Psi_m(\mathbf{v})\Omega^{\text{LS}}\Psi'_m(\mathbf{v}) + \Psi_m(\mathbf{v})\Sigma_{\hat{\alpha}, \gamma_m(\mathbf{v})} + \Sigma'_{\hat{\alpha}, \gamma_m(\mathbf{v})}\Psi'_m(\mathbf{v}) \text{ when } p + q > 0 \\ \text{and } \nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v}) &= \Sigma_{\gamma_m(\mathbf{v})} \text{ when } p = q = 0. \end{aligned}$$

Then:

$$\sqrt{N}\hat{\gamma}_{mT} \xrightarrow[N \rightarrow \infty]{d} \mathcal{N}(0, \nabla_{\hat{\gamma}}) \quad (4.24)$$

where the asymptotic covariance matrix  $\nabla_{\hat{\gamma}}$  is a block matrix, with the asymptotic variances given by  $\nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v})$ , for  $\mathbf{v} = 1, \dots, T$  and the asymptotic covariances given by

$$\nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v}') = \Sigma_{\gamma_m}(\mathbf{v}, \mathbf{v}') + \Psi_m(\mathbf{v})\Omega^{\text{LS}}\Psi'_m(\mathbf{v}') + \Psi_m(\mathbf{v})\Sigma_{\hat{\alpha}, \gamma_m}(\mathbf{v}, \mathbf{v}') + \Sigma'_{\hat{\alpha}, \gamma_m(\mathbf{v})}\Psi'_m(\mathbf{v}'), \quad \mathbf{v} \neq \mathbf{v}'.$$

We also have

$$\sqrt{N}\hat{\rho}_m(\mathbf{v}) \xrightarrow[N \rightarrow \infty]{d} \mathcal{N}(0, \nabla_{\hat{\rho}}(\mathbf{v}, \mathbf{v})) \text{ where } \nabla_{\hat{\rho}}(\mathbf{v}, \mathbf{v}) = D_m^{-1}(\mathbf{v})\nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v})D_m^{-1}(\mathbf{v}) \quad (4.25)$$

and

$$\sqrt{N}\hat{\rho}_{mT} \xrightarrow[N \rightarrow \infty]{d} \mathcal{N}(0, \nabla_{\hat{\rho}}) \text{ where } \nabla_{\hat{\rho}} = D_{mT}^{-1}\nabla_{\hat{\gamma}}D_{mT}^{-1}. \quad (4.26)$$

The detailed proof of this result is postponed to the Subsection 4.6.2 of Appendix.

The asymptotic variance matrices  $\nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v})$  and  $\nabla_{\hat{\rho}}(\mathbf{v}, \mathbf{v})$  depend on the unknown matrices  $\Xi(\mathbf{v})$ ,  $\Psi_m(\mathbf{v})$  and the scalar  $\sigma_{\mathbf{v}-h}$  for  $h = 1, \dots, m$ . Matrix  $\Psi_m(\mathbf{v})$  and  $\sigma_{\mathbf{v}-h}$  can be estimated by its empirical counterpart, respectively by

$$\hat{\Psi}_m(\mathbf{v}) = \frac{1}{N} \sum_{n=0}^{N-1} \left\{ (\hat{\epsilon}_{nT+\mathbf{v}-1}, \dots, \hat{\epsilon}_{nT+\mathbf{v}-m})' \frac{\partial \hat{\epsilon}_{nT+\mathbf{v}}}{\partial \alpha'} \right\} \text{ and } \hat{\sigma}_{\mathbf{v}-h} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} \hat{\epsilon}_{nT+\mathbf{v}-h}^2}.$$

Similarly the matrix  $J(\alpha_0, \omega^2)$  can be estimated empirically by the square matrix  $J_N(\hat{\alpha}_{\text{LS}}, \omega^2)$  of order  $(p+q)T$  defined by:

$$J_N(\hat{\alpha}_{\text{LS}}, \omega^2) = \sum_{\mathbf{v}=1}^T \omega_{\mathbf{v}}^2 \frac{1}{N} \sum_{n=0}^{N-1} \left\{ \frac{\partial}{\partial \alpha} e_{nT+\mathbf{v}}(\hat{\alpha}_{\text{LS}}) \right\} \left\{ \frac{\partial}{\partial \alpha} e_{nT+\mathbf{v}}(\hat{\alpha}_{\text{LS}}) \right\}'. \quad (4.27)$$

The estimation of the long-run variance  $\Xi(\mathbf{v})$  is more complicated. In the econometric literature the nonparametric kernel estimator, also called heteroscedastic autocorrelation consistent (HAC) estimator (see [39], or [3]), is widely used to estimate covariance matrices of the form  $\Xi(\mathbf{v})$ . Interpreting  $(2\pi)^{-1}\Xi(\mathbf{v})$  as the spectral density of the multivariate stationary periodic process  $(W_{nT+\mathbf{v}})_{n \in \mathbb{Z}} = (W'_{1, nT+\mathbf{v}}, W'_{2, nT+\mathbf{v}})'_{n \in \mathbb{Z}}$  evaluated at frequency 0 (see [15], p. 459 for more details), an alternative method consists in using a parametric vector AR (VAR) estimate of the spectral density of  $(W_{nT+\mathbf{v}})_{n \in \mathbb{Z}}$ .

The stationary process  $(W_{nT+\mathbf{v}})_{n \in \mathbb{Z}}$  admits the Wold decomposition

$$W_{nT+\mathbf{v}} = u_{nT+\mathbf{v}} + \sum_{i=1}^{\infty} \omega_i(\mathbf{v}) u_{nT+\mathbf{v}-i},$$

where  $u_{nT+\mathbf{v}}$  is a  $((p+q)T+m)$ -variate weak periodic white noise. Assume that the covariance matrix  $\Sigma_u(\mathbf{v}) := \text{Var}(u_{nT+\mathbf{v}})$  is non-singular, that  $\sum_{i=1}^{\infty} \|\omega_i(\mathbf{v})\| < \infty$ , where  $\|\cdot\|$  denotes any norm

on the space of the real  $((p+q)T+m) \times ((p+q)T+m)$  matrices, and that  $\det \{I_{(p+q)T+m} + \sum_{i=1}^{\infty} \omega_i(\nu) z^i\} \neq 0$  if  $|z| \leq 1$ . Then  $(W_{nT+\nu})_{n \in \mathbb{Z}}$  admits an AR( $\infty$ ) representation (see [2]) of the form

$$\Phi_\nu(L)W_{nT+\nu} = W_{nT+\nu} + \sum_{i=1}^{\infty} \Phi_i(\nu)W_{nT+\nu-i} = u_{nT+\nu}, \quad (4.28)$$

such that  $\sum_{i=1}^{\infty} \|\Phi_i(\nu)\| < \infty$  and  $\det \{\Phi_\nu(z)\} \neq 0$  if  $|z| \leq 1$ .

This parametric approach has been studied by [6] (see also [18]) is based on the expression

$$\Xi(\nu) = \Phi_\nu^{-1}(1)\Sigma_u(\nu)\Phi_\nu^{-1'}(1). \quad (4.29)$$

Since  $W_{nT+\nu}$  is unobserved we introduce  $\hat{W}_{nT+\nu} \in \mathbb{R}^{T(p+q)+m}$  obtained by replacing  $\epsilon_{nT+\nu}(\alpha_0)$  by  $e_{nT+\nu}(\hat{\alpha}_{LS})$  and  $J(\alpha_0, \omega^2)$  by its empirical or observable counterpart  $J_N(\hat{\alpha}_{LS}, \omega^2)$  in (4.22). Let  $\hat{\Phi}_{\nu,r}(z) = I_{(p+q)T+m} + \sum_{i=1}^r \hat{\Phi}_{r,i}(\nu)z^i$ , where  $\hat{\Phi}_{r,1}(\nu), \dots, \hat{\Phi}_{r,r}(\nu)$  denote the coefficients of the least squares regression of  $\hat{W}_{nT+\nu}$  on  $\hat{W}_{nT+\nu-1}, \dots, \hat{W}_{nT+\nu-r}$  for a specified season  $\nu$ . Let  $\hat{u}_{r,nT+\nu}$  be the residuals of this regression, and let  $\hat{\Sigma}_{\hat{u}_r}(\nu)$  be the empirical variance of  $\hat{u}_{r,1}, \dots, \hat{u}_{r,N}$ .

In the case of linear processes with independent innovations, [6] has shown that the spectral density can be consistently estimated by fitting autoregressive models of order  $r = r(N)$ , whenever  $r$  tends to infinity and  $r^3/N$  tends to 0 as  $N$  tends to infinity. There are differences with [6]:  $(W_{nT+\nu})_{n \in \mathbb{Z}}$  is multivariate, is not directly observed and is replaced by  $(\hat{W}_{nT+\nu})_{n \in \mathbb{Z}}$ . It is shown that this result remains valid for the multivariate linear process  $(W_{nT+\nu})_{n \in \mathbb{Z}}$  with non-independent innovations (see [10, 11, 21] for references in weak (multivariate) ARMA models). We will extend these results to weak PARMA models. We are now able to state the following theorem.

**Theorem 4.3.2** *In addition to the assumptions of Theorem 4.3.1, assume that the process  $(W_{nT+\nu})_{n \in \mathbb{Z}}$  defined in (4.22) admits a periodic VAR( $\infty$ ) representation (4.28) in which the roots of  $\det \Phi(z) = 0$  are outside the unit disk,  $\|\Phi_i\| = o(i^{-2})$ , and  $\Sigma_u(\nu) = \text{Var}(u_{nT+\nu})$  is non-singular. Moreover we assume that  $\mathbb{E} \|\epsilon_n\|^{8+4\kappa} < \infty$  and  $\sum_{k=0}^{\infty} \{\alpha_\epsilon(k)\}^{\kappa/(2+\kappa)} < \infty$  for some  $\kappa > 0$ . Then the spectral estimator of  $\Xi(\nu)$ :*

$$\hat{\Xi}^{\text{SP}}(\nu) := \hat{\Phi}_{\nu,r}^{-1}(1)\hat{\Sigma}_{\hat{u}_r}(\nu)\hat{\Phi}_{\nu,r}^{-1'}(1) \rightarrow \Xi(\nu)$$

*in probability when  $r = r(N) \rightarrow \infty$  and  $r^3/N \rightarrow 0$  as  $N \rightarrow \infty$ .*

The proof of this theorem is similar to the proof of Theorem 5.2 in [21] and it is omitted.

From Theorem 4.3.1 we can deduce the following result which gives the exact limiting distribution of the standard portmanteau statistics (4.15) and (4.16) under general assumptions on the innovation process of the fitted PARMA( $p; q$ ) model.

**Theorem 4.3.3** *Under Assumptions of Theorem 4.3.1 and H0, the statistics  $Q_m^{\text{LB}}(\nu)$  and  $Q_m^{\text{LBM}}(\nu)$  defined in (4.15) converge in distribution, as  $N \rightarrow \infty$ , to*

$$Z_m^\nu(\xi_m) = \sum_{i=1}^m \xi_{i,m}^\nu Z_i^2$$

*where  $\xi_m^\nu = (\xi_{1,m}^\nu, \dots, \xi_{m,m}^\nu)'$  is the vector of the eigenvalues of the matrix  $\nabla_{\hat{\rho}}(\nu, \nu)$  given in (4.25) and  $Z_1, \dots, Z_m$  are independent  $\mathcal{N}(0, 1)$  variables.*

*The asymptotic distribution of the global portmanteau test statistics (that takes into account all the seasons)  $Q_m^{\text{LB}}$  and  $Q_m^{\text{LBM}}$  defined in (4.16) is also a weighted sum of chi-square random variables:*

$$Z_m(\xi_m) = \sum_{i=1}^{mT} \xi_{i,mT} Z_i^2$$

*where  $\xi_{mT} = (\xi_{1,mT}, \dots, \xi_{mT,mT})'$  denotes the vector of the eigenvalues of the matrix  $\nabla_{\hat{\rho}}$  given in (4.26).*

As mentioned below, Theorem 4.3.3 shows that: for the asymptotic distribution of (4.15) and (4.16) the  $\chi_{m-(p+q)}^2$  and  $\chi_{mT-(p+q)T}^2$  approximations are no longer valid in the framework of weak PARMA( $p; q$ ) models. The trues asymptotic distributions depend on the periodic nuisance parameters involving  $\sigma_{v-h}$  for  $h = 1, \dots, m$  and the elements of  $\Xi(v)$ . Consequently in order to obtain the asymptotic distribution of the portmanteau statistics (4.15) (resp. (4.16)) under weak assumptions on the periodic noise, one needs a consistent estimator of the asymptotic covariance matrix  $\nabla_{\hat{\rho}}(v, v)$  (resp.  $\nabla_{\hat{\rho}}$ ). We let  $\hat{\nabla}_{\hat{\rho}}(v, v)$  (resp.  $\hat{\nabla}_{\hat{\rho}}$ ) the matrix obtained by replacing  $\Xi(v)$  by  $\hat{\Xi}(v)$ ,  $\Psi_m(v)$  by  $\hat{\Psi}_m(v)$  and  $\sigma_{v-h}$  by  $\hat{\sigma}_{v-h}$  in  $\nabla_{\hat{\rho}}(v, v)$  (resp.  $\nabla_{\hat{\rho}}$ ). Denote by  $\hat{\xi}_m^v = (\hat{\xi}_{1,m}^v, \dots, \hat{\xi}_{m,m}^v)'$  (resp.  $\hat{\xi}_m = (\hat{\xi}_{1,mT}, \dots, \hat{\xi}_{mT,mT})'$ ) the vector of the eigenvalues of  $\hat{\nabla}_{\hat{\rho}}(v, v)$  (resp.  $\hat{\nabla}_{\hat{\rho}}$ ). At the asymptotic level  $\alpha$  and for a specified season  $v$ , the LBM test (resp. the BP test) consists in rejecting the adequacy of the weak PARMA<sub>T</sub>( $p; q$ ) model when

$$\lim_{N \rightarrow \infty} \mathbb{P} \left( Q_m^{\text{BP}}(v) > S_m^v(1 - \alpha) \right) = \lim_{N \rightarrow \infty} \mathbb{P} \left( Q_m^{\text{LBM}}(v) > S_m^v(1 - \alpha) \right) = \alpha,$$

where  $S_m^v(1 - \alpha)$  is such that  $\mathbb{P} \{ Z_m^v(\hat{\xi}_m^v) > S_m^v(1 - \alpha) \} = \alpha$ . We emphasize the fact that the proposed modified versions of the BP and LBM statistics are more difficult to implement because their critical values have to be computed from the data while the critical values of the standard method are simply deduced from a  $\chi^2$ -table. We shall evaluate the  $p$ -values

$$\mathbb{P} \{ Z_m(\hat{\xi}_m^v) > Q_m^{\text{BP}}(v) \} \quad \text{and} \quad \mathbb{P} \{ Z_m(\hat{\xi}_m^v) > Q_m^{\text{LBM}}(v) \}, \quad \text{with} \quad Z_m(\hat{\xi}_m^v) = \sum_{i=1}^m \hat{\xi}_{i,m}^v Z_i^2,$$

by means of the Imhof algorithm (see [27]) or other exact methods.

The test procedures for  $Q_m^{\text{BP}}$  and  $Q_m^{\text{LBM}}$  defined in (4.16) are similar but they are based on the  $mT$  empirical eigenvalues of  $\hat{\nabla}_{\hat{\rho}}$ .

### 4.3.2 Self-normalized asymptotic distribution of the residual autocorrelations

The nonparametric kernel estimator (see [3, 39]) used to estimate the matrix  $\Xi(v)$  causes serious difficulties regarding the choice of the sequence of weights. The parametric approach based on an autoregressive estimate of the spectral density of  $(W_{nT+1})_{n \in \mathbb{Z}}$  studied in [6, 18] is also facing the problem of choosing the truncation parameter. So the choice of the order of truncation is often crucial and difficult.

In this section, we propose an alternative method where we do not estimate an asymptotic covariance matrix which is an extension to the results obtained by [13]. It is based on a self-normalization approach to construct a test-statistic which is asymptotically distribution-free under the null hypothesis. This approach has been studied by [13] in the weak ARMA case by proposing new portmanteau statistics. In this case the critical values are not computed from the data since they are tabulated by [33]. In some sense this method is finally closer to the standard method in which the critical values are simply deduced from a  $\mathcal{X}^2$ -table. The idea comes from [33] and has been already extended by [13], [30], [43], [44] and [45] to name a few in more general frameworks. See also [46] for a review on some recent developments on the inference of time series data using the self-normalized approach.

Let  $(B_K(r))_{r \geq 0}$  be a  $K$ -dimensional Brownian motion starting from 0. For  $K \geq 1$ , we denote  $\mathcal{U}_K$  the random variable defined by

$$\mathcal{U}_K = B_K'(1) V_K^{-1} B_K(1) \quad \text{where} \quad V_K = \int_0^1 (B_K(r) - r B_K(1)) (B_K(r) - r B_K(1))' dr.$$

We denote  $\Lambda(v)$  the matrix in  $\mathbb{R}^{m \times (T(p+q)+m)}$  defined in block formed by  $\Lambda(v) = (\Psi_m(v) | I_m)$ . In view of (4.19) and (4.21) we deduce that  $\sqrt{N} \hat{\gamma}_m(v) = N^{-1/2} \sum_{n=1}^{N-1} \Lambda(v) W_{nT+v} + o_{\mathbb{P}}(1)$ .

Contrary to Subsection 4.3.1 we do not rely on the classical method that would consist in estimating the asymptotic covariance matrix of  $(\Lambda(\nu)W_{nT+\nu})_{n \in \mathbb{Z}}$ . We need to apply the functional central limit theorem holds for the process  $(W_{nT+\nu})_{n \in \mathbb{Z}}$  (see Lemma 1 in [33]).

Finally we define the normalization matrix  $C_m(\nu, \nu) \in \mathbb{R}^{m \times m}$  for a specified season  $\nu$  by

$$C_m(\nu, \nu) = \frac{1}{N^2} \sum_{n=0}^{N-1} S_n(\nu) S_n'(\nu) \text{ where } S_n(\nu) = \sum_{j=0}^n (\Lambda(\nu)W_{jT+\nu} - \Upsilon_m(\nu)). \quad (4.30)$$

Similarly we denote by  $C_{mT} \in \mathbb{R}^{mT \times mT}$  the global normalization block matrix across all the seasons with diagonal block given by (4.30) for  $\nu = 1, \dots, T$  and

$$C_m(\nu, \nu') = \frac{1}{N^2} \sum_{n=0}^{N-1} S_n(\nu) S_n'(\nu') \text{ where } S_n(\nu') = \sum_{j=0}^n (\Lambda(\nu')W_{jT+\nu'} - \Upsilon_m(\nu')) \text{ for } \nu \neq \nu'.$$

To ensure the invertibility of the normalization matrix  $C_m(\nu, \nu)$  we need the following technical assumption on the distribution of  $\epsilon_n$ .

**(A5):** The process  $(\epsilon_n)_{n \in \mathbb{Z}}$  has a positive density on some neighborhood of zero.

The following proposition gives the invertibility of the normalization matrices  $C_m(\nu, \nu)$  and  $C_{mT}$ .

**Proposition 4.3.2** *Under the assumptions of Theorem 4.3.1 and (A5), the matrices  $C_m(\nu, \nu)$  and  $C_{mT}$  are almost surely non singular.*

The proof of this proposition is given in Subsection 4.6.3 of the appendix.

The following theorem states the asymptotic distributions of the residual autocovariances and autocorrelations of weak PARMA models for a specified season  $\nu$ . We also provide the limit distribution of the global residual autocovariances and autocorrelations of weak PARMA models across all the seasons.

**Theorem 4.3.4** *We assume that  $p + q > 0$ . Under Assumptions of Theorem 4.3.1, (A5) and under the null hypothesis (H0) we have for a specified season  $\nu$*

$$N\hat{\Upsilon}'_m(\nu)C_m^{-1}(\nu, \nu)\hat{\Upsilon}_m(\nu) \xrightarrow[N \rightarrow \infty]{d} \mathcal{U}_m \text{ and } N\hat{\rho}'_m(\nu)D_m(\nu)C_m^{-1}(\nu, \nu)D_m(\nu)\hat{\rho}_m(\nu) \xrightarrow[N \rightarrow \infty]{d} \mathcal{U}_m.$$

For the global residual autocovariances and autocorrelations across all the seasons we also have

$$N\hat{\Upsilon}'_{mT}C_{mT}^{-1}\hat{\Upsilon}_{mT} \xrightarrow[N \rightarrow \infty]{d} \mathcal{U}_{mT} \text{ and } N\hat{\rho}'_{mT}D_{mT}C_{mT}^{-1}D_{mT}\hat{\rho}_{mT} \xrightarrow[N \rightarrow \infty]{d} \mathcal{U}_{mT}.$$

The proof of this result is postponed to Section 4.6.

**Remarque 4.3.1** *When  $p = q = 0$  we do no more need to estimate the unknown parameter  $\alpha_0$ . Thus a careful reading of the proofs shows that the vector  $W_{nT+\nu}$  is replaced by*

$$\tilde{W}_{nT+\nu} = W_{2,nT+\nu} = (\epsilon_{nT+\nu}\epsilon_{nT+\nu-1}, \dots, \epsilon_{nT+\nu}\epsilon_{nT+\nu-m})'.$$

Then we generalized for weak periodic noise the result obtained by [13].

Of course the above theorem is useless for practical purpose because the normalization matrix  $C_m(\nu, \nu)$ ,  $C_{mT}$  and the parameter  $\sigma_{\nu-h}$  are not observable. This gap will be fixed below (see Theorem 4.3.5) when one replaces the matrix  $C_m(\nu, \nu)$ ,  $C_{mT}$  and the scalar  $\sigma_{\nu-h}$  by their empirical or observable counterparts. Then we denote

$$\hat{C}_m(\nu, \nu) = \frac{1}{N^2} \sum_{n=0}^{N-1} \hat{S}_n(\nu) \hat{S}_n'(\nu) \text{ where } \hat{S}_n(\nu) = \sum_{j=0}^n (\hat{\Lambda}(\nu)\hat{W}_{jT+\nu} - \hat{\Upsilon}_m(\nu)),$$

with  $\hat{\Lambda}(\nu) = (\hat{\Psi}_m(\nu)|I_m)$  and where  $\hat{W}_{nT+\nu}$  and  $\hat{\sigma}_{\nu-h}$  are defined in Subsection 4.3.1.

We are able to state the following result which is the applicable counterpart of Theorem 4.3.4.

**Theorem 4.3.5** *Under the assumptions of Theorem 4.3.4, for a specified season  $\nu$  we have*

$$N\hat{\gamma}'_m(\nu)\hat{C}_m^{-1}(\nu, \nu)\hat{\gamma}_m(\nu) \xrightarrow[N \rightarrow \infty]{d} \mathcal{U}_m \text{ and } Q_m^{\text{SN}}(\nu) =: N\hat{\rho}'_m(\nu)\hat{D}_m(\nu)\hat{C}_m^{-1}(\nu, \nu)\hat{D}_m(\nu)\hat{\rho}_m(\nu) \xrightarrow[N \rightarrow \infty]{d} \mathcal{U}_m.$$

We also have for the global seasons

$$N\hat{\gamma}'_{mT}\hat{C}_{mT}^{-1}\hat{\gamma}_{mT} \xrightarrow[N \rightarrow \infty]{d} \mathcal{U}_{mT} \text{ and } Q_m^{\text{SN}} =: N\hat{\rho}'_{mT}\hat{D}_{mT}\hat{C}_{mT}^{-1}\hat{D}_{mT}\hat{\rho}_{mT} \xrightarrow[N \rightarrow \infty]{d} \mathcal{U}_{mT}.$$

The proof of this result is postponed in Subsection 4.6.5 of Appendix.

Based on the above result we propose a modified version of the LBM statistic for a specified season  $\nu$  when one uses the statistic

$$\tilde{Q}_m^{\text{SN}}(\nu) = N\hat{\rho}'_m(\nu)M_m^{1/2}(\nu, N, T)\hat{D}_m(\nu)\hat{C}_m^{-1}(\nu, \nu)\hat{D}_m(\nu)M_m^{1/2}(\nu, N, T)\hat{\rho}_m(\nu), \quad (4.31)$$

where the matrix  $M_m(\nu, N, T) \in \mathbb{R}^{m \times m}$  is diagonal with  $(\varrho(1, \nu, N, T), \dots, \varrho(m, \nu, N, T))$  as diagonal terms.

Similarly we propose a modified version of the LBM statistic for the global seasons when one uses the following statistic

$$\tilde{Q}_m^{\text{SN}} = N\hat{\rho}'_{mT}M_{mT}^{1/2}\hat{D}_{mT}\hat{C}_{mT}^{-1}\hat{D}_{mT}M_{mT}^{1/2}\hat{\rho}_{mT}, \quad (4.32)$$

where the matrix  $M_{mT} \in \mathbb{R}^{mT \times mT}$  is diagonal with

$$(\varrho(1, 1, N, T), \dots, \varrho(m, 1, N, T), \dots, \varrho(1, T, N, T), \dots, \varrho(m, T, N, T))$$

as diagonal terms.

At the asymptotic level  $\alpha$  and for a specified season  $\nu$ , the self-normalized LBM test (resp. the BP test) consists in rejecting the adequacy of the weak PARMA $_T(p; q)$  model when

$$\tilde{Q}_m^{\text{SN}}(\nu) > \mathcal{U}_m(1 - \alpha) \quad (\text{resp. } Q_m^{\text{SN}}(\nu) > \mathcal{U}_m(1 - \alpha)),$$

where the critical values  $\mathcal{U}_m(1 - \alpha)$  are tabulated in [33, Table 1]. Similarly for the global test we have

$$\tilde{Q}_m^{\text{SN}} > \mathcal{U}_{mT}(1 - \alpha) \quad (\text{resp. } Q_m^{\text{SN}} > \mathcal{U}_{mT}(1 - \alpha)).$$

## 4.4 Numerical illustration

By means of Monte Carlo experiments, we investigate the finite sample properties of the tests introduced in this paper. The numerical illustrations of this section are made with the free statistical R software (see <https://www.r-project.org/>).

We indicate the conventions that we adopt in the discussion and in the tables. One refers to

- BP $_S(\nu)$  for the standard Box-Pierce test using the statistic  $Q_m^{\text{BP}}(\nu)$ ;
- LBM $_S(\nu)$  for the standard Ljung-Box-McLeod test using the statistic  $Q_m^{\text{LBM}}(\nu)$ ;
- BP $_{\text{SN}}(\nu)$  for the modified test using the statistic  $Q_m^{\text{SN}}(\nu)$  with the values of the quantiles of  $\mathcal{U}_m$  simulated in Table 1 of [33];
- LBM $_{\text{SN}}(\nu)$  for the modified test using the statistic  $\tilde{Q}_m^{\text{SN}}(\nu)$  with the values of the quantiles of  $\mathcal{U}_m$ ;
- BP $_W(\nu)$  for the modified Box-Pierce test (proposed in Section 4.3.1) using the statistic  $Q_m^{\text{BP}}(\nu)$ ;

- $LBM_W(v)$  for the modified Ljung-Box-McLeod test (see Section 4.3.1) using the statistic  $Q_m^{LBM}(v)$ ;
- $BP_S$  for the standard global Box-Pierce test using the statistic  $Q_m^{BP}$ ;
- $LBM_S$  for the standard global Ljung-Box-McLeod test using the statistic  $Q_m^{LBM}$ ;
- $BP_{SN}$  for the modified global test using the statistic  $Q_m^{SN}$  with the values of the quantiles of  $\mathcal{U}_{mT}$ ;
- $LBM_{SN}$  for the modified global test using the statistic  $\tilde{Q}_m^{SN}$  with the values of the quantiles of  $\mathcal{U}_{mT}$ ;
- $BP_W$  for the modified global Box-Pierce test (proposed in Section 4.3.1) using the statistic  $Q_m^{BP}$ ;
- $LBM_W$  for the modified global Ljung-Box-McLeod test (see Section 4.3.1) using the statistic  $Q_m^{LBM}$ .

We will see in the tables that the numerical results using the Ljung-Box-McLeod tests are very close to those of the Box-Pierce tests. Nevertheless they are still presented here for the sake of completeness.

We study numerically the behavior of the quasi-generalized least squares (QLS) estimator for a  $PARMA_2(1, 1)$  models of the form

$$X_{2n+v} - \phi(v)X_{2n+v-1} = \epsilon_{2n+v} - \theta(v)\epsilon_{2n+v-1}, \quad v = 1, 2 \quad (4.33)$$

where the unknown parameter is  $\alpha_0 = (\phi(1), \phi(2), \theta(1), \theta(2))'$ . Three different periodic noises are considered. First we assume that in (4.33) the innovation process  $(\epsilon_{2n+v})_{n \in \mathbb{Z}}$  is an iid centered Gaussian process with variance  $\sigma_v^2 = \mathbb{E}(\epsilon_{2n+v}^2)$  where  $\sigma_1 = 0.9$  and  $\sigma_2 = 1.5$  which corresponds to the strong PARMA case. For the weak PARMA cases, we consider that in (4.33) the innovation process  $(\epsilon_{2n+v})_{n \in \mathbb{Z}}$  follows firstly a periodic ARCH(1) given by

$$\epsilon_{2n+v} = \sqrt{\omega(v) + a(v)\epsilon_{2n+v-1}^2} \eta_{2n+v}, \quad v = 1, 2 \quad (4.34)$$

with  $(\omega(1), \omega(2)) = (0.2, 0.4)$ ,  $(a(1), a(2)) = (0.4, 0.45)$  and where  $(\eta_{2n+v})_{n \in \mathbb{Z}}$  is a sequence of iid centered Gaussian random variables with variance 1. Secondly a noise defined by a following PGARCH(1, 1)

$$\begin{cases} \epsilon_{2n+v} = \sqrt{h_{2n+v}} \eta_{2n+v} \\ h_{2n+v} = \omega(v) + a(v)\epsilon_{2n+v-1}^2 + b(v)h_{2n+v-1}, \quad v = 1, 2 \end{cases} \quad (4.35)$$

where  $(\omega(1), \omega(2)) = (0.2, 0.4)$ ,  $(a(1), a(2)) = (0.1, 0.15)$  and  $(b(1), b(2)) = (0.83, 0.82)$ .

#### 4.4.1 Empirical size

For each of these three different periodic white noises 1,000 independent replications of length  $(N + 100) \times 2$ . These sequences were plugged into the model (4.33) yielding 1,000 independent replications of the periodic process  $(X_{2n+v})_n$  of length  $(N + 100) \times 2$ . Initial values were set to zero and in order to achieve periodic stationarity the first 200 observations were dropped. For each replication of size  $NT = N \times 2$  we use the QLS method to estimate the corresponding coefficient  $\alpha_0$ . After estimating Model (4.33) we apply portmanteau test to the residuals for different values of  $m \in \{1, \dots, 12\}$ , where  $m$  is the number of autocorrelations used in the portmanteau test statistic. In the experiment we considered two values of  $N$ : 400 and 2,000. The nominal asymptotic level considered of the tests is  $\alpha = 5\%$ . Note that for this level, the empirical relative frequency of rejection size over the 1,000 independent replications should vary within the confidences intervals [3.6%, 6.4%] with probability 95% and [3.3%, 6.9%] with probability 99%.



### Strong PARMA<sub>2</sub>(1, 1) model case

Table 4.1 (resp. Table 4.4 and Table 4.7) displays relative frequencies (in %) of rejection (over the 1,000 replications) of the null hypothesis **H0** that the data generating process (DGP for short) follows a strong PARMA<sub>2</sub>(1, 1) given by (4.33) with the parameter  $\alpha_0 = (0, 0, 0, 0)'$  (resp.  $\alpha_0 = (0.8, 0.9, 0, 0)'$  and  $\alpha_0 = (0.8, 0.9, -0.5, -0.45)'$ ). For the strong periodic white noise, all the relative rejection frequencies (for global or a specified season  $\nu$ ) are inside the significant limits. Thus the error of first kind is well controlled by all the tests in this case. Note that the empirical size is not available (n.a.) for the standard BP and LBM tests because these tests are not applicable to  $m = 1$  (see Table 4.4) and  $m \leq 2$  (see Table 4.7). As expected the observed relative rejection frequency of the standard tests is very far from the nominal  $\alpha = 5\%$  for small  $m$  when  $\alpha_0 \neq (0, 0, 0, 0)'$ . The results are worse for  $N = 2,000$  than for  $N = 400$ . This is in accordance with the results in the literature on the strong PARMA models. The theory that the asymptotic distributions of (4.15) and (4.16), namely,  $\chi_{m-(p+q)}^2$  and  $\chi_{2m-2(p+q)}^2$  approximations are better for larger  $m$  is confirmed. In contrast, the proposed tests well control the error of first kind, even when  $m$  is small.

From these examples we draw the conclusion that the proposed versions are preferable to the standard ones, when the number  $m$  of autocorrelations used is small.

### Weak PARMA<sub>2</sub>(1, 1) model case

We now repeat the same experiment on Model (4.33) by assuming that the innovation process  $(\epsilon_{2n+\nu})_n$  follows first (4.34) and secondly (4.35).

Tables 4.2 and 4.3 (resp. Tables 4.5 and 4.6, Tables 4.8 and 4.9) display relative frequencies (in %) of rejection (over the 1,000 replications) of the null hypothesis **H0** that the DGP follows two weak PARMA<sub>2</sub>(1, 1) given by (4.33)–(4.34) (resp. (4.33)–(4.35)) with the parameter  $\alpha_0 = (0, 0, 0, 0)'$  (resp.  $\alpha_0 = (0.8, 0.9, 0, 0)'$  and  $\alpha_0 = (0.8, 0.9, -0.5, -0.45)'$ ). Note also that the empirical size is not available (n.a.) for the standard BP and LBM tests because these tests are not applicable to  $m = 1$  (see Tables 4.5 and 4.6) and  $m \leq 2$  (see Tables 4.8 and 4.9). As expected, Tables 4.2, 4.3, 4.5, 4.6, 4.8 and 4.9 show that the standard tests poorly perform to assess the adequacy of these weak PARMA models. The (global or specified season  $\nu$ ) observed relative rejection frequencies of the standard tests are definitely outside the significant limits. Thus the standard tests reject very often the true weak periodic noises (see Tables 4.2 and 4.3) or the true weak PARMA<sub>2</sub>(1, 1) models (see Tables 4.5 and 4.6, Tables 4.8 and 4.9). By contrast, the error of first kind is globally well controlled by the proposed tests when  $N$  increases. We draw the conclusion that for these particular weak PARMA models the proposed tests are clearly preferable to the standard ones.

#### 4.4.2 Empirical power

We now repeat the same experiments for  $N = 1,000$  to examine the empirical power of the standard and proposed tests: first for the null hypothesis of a periodic noise against a PARMA<sub>2</sub>(1, 1) with  $\alpha_0 = (0.8, 0.9, 0.0, 0.0)'$  alternative given by (4.33). Second for the null hypothesis of a PAR<sub>2</sub>(1) against a PARMA<sub>2</sub>(1, 1) with  $\alpha_0 = (0.8, 0.9, 0.5, 0.45)'$  alternative given by (4.33).

Tables 4.10 and 4.11 compare the empirical powers of Model (4.33) for the three different periodic noises above introduced over the  $N \times 2$  independent replications at the asymptotic level  $\alpha = 5\%$ . Model I with innovation  $\epsilon_t \sim \mathcal{N}(0, 1)$ ; Model II with innovation  $\epsilon_t \sim (4.34)$  and Model III with innovation  $\epsilon_t \sim (4.35)$ .

For these particular PARMA<sub>2</sub>(1, 1) models, we notice that the standard and our proposed tests have very similar powers except for BP<sub>SN</sub>( $\nu$ ) (resp. BP<sub>SN</sub>) and LBM<sub>SN</sub>( $\nu$ ) (resp. LBM<sub>SN</sub>) in the case of Model III.



Table 4.1 – Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a strong PARMA<sub>2</sub>(1, 1) given by (4.33) with the parameter  $\alpha_0 = (0, 0, 0, 0)'$  and where the innovation  $\epsilon_t \sim \mathcal{N}(0, 1)$ .

Length N	Tests	Lag $m$											
		1	2	3	4	5	6	7	8	9	10	11	12
N = 400	BP <sub>S</sub> (1)	5.0	4.5	5.2	4.7	5.2	5.2	5.5	5.7	5.7	5.8	5.9	6.0
	BP <sub>S</sub> (2)	4.9	6.5	5.5	5.5	5.9	5.8	6.1	5.9	6.1	6.2	6.0	6.4
	LBM <sub>S</sub> (1)	5.0	4.5	5.3	4.7	5.2	5.2	5.7	5.8	5.8	6.2	6.1	6.2
	LBM <sub>S</sub> (2)	4.9	6.7	5.7	5.7	6.0	6.0	6.1	6.0	6.2	6.4	6.1	6.7
	BP <sub>W</sub> (1)	4.3	3.6	3.6	3.5	3.6	4.0	4.2	4.0	3.8	4.4	4.4	4.1
	BP <sub>W</sub> (2)	4.8	5.5	5.1	5.4	5.0	5.4	5.1	4.7	4.8	5.1	4.5	4.2
	LBM <sub>W</sub> (1)	4.3	3.6	3.8	3.6	3.6	4.0	4.3	4.1	3.9	4.4	4.7	4.7
	LBM <sub>W</sub> (2)	4.8	5.6	5.2	5.5	5.2	5.6	5.2	5.0	5.2	5.4	4.8	4.6
	BP <sub>SN</sub> (1)	4.8	4.5	3.9	4.4	4.6	5.1	4.9	5.5	4.9	5.0	5.8	5.7
	BP <sub>SN</sub> (2)	6.4	5.6	5.3	4.4	5.4	5.3	6.4	5.4	5.3	4.8	5.0	4.9
	LBM <sub>SN</sub> (1)	4.8	4.5	4.0	4.4	4.7	5.1	5.0	5.7	4.9	5.0	5.9	6.2
	LBM <sub>SN</sub> (2)	6.4	5.6	5.3	4.4	5.4	5.6	6.4	5.7	5.4	4.8	5.0	5.1
	BP <sub>S</sub>	4.9	5.6	5.6	5.6	5.9	5.3	5.6	5.9	5.7	6.4	6.3	6.4
	LBM <sub>S</sub>	4.9	5.6	5.6	5.7	6.2	5.6	5.7	6.4	6.1	7.0	6.7	6.9
	BP <sub>W</sub>	4.3	4.4	4.2	4.4	4.2	3.6	3.7	3.7	3.9	3.7	3.1	3.4
	LBM <sub>W</sub>	4.3	4.4	4.2	4.7	4.3	3.7	3.9	4.0	4.2	4.3	3.4	3.5
	BP <sub>SN</sub>	5.9	4.6	4.9	5.4	5.4	5.4	5.8	5.5	5.4	4.9	5.4	4.8
	LBM <sub>SN</sub>	6.1	4.7	5.1	5.4	5.4	5.8	5.8	5.8	5.5	5.1	5.5	5.2
N = 2,000		1	2	3	4	5	6	7	8	9	10	11	12
	BP <sub>S</sub> (1)	5.8	5.8	5.0	5.3	5.7	5.3	6.0	5.6	5.4	5.6	5.9	6.7
	BP <sub>S</sub> (2)	5.3	5.3	5.5	5.3	5.3	5.7	5.8	6.3	6.5	5.9	5.4	4.3
	LBM <sub>S</sub> (1)	5.8	5.8	5.0	5.3	5.7	5.3	6.0	5.7	5.6	5.6	6.0	6.7
	LBM <sub>S</sub> (2)	5.3	5.3	5.5	5.3	5.4	5.7	5.8	6.3	6.6	6.0	5.6	4.3
	BP <sub>W</sub> (1)	5.5	5.9	4.9	5.5	5.5	4.9	5.6	5.6	5.0	5.4	5.7	5.8
	BP <sub>W</sub> (2)	4.8	5.5	5.9	5.1	5.5	5.3	5.5	5.9	6.2	5.8	4.9	4.1
	LBM <sub>W</sub> (1)	5.5	5.9	4.9	5.6	5.5	4.9	5.6	5.6	5.0	5.4	5.8	5.8
	LBM <sub>W</sub> (2)	4.8	5.5	5.9	5.1	5.5	5.3	5.5	6.0	6.2	5.8	4.9	4.1
	BP <sub>SN</sub> (1)	5.0	5.0	6.0	4.9	4.8	4.7	5.6	5.6	6.4	6.0	5.8	6.7
	BP <sub>SN</sub> (2)	5.1	4.3	4.3	5.5	5.1	4.8	5.2	7.3	6.9	6.3	6.3	5.2
	LBM <sub>SN</sub> (1)	5.0	5.0	6.0	5.0	4.8	4.7	5.6	5.7	6.4	6.0	5.8	6.7
	LBM <sub>SN</sub> (2)	5.1	4.3	4.3	5.5	5.1	4.8	5.2	7.3	6.9	6.3	6.3	5.3
	BP <sub>S</sub>	5.1	5.9	5.0	5.1	5.0	5.6	5.6	4.8	5.0	5.2	5.3	4.8
	LBM <sub>S</sub>	5.1	5.9	5.1	5.1	5.0	5.6	5.7	4.8	5.0	5.3	5.4	4.8
	BP <sub>W</sub>	5.6	6.0	4.8	5.0	5.2	5.3	4.9	4.6	4.3	4.6	4.9	4.3
	LBM <sub>W</sub>	5.6	6.0	4.8	5.1	5.3	5.3	4.9	4.6	4.4	4.6	5.0	4.4
	BP <sub>SN</sub>	5.2	4.5	5.2	4.8	4.6	4.9	4.7	5.5	6.0	6.8	6.9	5.6
LBM <sub>SN</sub>	5.2	4.5	5.2	4.8	4.6	5.0	4.8	5.6	6.0	6.8	6.9	5.7	

Table 4.2 – Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a weak PARMA<sub>2</sub>(1, 1) given by (4.33) with the parameter  $\alpha_0 = (0, 0, 0, 0)'$  and where the innovation  $\epsilon_t \sim (4.34)$ .

Length N	Tests	Lag $m$											
		1	2	3	4	5	6	7	8	9	10	11	12
N = 400	BP <sub>S</sub> (1)	13.2	12.5	10.8	9.5	9.4	9.0	9.0	8.2	7.3	7.5	7.4	7.4
	BP <sub>S</sub> (2)	10.5	8.9	8.0	7.2	7.9	8.1	7.9	7.5	6.9	7.6	7.7	7.9
	LBM <sub>S</sub> (1)	13.2	12.9	10.8	9.6	9.4	9.1	9.1	8.3	7.5	7.8	7.7	8.0
	LBM <sub>S</sub> (2)	10.5	8.9	8.3	7.2	8.1	8.1	8.0	7.7	7.3	7.8	8.3	8.4
	BP <sub>W</sub> (1)	4.4	3.9	3.4	3.1	3.3	3.0	2.5	2.7	2.8	2.6	2.6	3.0
	BP <sub>W</sub> (2)	3.6	4.0	3.8	3.5	3.3	3.2	3.6	3.2	3.0	2.4	2.7	2.9
	LBM <sub>W</sub> (1)	4.4	4.0	3.5	3.3	3.6	3.0	2.7	2.7	2.9	2.9	2.8	3.0
	LBM <sub>W</sub> (2)	3.6	4.0	3.8	3.7	3.3	3.4	3.7	3.3	3.0	2.7	2.8	3.0
	BP <sub>SN</sub> (1)	4.8	4.3	4.4	5.2	4.6	4.4	5.2	4.5	3.9	3.7	3.6	3.3
	BP <sub>SN</sub> (2)	3.9	4.1	3.8	4.0	3.5	4.3	4.7	4.8	3.9	3.9	4.0	4.6
	LBM <sub>SN</sub> (1)	4.8	4.3	4.4	5.2	4.7	4.4	5.4	4.5	4.1	4.0	3.7	3.3
	LBM <sub>SN</sub> (2)	3.9	4.1	3.8	4.0	3.6	4.5	4.8	5.1	3.9	4.1	4.0	4.7
	BP <sub>S</sub>	14.3	12.6	11.6	10.2	10.4	10.3	10.5	10.3	10.0	9.8	8.7	9.1
	LBM <sub>S</sub>	14.3	12.7	11.7	10.7	10.6	10.5	10.6	10.5	10.3	10.3	9.1	9.4
	BP <sub>W</sub>	3.2	2.7	2.3	2.4	3.4	2.7	1.7	2.0	2.5	1.5	2.2	1.5
	LBM <sub>W</sub>	3.2	2.8	2.3	2.4	3.4	2.8	2.0	2.2	2.5	2.1	2.5	1.7
	BP <sub>SN</sub>	4.6	5.5	4.3	3.4	3.5	3.5	3.1	3.3	2.8	2.0	1.4	1.2
	LBM <sub>SN</sub>	4.7	5.5	4.4	3.5	3.6	3.7	3.6	3.4	3.2	2.2	1.4	1.3
N = 2,000		1	2	3	4	5	6	7	8	9	10	11	12
	BP <sub>S</sub> (1)	16.6	14.8	12.9	11.7	12.1	11.3	11.8	11.9	10.9	10.6	10.2	10.3
	BP <sub>S</sub> (2)	12.8	10.5	9.9	9.9	8.8	8.5	8.1	7.6	7.1	6.9	6.3	6.6
	LBM <sub>S</sub> (1)	16.6	14.8	12.9	11.8	12.2	11.4	11.8	11.9	11.0	10.6	10.3	10.4
	LBM <sub>S</sub> (2)	12.8	10.5	9.9	9.9	8.9	8.6	8.1	7.6	7.1	6.9	6.3	6.7
	BP <sub>W</sub> (1)	5.4	4.0	4.1	4.5	4.9	5.0	4.5	4.6	4.5	4.8	4.7	5.0
	BP <sub>W</sub> (2)	5.5	5.3	4.6	4.5	4.0	4.4	4.1	3.6	3.7	3.3	3.2	3.4
	LBM <sub>W</sub> (1)	5.4	4.0	4.1	4.5	4.9	5.0	4.5	4.6	4.5	4.8	4.8	5.1
	LBM <sub>W</sub> (2)	5.5	5.3	4.6	4.5	4.0	4.4	4.2	3.6	3.7	3.4	3.2	3.4
	BP <sub>SN</sub> (1)	5.7	4.3	4.7	4.6	5.1	5.0	5.4	5.6	5.4	5.5	5.2	5.4
	BP <sub>SN</sub> (2)	6.0	5.5	5.0	4.6	5.1	6.0	5.1	5.3	4.8	4.9	5.2	5.1
	LBM <sub>SN</sub> (1)	5.7	4.3	4.7	4.6	5.2	5.0	5.4	5.6	5.5	5.6	5.3	5.4
	LBM <sub>SN</sub> (2)	6.0	5.5	5.0	4.6	5.1	6.0	5.1	5.4	4.8	4.9	5.2	5.1
	BP <sub>S</sub>	18.3	15.9	13.9	12.5	12.1	11.9	11.2	11.3	10.6	9.9	10.0	10.4
	LBM <sub>S</sub>	18.3	15.9	13.9	12.5	12.1	11.9	11.3	11.3	10.6	10.0	10.1	10.4
	BP <sub>W</sub>	4.9	4.1	3.9	4.2	4.3	3.8	4.0	4.0	3.7	3.7	3.5	3.7
	LBM <sub>W</sub>	4.9	4.1	3.9	4.2	4.3	3.8	4.0	4.0	3.7	3.7	3.6	3.7
	BP <sub>SN</sub>	5.6	4.5	4.7	4.8	5.8	4.7	5.3	5.4	4.5	4.1	4.7	5.2
LBM <sub>SN</sub>	5.6	4.5	4.7	4.8	5.8	4.7	5.4	5.5	4.5	4.1	4.7	5.2	

Table 4.3 – Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a weak PARMA<sub>2</sub>(1, 1) given by (4.33) with the parameter  $\alpha_0 = (0, 0, 0, 0)'$  and where the innovation  $\epsilon_t \sim (4.35)$ .

Length N	Tests	Lag $m$											
		1	2	3	4	5	6	7	8	9	10	11	12
N = 400	BP <sub>S</sub> (1)	12.6	13.8	15.7	17.1	16.9	18.2	18.7	19.2	20.1	19.0	18.7	19.5
	BP <sub>S</sub> (2)	9.6	12.1	12.7	15.1	14.9	15.1	15.3	15.8	16.5	17.1	16.6	16.3
	LBM <sub>S</sub> (1)	12.6	14.0	15.7	17.3	17.2	18.5	18.9	19.4	20.4	19.6	19.2	20.1
	LBM <sub>S</sub> (2)	9.6	12.2	12.7	15.1	15.4	15.4	15.5	16.4	16.8	17.7	17.2	16.8
	BP <sub>W</sub> (1)	5.1	4.4	3.7	3.3	3.5	3.1	3.3	3.4	2.6	2.9	3.2	3.0
	BP <sub>W</sub> (2)	3.4	3.5	2.7	2.7	2.7	3.1	3.2	3.1	2.5	2.6	2.4	2.4
	LBM <sub>W</sub> (1)	5.1	4.4	3.7	3.3	3.5	3.3	3.4	3.6	3.0	2.9	3.4	3.2
	LBM <sub>W</sub> (2)	3.4	3.5	2.8	2.8	2.7	3.2	3.2	3.2	2.6	2.7	2.7	2.5
	BP <sub>SN</sub> (1)	4.6	3.4	3.8	3.4	3.1	2.4	2.5	2.2	2.0	1.1	1.5	1.4
	BP <sub>SN</sub> (2)	2.8	2.9	2.7	3.5	2.1	2.2	2.6	2.1	1.5	1.4	1.4	1.4
	LBM <sub>SN</sub> (1)	4.6	3.4	3.9	3.4	3.1	2.5	2.6	2.2	2.0	1.1	1.5	1.4
	LBM <sub>SN</sub> (2)	2.8	2.9	2.8	3.6	2.1	2.2	2.6	2.1	1.6	1.5	1.4	1.6
	BP <sub>S</sub>	13.6	15.9	20.0	20.9	23.6	23.9	24.7	25.9	26.8	26.4	25.9	25.6
	LBM <sub>S</sub>	13.6	16.0	20.4	21.2	23.8	24.3	25.0	26.3	27.5	27.2	26.7	26.7
	BP <sub>W</sub>	3.4	2.8	2.0	2.5	2.5	2.4	2.0	2.5	2.3	2.2	2.3	1.8
	LBM <sub>W</sub>	3.4	2.8	2.2	2.5	2.5	2.7	2.1	2.5	2.3	2.3	2.3	2.0
	BP <sub>SN</sub>	3.3	3.2	2.8	1.4	0.9	1.2	0.7	0.8	0.6	0.6	0.6	0.3
	LBM <sub>SN</sub>	3.3	3.2	2.8	1.4	0.9	1.2	0.7	0.8	0.6	0.6	0.6	0.3
N = 2,000		1	2	3	4	5	6	7	8	9	10	11	12
	BP <sub>S</sub> (1)	15.5	17.7	19.3	21.2	23.7	24.4	25.3	26.5	27.3	26.6	27.9	28.4
	BP <sub>S</sub> (2)	13.1	15.6	18.5	19.9	20.4	21.9	22.9	22.0	23.3	22.2	23.9	23.8
	LBM <sub>S</sub> (1)	15.5	17.7	19.3	21.3	23.7	24.5	25.3	26.5	27.3	26.7	28.4	28.6
	LBM <sub>S</sub> (2)	13.1	15.7	18.5	20.0	20.4	21.9	22.9	22.0	23.4	22.6	24.1	23.9
	BP <sub>W</sub> (1)	5.1	5.2	4.5	5.0	4.4	4.0	4.6	4.5	4.7	5.2	4.9	4.0
	BP <sub>W</sub> (2)	5.1	5.3	4.4	3.7	3.7	4.4	4.1	3.9	3.3	3.7	3.5	3.3
	LBM <sub>W</sub> (1)	5.1	5.2	4.5	5.0	4.4	4.0	4.6	4.5	4.7	5.3	5.1	4.0
	LBM <sub>W</sub> (2)	5.1	5.3	4.4	3.8	3.8	4.4	4.1	3.9	3.4	3.8	3.5	3.4
	BP <sub>SN</sub> (1)	5.4	5.1	4.5	3.4	4.8	4.7	4.3	4.7	4.1	4.0	3.7	3.2
	BP <sub>SN</sub> (2)	5.8	5.1	4.5	3.6	4.3	3.7	4.5	4.1	3.9	3.8	3.9	3.2
	LBM <sub>SN</sub> (1)	5.4	5.1	4.5	3.5	4.8	4.7	4.3	4.7	4.1	4.0	3.7	3.2
	LBM <sub>SN</sub> (2)	5.8	5.1	4.5	3.6	4.3	3.7	4.5	4.1	3.9	3.9	4.0	3.3
	BP <sub>S</sub>	19.1	21.9	26.1	29.0	30.1	30.7	31.3	34.1	34.7	34.6	36.4	36.5
	LBM <sub>S</sub>	19.1	22.0	26.1	29.0	30.3	30.7	31.6	34.1	35.0	34.8	36.5	36.8
	BP <sub>W</sub>	5.3	5.0	4.2	3.8	3.9	3.9	4.1	4.1	4.0	3.4	3.4	3.9
	LBM <sub>W</sub>	5.3	5.0	4.2	3.8	3.9	3.9	4.2	4.2	4.0	3.4	3.4	4.1
	BP <sub>SN</sub>	6.2	5.0	4.4	4.1	3.7	4.0	3.9	3.8	3.0	1.7	1.3	1.3
LBM <sub>SN</sub>	6.2	5.0	4.4	4.1	3.7	4.0	3.9	3.8	3.1	1.7	1.4	1.3	

Table 4.4 – Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a strong PARMA<sub>2</sub>(1, 1) given by (4.33) with the parameter  $\alpha_0 = (0.8, 0.9, 0, 0)'$  and where the innovation  $\epsilon_t \sim \mathcal{N}(0, 1)$ .

Length N	Tests	Lag $m$											
		1	2	3	4	5	6	7	8	9	10	11	12
N = 400	BP <sub>S</sub> (1)	n.a.	11.8	9.1	8.5	7.6	7.2	7.5	6.5	6.4	6.8	6.3	6.4
	BP <sub>S</sub> (2)	n.a.	8.4	7.0	6.7	6.7	7.0	7.2	7.4	6.6	7.4	6.8	6.9
	LBM <sub>S</sub> (1)	n.a.	11.9	9.1	8.8	7.6	7.4	7.7	6.6	6.9	7.0	7.0	6.6
	LBM <sub>S</sub> (2)	n.a.	8.5	7.3	6.9	6.8	7.1	7.2	7.5	6.9	7.6	7.2	7.3
	BP <sub>W</sub> (1)	5.5	3.9	3.8	3.1	2.8	2.0	2.5	2.2	1.7	1.9	2.0	1.5
	BP <sub>W</sub> (2)	4.2	4.2	4.4	2.8	2.9	2.3	2.5	2.1	1.5	1.7	1.0	0.9
	LBM <sub>W</sub> (1)	5.5	3.9	3.8	3.1	2.8	2.1	2.6	2.3	1.8	2.0	2.1	1.8
	LBM <sub>W</sub> (2)	4.2	4.5	4.4	3.1	3.1	2.4	2.8	2.3	1.6	2.2	1.4	1.0
	BP <sub>SN</sub> (1)	5.4	5.3	5.8	4.9	4.6	5.7	6.0	5.5	5.3	6.1	5.1	5.6
	BP <sub>SN</sub> (2)	4.8	6.0	5.9	5.8	5.6	5.6	5.2	6.1	5.1	5.4	5.9	6.3
	LBM <sub>SN</sub> (1)	5.4	5.4	5.8	4.9	4.6	5.9	6.3	5.6	5.5	6.2	5.3	5.7
	LBM <sub>SN</sub> (2)	4.8	6.0	6.3	5.9	5.6	5.7	5.3	6.1	5.3	5.6	6.2	6.5
	BP <sub>S</sub>	n.a.	12.3	9.8	8.8	9.0	8.9	8.5	8.1	8.1	8.5	7.8	8.1
	LBM <sub>S</sub>	n.a.	12.4	9.9	9.0	9.1	8.9	8.8	8.6	8.3	8.9	8.6	8.5
	BP <sub>W</sub>	4.4	4.0	3.3	3.4	2.6	1.7	1.4	1.4	1.1	1.1	0.9	0.5
	LBM <sub>W</sub>	4.4	4.1	3.3	3.5	2.6	2.0	1.5	1.4	1.2	1.2	1.0	0.5
	BP <sub>SN</sub>	5.0	5.8	6.4	5.6	5.9	5.9	5.3	5.5	6.0	5.9	4.9	5.0
	LBM <sub>SN</sub>	5.0	5.8	6.4	5.7	6.0	6.3	5.7	5.6	6.1	6.2	5.2	5.3
N = 2,000		1	2	3	4	5	6	7	8	9	10	11	12
	BP <sub>S</sub> (1)	n.a.	13.1	9.6	9.7	8.3	7.1	7.5	7.2	7.2	7.7	8.2	8.3
	BP <sub>S</sub> (2)	n.a.	9.5	7.3	6.9	6.9	7.0	6.3	6.4	6.4	5.3	4.6	4.5
	LBM <sub>S</sub> (1)	n.a.	13.1	9.6	9.7	8.3	7.1	7.7	7.2	7.2	7.7	8.2	8.3
	LBM <sub>S</sub> (2)	n.a.	9.5	7.3	6.9	7.0	7.0	6.4	6.4	6.5	5.4	4.8	4.6
	BP <sub>W</sub> (1)	6.5	5.3	4.8	5.0	4.9	4.3	3.9	4.1	4.3	4.6	5.4	5.7
	BP <sub>W</sub> (2)	5.1	5.6	5.8	4.8	5.4	5.2	5.1	4.7	4.3	3.0	3.0	2.9
	LBM <sub>W</sub> (1)	6.5	5.3	4.8	5.0	4.9	4.3	3.9	4.1	4.3	4.6	5.5	5.7
	LBM <sub>W</sub> (2)	5.1	5.6	5.8	4.8	5.4	5.2	5.1	4.8	4.4	3.1	3.0	3.0
	BP <sub>SN</sub> (1)	6.1	5.1	5.9	5.6	6.1	6.4	6.1	6.7	6.5	7.0	6.2	6.3
	BP <sub>SN</sub> (2)	5.2	5.5	4.2	5.3	5.7	6.3	6.0	6.0	6.1	6.2	6.4	6.1
	LBM <sub>SN</sub> (1)	6.1	5.1	5.9	5.6	6.1	6.4	6.1	6.7	6.5	7.0	6.2	6.4
	LBM <sub>SN</sub> (2)	5.2	5.5	4.2	5.3	5.7	6.3	6.1	6.0	6.2	6.4	6.4	6.1
	BP <sub>S</sub>	n.a.	15.3	10.5	9.8	8.7	8.7	7.6	6.8	7.5	7.2	7.5	8.2
	LBM <sub>S</sub>	n.a.	15.3	10.5	9.8	8.9	8.7	7.7	6.9	7.5	7.3	7.5	8.3
	BP <sub>W</sub>	5.9	5.5	4.9	4.8	4.4	4.0	4.2	4.1	3.5	4.0	3.3	4.0
	LBM <sub>W</sub>	5.9	5.5	4.9	4.8	4.4	4.0	4.2	4.1	3.6	4.2	3.4	4.0
	BP <sub>SN</sub>	6.6	5.6	4.5	4.7	5.0	6.0	6.0	6.2	5.3	5.9	6.3	6.5
LBM <sub>SN</sub>	6.6	5.6	4.5	4.7	5.0	6.0	6.0	6.2	5.3	6.2	6.3	6.5	

Table 4.5 – Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a weak PARMA<sub>2</sub>(1, 1) given by (4.33) with the parameter  $\alpha_0 = (0.8, 0.9, 0, 0)'$  and where the innovation  $\epsilon_t \sim (4.34)$ .

Length N	Tests	Lag $m$											
		1	2	3	4	5	6	7	8	9	10	11	12
N = 400	BP <sub>S</sub> (1)	n.a.	18.5	14.5	12.6	12.2	10.8	10.0	9.9	10.8	9.5	8.6	8.5
	BP <sub>S</sub> (2)	n.a.	15.9	13.5	9.6	9.8	9.2	9.4	9.3	8.6	8.2	8.9	8.6
	LBM <sub>S</sub> (1)	n.a.	18.6	14.5	13.0	12.3	11.0	10.2	10.0	11.1	9.7	8.9	9.1
	LBM <sub>S</sub> (2)	n.a.	16.1	13.5	9.7	10.0	9.3	9.4	9.6	8.8	9.2	9.0	8.8
	BP <sub>W</sub> (1)	4.3	3.1	2.5	2.9	2.7	2.1	1.8	1.9	1.4	1.2	1.1	1.0
	BP <sub>W</sub> (2)	3.6	3.5	2.8	2.6	1.9	1.9	2.0	1.3	1.1	1.1	0.8	0.7
	LBM <sub>W</sub> (1)	4.3	3.1	2.5	2.9	2.7	2.2	1.8	1.9	1.4	1.2	1.3	1.0
	LBM <sub>W</sub> (2)	3.6	3.6	2.8	2.7	2.1	1.9	2.2	1.3	1.2	1.2	1.0	1.1
	BP <sub>SN</sub> (1)	4.4	3.7	3.8	4.4	3.9	4.2	5.0	4.5	3.9	4.3	3.7	3.5
	BP <sub>SN</sub> (2)	3.5	4.8	6.4	5.4	4.5	5.8	4.7	5.5	5.0	4.0	3.7	4.8
	LBM <sub>SN</sub> (1)	4.4	3.7	3.8	4.5	4.0	4.2	5.2	4.5	3.9	4.7	3.8	3.6
	LBM <sub>SN</sub> (2)	3.5	4.8	6.4	5.4	4.6	5.8	4.7	5.7	5.0	4.4	4.1	5.0
	BP <sub>S</sub>	n.a.	24.4	19.1	16.0	14.6	13.6	13.3	13.6	12.4	12.5	11.5	11.1
	LBM <sub>S</sub>	n.a.	24.4	19.3	16.1	15.1	14.1	13.3	13.9	13.0	12.9	12.1	11.8
	BP <sub>W</sub>	2.7	2.0	2.0	2.2	1.6	1.1	0.8	0.9	0.7	0.5	0.4	0.4
	LBM <sub>W</sub>	2.7	2.0	2.0	2.2	1.7	1.1	0.8	0.9	0.8	0.5	0.4	0.4
	BP <sub>SN</sub>	4.0	4.4	5.0	4.7	5.3	4.4	3.7	3.9	3.2	2.6	1.4	1.8
	LBM <sub>SN</sub>	4.0	4.5	5.0	4.8	5.5	4.4	4.0	4.0	3.4	2.7	1.6	2.1
N = 2,000		1	2	3	4	5	6	7	8	9	10	11	12
	BP <sub>S</sub> (1)	n.a.	23.6	17.3	15.7	15.3	14.4	14.7	14.0	12.7	11.9	11.6	11.3
	BP <sub>S</sub> (2)	n.a.	18.4	14.2	12.2	11.3	10.2	8.8	8.9	8.5	7.5	7.5	7.3
	LBM <sub>S</sub> (1)	n.a.	23.6	17.4	15.7	15.3	14.4	14.7	14.1	13.0	12.1	11.6	11.3
	LBM <sub>S</sub> (2)	n.a.	18.4	14.2	12.2	11.3	10.2	8.8	9.0	8.5	7.5	7.5	7.3
	BP <sub>W</sub> (1)	6.2	5.1	4.2	3.9	4.0	4.6	4.2	3.6	3.7	4.1	3.8	4.0
	BP <sub>W</sub> (2)	5.1	4.9	4.6	4.7	3.5	3.2	3.5	3.5	3.5	2.9	2.6	3.0
	LBM <sub>W</sub> (1)	6.2	5.1	4.2	3.9	4.0	4.6	4.3	3.6	3.8	4.1	3.8	4.1
	LBM <sub>W</sub> (2)	5.1	4.9	4.6	4.7	3.5	3.3	3.5	3.5	3.7	3.0	2.6	3.2
	BP <sub>SN</sub> (1)	4.9	4.8	5.3	5.4	5.8	5.5	5.9	5.7	5.4	5.4	5.5	5.3
	BP <sub>SN</sub> (2)	4.8	5.2	4.8	4.6	5.1	6.2	5.0	5.4	5.7	5.7	6.0	5.3
	LBM <sub>SN</sub> (1)	4.9	4.8	5.3	5.4	5.8	5.6	6.0	5.7	5.5	5.5	5.5	5.4
	LBM <sub>SN</sub> (2)	4.8	5.2	4.8	4.6	5.1	6.2	5.1	5.4	5.7	5.8	6.0	5.4
	BP <sub>S</sub>	n.a.	30.8	23.4	20.1	17.3	16.3	15.5	14.1	12.5	12.3	11.7	12.1
	LBM <sub>S</sub>	n.a.	30.8	23.5	20.2	17.4	16.3	15.6	14.1	12.5	12.4	11.8	12.3
	BP <sub>W</sub>	4.7	4.3	3.6	3.7	3.6	2.9	3.1	2.6	2.8	2.9	3.0	3.4
	LBM <sub>W</sub>	4.7	4.3	3.6	3.8	3.6	2.9	3.1	2.6	2.8	2.9	3.0	3.4
	BP <sub>SN</sub>	5.6	6.2	5.7	5.2	5.6	5.5	5.2	6.1	5.6	5.1	4.9	5.8
LBM <sub>SN</sub>	5.6	6.2	5.7	5.2	5.6	5.5	5.2	6.2	5.7	5.1	4.9	5.9	

Table 4.6 – Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a weak PARMA<sub>2</sub>(1, 1) given by (4.33) with the parameter  $\alpha_0 = (0.8, 0.9, 0, 0)'$  and where the innovation  $\epsilon_t \sim (4.35)$ .

Length N	Tests	Lag $m$											
		1	2	3	4	5	6	7	8	9	10	11	12
N = 400	BP <sub>S</sub> (1)	n.a.	22.2	21.5	21.2	20.4	21.2	21.3	21.0	20.7	20.5	21.4	20.3
	BP <sub>S</sub> (2)	n.a.	20.9	18.3	18.3	19.4	19.5	18.4	19.5	19.0	19.0	19.6	18.5
	LBM <sub>S</sub> (1)	n.a.	22.2	21.6	21.6	20.5	21.6	21.3	21.8	21.4	20.9	21.7	21.3
	LBM <sub>S</sub> (2)	n.a.	20.9	18.3	18.4	19.7	19.8	18.5	19.7	19.3	19.6	20.0	19.2
	BP <sub>W</sub> (1)	5.0	3.1	3.8	2.8	1.9	1.5	1.4	1.3	0.9	0.8	0.6	0.3
	BP <sub>W</sub> (2)	4.1	3.1	2.5	2.1	1.4	0.8	0.8	0.9	0.5	0.4	0.6	0.2
	LBM <sub>W</sub> (1)	5.0	3.2	3.8	3.0	2.1	1.7	1.5	1.4	0.9	1.0	0.8	0.6
	LBM <sub>W</sub> (2)	4.1	3.1	2.5	2.1	1.5	0.8	0.8	1.0	0.7	0.6	0.6	0.2
	BP <sub>SN</sub> (1)	4.2	3.7	3.6	3.3	3.4	2.3	2.7	2.7	1.9	1.5	1.7	1.7
	BP <sub>SN</sub> (2)	3.8	3.3	3.3	3.7	2.9	2.4	3.2	2.9	1.8	1.7	1.6	1.7
	LBM <sub>SN</sub> (1)	4.2	3.7	3.6	3.4	3.4	2.4	2.7	2.7	1.9	1.6	1.8	1.8
	LBM <sub>SN</sub> (2)	3.8	3.4	3.3	3.7	2.9	2.5	3.3	3.3	1.8	1.7	1.7	1.8
	BP <sub>S</sub>	n.a.	31.4	28.8	28.2	28.4	29.3	28.3	28.5	29.3	28.7	29.4	29.0
	LBM <sub>S</sub>	n.a.	31.5	28.9	28.5	28.7	29.8	28.4	29.4	30.3	29.4	30.4	30.1
	BP <sub>W</sub>	3.9	2.2	2.1	1.7	0.6	0.7	0.1	0.4	0.3	0.1	0.1	0.0
	LBM <sub>W</sub>	3.9	2.2	2.2	1.9	0.7	0.7	0.1	0.5	0.3	0.1	0.1	0.0
	BP <sub>SN</sub>	3.1	3.9	2.9	2.4	2.1	1.5	1.0	1.2	0.6	0.5	0.5	0.5
	LBM <sub>SN</sub>	3.1	3.9	2.9	2.4	2.2	1.6	1.0	1.2	0.6	0.6	0.6	0.5
N = 2,000		1	2	3	4	5	6	7	8	9	10	11	12
	BP <sub>S</sub> (1)	n.a.	24.4	23.5	25.5	26.8	26.8	27.5	28.2	29.2	28.6	29.6	30.1
	BP <sub>S</sub> (2)	n.a.	24.1	24.3	24.7	22.8	23.5	24.3	25.2	25.3	25.3	25.6	25.5
	LBM <sub>S</sub> (1)	n.a.	24.6	23.5	25.5	26.9	27.0	27.5	28.2	29.6	28.6	29.7	30.2
	LBM <sub>S</sub> (2)	n.a.	24.1	24.3	24.8	22.8	23.6	24.4	25.3	25.3	25.4	25.7	25.5
	BP <sub>W</sub> (1)	4.6	5.3	4.7	3.7	3.7	2.9	3.1	3.2	2.7	3.6	3.3	2.8
	BP <sub>W</sub> (2)	5.4	4.9	4.5	4.0	3.3	3.5	3.0	3.2	2.8	2.3	1.8	1.7
	LBM <sub>W</sub> (1)	4.6	5.3	4.7	3.7	3.7	2.9	3.1	3.2	2.7	3.7	3.3	2.8
	LBM <sub>W</sub> (2)	5.4	4.9	4.5	4.0	3.3	3.6	3.0	3.2	2.8	2.4	1.8	1.7
	BP <sub>SN</sub> (1)	5.0	4.8	4.2	4.6	5.0	4.6	4.5	5.1	4.5	3.8	3.9	3.7
	BP <sub>SN</sub> (2)	4.5	4.2	4.5	4.3	4.4	3.9	3.8	4.1	3.6	4.0	3.8	3.0
	LBM <sub>SN</sub> (1)	5.0	4.8	4.2	4.6	5.0	4.6	4.5	5.1	4.5	3.8	3.9	3.7
	LBM <sub>SN</sub> (2)	4.5	4.3	4.5	4.3	4.4	3.9	3.9	4.1	3.7	4.1	3.9	3.1
	BP <sub>S</sub>	n.a.	38.2	36.2	37.1	37.3	36.7	36.5	38.7	39.9	39.2	39.5	38.5
	LBM <sub>S</sub>	n.a.	38.2	36.4	37.2	37.5	37.0	36.6	38.8	40.0	39.3	39.7	38.6
	BP <sub>W</sub>	5.3	4.7	3.1	2.7	2.8	2.8	2.2	2.3	2.2	1.9	1.8	2.0
	LBM <sub>W</sub>	5.3	4.7	3.1	2.7	2.8	2.8	2.2	2.3	2.2	1.9	1.9	2.0
	BP <sub>SN</sub>	5.6	5.3	4.6	5.0	4.2	3.8	3.9	3.4	2.9	2.2	1.7	1.9
LBM <sub>SN</sub>	5.6	5.3	4.6	5.0	4.4	3.8	3.9	3.4	2.9	2.2	1.8	1.9	

Table 4.7 – Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a strong PARMA<sub>2</sub>(1, 1) given by (4.33) with the parameter  $\alpha_0 = (0.8, 0.9, -0.5, -0.45)'$  and where the innovation  $\epsilon_t \sim \mathcal{N}(0, 1)$ .

Length N	Tests	Lag $m$											
		1	2	3	4	5	6	7	8	9	10	11	12
N = 400	BP <sub>S</sub> (1)	n.a.	n.a.	17.2	13.3	10.9	10.2	10.0	9.1	9.8	9.3	9.2	8.6
	BP <sub>S</sub> (2)	n.a.	n.a.	19.9	13.5	12.1	11.3	11.2	10.1	9.8	10.2	10.2	9.4
	LBM <sub>S</sub> (1)	n.a.	n.a.	17.3	13.4	11.1	10.2	10.1	9.4	10.3	9.9	9.9	9.3
	LBM <sub>S</sub> (2)	n.a.	n.a.	20.1	13.7	12.3	11.6	11.5	10.3	10.2	10.3	11.0	9.6
	BP <sub>W</sub> (1)	4.1	3.9	3.5	2.8	2.6	2.7	2.7	2.5	2.1	2.8	1.9	1.9
	BP <sub>W</sub> (2)	5.0	5.4	5.2	3.9	3.5	3.1	3.2	2.8	2.0	2.0	1.9	1.6
	LBM <sub>W</sub> (1)	4.1	3.9	3.7	2.8	2.8	2.7	2.7	2.8	2.5	3.1	2.0	2.0
	LBM <sub>W</sub> (2)	5.0	5.4	5.2	3.9	3.6	3.1	3.4	2.9	2.3	2.0	1.9	1.6
	BP <sub>SN</sub> (1)	5.3	5.3	4.6	3.9	4.3	5.2	5.1	5.0	5.4	5.3	6.0	6.2
	BP <sub>SN</sub> (2)	6.0	5.6	5.3	5.1	4.6	6.6	5.8	5.6	5.0	5.6	5.1	5.5
	LBM <sub>SN</sub> (1)	5.3	5.3	4.6	3.9	4.3	5.2	5.2	5.2	5.6	5.3	6.3	6.5
	LBM <sub>SN</sub> (2)	6.0	5.7	5.4	5.1	4.7	6.9	6.2	5.8	5.1	5.9	5.3	5.9
	BP <sub>S</sub>	n.a.	n.a.	26.8	20.3	17.1	16.3	15.8	13.0	12.1	13.3	12.0	11.5
	LBM <sub>S</sub>	n.a.	n.a.	27.2	20.6	17.5	16.8	16.5	13.5	13.0	14.2	12.6	12.8
	BP <sub>W</sub>	4.8	3.7	3.8	2.8	2.3	1.9	1.8	1.1	1.6	1.5	1.1	0.8
	LBM <sub>W</sub>	4.8	3.7	3.9	3.0	2.3	1.9	1.8	1.2	1.8	1.5	1.2	0.9
	BP <sub>SN</sub>	3.9	3.6	4.9	4.9	5.4	5.9	6.1	6.7	5.8	4.8	5.1	4.5
	LBM <sub>SN</sub>	3.9	3.8	4.9	4.9	5.4	6.0	6.2	6.7	6.0	5.1	5.2	5.0
N = 2,000		1	2	3	4	5	6	7	8	9	10	11	12
	BP <sub>S</sub> (1)	n.a.	n.a.	17.9	12.1	10.2	8.2	8.0	8.5	7.6	8.1	7.8	8.1
	BP <sub>S</sub> (2)	n.a.	n.a.	18.3	12.9	11.9	11.2	9.5	10.1	9.2	9.2	8.7	8.2
	LBM <sub>S</sub> (1)	n.a.	n.a.	18.0	12.1	10.2	8.3	8.0	8.6	7.6	8.2	7.8	8.1
	LBM <sub>S</sub> (2)	n.a.	n.a.	18.3	13.0	11.9	11.2	9.6	10.1	9.3	9.2	8.9	8.2
	BP <sub>W</sub> (1)	4.3	4.9	4.5	4.7	4.3	3.9	3.8	4.4	3.6	3.7	4.1	5.1
	BP <sub>W</sub> (2)	5.4	5.9	5.3	4.4	4.7	4.6	5.3	4.9	5.0	4.7	3.9	3.0
	LBM <sub>W</sub> (1)	4.3	4.9	4.5	4.7	4.3	3.9	3.8	4.4	3.6	3.7	4.1	5.1
	LBM <sub>W</sub> (2)	5.4	5.9	5.3	4.4	4.7	4.6	5.3	4.9	5.0	4.9	4.0	3.0
	BP <sub>SN</sub> (1)	4.7	4.8	5.4	4.2	5.2	5.6	5.7	5.5	5.2	5.2	5.6	5.6
	BP <sub>SN</sub> (2)	5.9	5.0	4.1	5.7	5.5	4.9	5.5	6.1	6.5	6.4	5.9	5.3
	LBM <sub>SN</sub> (1)	4.7	4.8	5.4	4.2	5.2	5.6	5.8	5.5	5.2	5.2	5.9	5.7
	LBM <sub>SN</sub> (2)	5.9	5.0	4.1	5.7	5.5	4.9	5.5	6.1	6.5	6.4	6.1	5.3
	BP <sub>S</sub>	n.a.	n.a.	26.5	18.3	15.2	13.1	13.4	11.7	9.4	10.2	9.8	8.7
	LBM <sub>S</sub>	n.a.	n.a.	26.5	18.3	15.2	13.1	13.6	11.7	9.4	10.3	10.0	8.7
	BP <sub>W</sub>	4.6	4.9	4.2	4.3	4.2	4.0	3.6	3.2	3.5	3.6	3.4	3.0
	LBM <sub>W</sub>	4.6	4.9	4.2	4.3	4.2	4.0	3.6	3.3	3.7	3.7	3.4	3.0
	BP <sub>SN</sub>	4.7	5.4	5.4	4.7	4.6	4.7	5.4	5.3	4.9	6.7	6.8	6.5
LBM <sub>SN</sub>	4.7	5.4	5.4	4.7	4.6	4.8	5.4	5.4	5.0	6.7	6.9	6.5	



Table 4.8 – Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a weak  $\text{PARMA}_2(1, 1)$  given by (4.33) with the parameter  $\alpha_0 = (0.8, 0.9, -0.5, -0.45)'$  and where the innovation  $\epsilon_t \sim (4.34)$ .

Length N	Tests	Lag $m$											
		1	2	3	4	5	6	7	8	9	10	11	12
N = 400	$\text{BP}_S(1)$	n.a.	n.a.	24.8	16.9	14.2	11.3	10.1	10.4	10.8	10.4	9.8	9.5
	$\text{BP}_S(2)$	n.a.	n.a.	22.5	17.5	13.6	12.4	12.1	12.1	11.1	10.8	10.4	10.3
	$\text{LBM}_S(1)$	n.a.	n.a.	24.8	17.2	14.2	11.5	10.2	10.6	11.1	10.7	10.2	9.8
	$\text{LBM}_S(2)$	n.a.	n.a.	22.6	17.7	13.8	12.8	12.3	12.3	11.2	11.1	10.9	10.7
	$\text{BP}_W(1)$	4.6	3.4	3.0	2.8	3.1	2.8	2.0	2.3	2.1	1.7	1.3	0.9
	$\text{BP}_W(2)$	3.4	3.2	3.5	3.2	2.6	2.6	2.0	1.2	1.2	0.8	1.0	0.8
	$\text{LBM}_W(1)$	4.6	3.4	3.1	2.8	3.2	2.9	2.1	2.3	2.2	2.0	1.4	1.0
	$\text{LBM}_W(2)$	3.4	3.2	3.6	3.3	2.6	2.8	2.2	1.5	1.3	0.9	1.1	0.9
	$\text{BP}_{SN}(1)$	4.5	4.0	4.4	4.1	4.6	4.4	4.6	4.1	4.7	4.4	3.9	3.8
	$\text{BP}_{SN}(2)$	4.0	4.5	4.4	5.0	4.8	4.5	5.2	5.5	5.4	4.8	4.1	4.4
	$\text{LBM}_{SN}(1)$	4.5	4.0	4.5	4.1	4.7	4.6	4.9	4.3	4.7	4.4	4.1	3.9
	$\text{LBM}_{SN}(2)$	4.0	4.6	4.4	5.1	4.9	4.6	5.2	5.7	5.7	5.0	4.2	4.5
	$\text{BP}_S$	n.a.	n.a.	35.1	24.7	20.2	17.7	16.6	14.9	15.2	15.0	14.2	13.5
	$\text{LBM}_S$	n.a.	n.a.	35.1	24.8	20.3	18.1	16.9	15.4	15.4	15.7	15.0	14.8
	$\text{BP}_W$	3.4	2.7	2.8	2.5	1.4	1.3	1.2	1.1	0.8	0.5	0.7	0.4
	$\text{LBM}_W$	3.4	2.7	2.9	2.5	1.5	1.4	1.2	1.1	0.8	0.7	0.7	0.5
	$\text{BP}_{SN}$	3.4	4.7	4.0	3.8	3.6	3.9	3.3	3.4	2.8	2.5	1.9	1.5
	$\text{LBM}_{SN}$	3.4	4.7	4.0	3.9	3.7	3.9	3.4	3.4	3.0	2.7	2.1	1.8
N = 2,000		1	2	3	4	5	6	7	8	9	10	11	12
	$\text{BP}_S(1)$	n.a.	n.a.	28.2	21.3	18.8	16.5	16.6	14.1	13.4	14.4	12.8	12.6
	$\text{BP}_S(2)$	n.a.	n.a.	25.8	19.3	15.1	13.3	12.6	10.7	10.2	10.0	8.7	8.2
	$\text{LBM}_S(1)$	n.a.	n.a.	28.2	21.3	18.8	16.5	16.6	14.1	13.7	14.4	13.1	12.6
	$\text{LBM}_S(2)$	n.a.	n.a.	25.8	19.3	15.1	13.5	12.6	10.7	10.2	10.0	8.7	8.2
	$\text{BP}_W(1)$	5.0	4.3	4.3	4.6	4.7	4.6	4.4	4.3	4.3	4.1	4.3	4.8
	$\text{BP}_W(2)$	5.0	5.0	4.3	4.4	3.7	3.8	3.5	3.2	3.5	2.8	3.2	3.5
	$\text{LBM}_W(1)$	5.0	4.3	4.3	4.6	4.7	4.6	4.4	4.4	4.3	4.3	4.3	4.8
	$\text{LBM}_W(2)$	5.0	5.1	4.3	4.4	3.8	3.9	3.5	3.2	3.5	2.8	3.2	3.5
	$\text{BP}_{SN}(1)$	5.5	4.9	5.0	4.3	5.5	5.2	5.5	5.3	5.3	4.9	4.1	5.0
	$\text{BP}_{SN}(2)$	5.3	5.1	5.7	5.3	5.3	5.3	5.2	5.7	5.4	5.7	5.8	5.6
	$\text{LBM}_{SN}(1)$	5.5	5.0	5.0	4.4	5.5	5.2	5.6	5.3	5.3	5.0	4.1	5.0
	$\text{LBM}_{SN}(2)$	5.3	5.1	5.7	5.3	5.3	5.3	5.2	5.7	5.5	5.7	5.8	5.6
	$\text{BP}_S$	n.a.	n.a.	40.4	29.1	24.8	22.3	19.7	18.4	16.2	15.1	15.1	14.9
	$\text{LBM}_S$	n.a.	n.a.	40.5	29.1	24.9	22.5	19.8	18.5	16.2	15.1	15.1	15.3
	$\text{BP}_W$	4.8	4.0	3.9	3.9	3.6	2.9	3.1	2.3	2.4	2.9	3.0	3.8
	$\text{LBM}_W$	4.8	4.0	3.9	3.9	3.7	2.9	3.1	2.4	2.4	2.9	3.0	3.8
	$\text{BP}_{SN}$	5.5	4.8	5.0	5.0	5.3	5.1	5.6	5.1	4.6	4.4	4.6	5.0
$\text{LBM}_{SN}$	5.5	4.8	5.0	5.0	5.3	5.1	5.7	5.1	4.6	4.5	4.6	5.0	

Table 4.9 – Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a weak  $\text{PARMA}_2(1, 1)$  given by (4.33) with the parameter  $\alpha_0 = (0.8, 0.9, -0.5, -0.45)'$  and where the innovation  $\epsilon_t \sim (4.35)$ .

Length N	Tests	Lag $m$											
		1	2	3	4	5	6	7	8	9	10	11	12
N = 400	$\text{BP}_S(1)$	n.a.	n.a.	35.7	27.8	25.9	25.4	25.1	23.5	23.6	22.6	22.7	23.1
	$\text{BP}_S(2)$	n.a.	n.a.	31.3	26.8	23.7	23.4	23.0	21.6	21.1	21.0	21.3	20.2
	$\text{LBM}_S(1)$	n.a.	n.a.	35.7	28.2	26.0	25.7	25.4	24.1	23.9	23.0	23.0	23.4
	$\text{LBM}_S(2)$	n.a.	n.a.	31.4	26.8	24.3	23.9	23.3	21.9	21.8	21.5	21.8	21.1
	$\text{BP}_W(1)$	4.0	3.8	3.5	3.1	2.3	2.4	1.5	1.5	1.0	0.9	0.7	0.4
	$\text{BP}_W(2)$	3.3	2.5	2.1	1.5	1.0	1.0	1.1	0.8	0.6	0.6	0.5	0.3
	$\text{LBM}_W(1)$	4.0	3.9	3.6	3.1	2.4	2.5	1.6	1.7	1.0	0.9	0.8	0.6
	$\text{LBM}_W(2)$	3.3	2.5	2.2	1.7	1.2	1.0	1.2	0.9	0.6	0.6	0.6	0.3
	$\text{BP}_{SN}(1)$	4.0	3.4	3.3	3.4	3.4	2.4	2.9	2.0	1.8	1.4	1.7	1.5
	$\text{BP}_{SN}(2)$	3.0	2.9	3.3	3.2	2.4	2.6	3.0	3.0	2.1	1.9	1.7	1.8
	$\text{LBM}_{SN}(1)$	4.0	3.5	3.3	3.5	3.5	2.4	2.9	2.1	1.8	1.4	1.7	1.5
	$\text{LBM}_{SN}(2)$	3.0	2.9	3.4	3.2	2.4	2.6	3.0	3.0	2.1	2.0	1.7	1.9
	$\text{BP}_S$	n.a.	n.a.	48.3	40.4	36.7	36.4	34.5	33.8	34.4	33.6	31.4	31.6
	$\text{LBM}_S$	n.a.	n.a.	48.5	40.8	37	37.1	34.8	34.2	35.3	34.4	32.4	33.1
	$\text{BP}_W$	2.7	2.6	2.2	2.0	1.2	0.8	0.4	0.4	0.3	0.2	0.0	0.1
	$\text{LBM}_W$	2.7	2.6	2.3	2.0	1.3	0.8	0.4	0.5	0.3	0.2	0.1	0.1
	$\text{BP}_{SN}$	2.7	2.8	2.4	1.9	1.4	1.3	0.9	0.8	0.4	0.6	0.5	0.5
	$\text{LBM}_{SN}$	2.8	2.9	2.5	1.9	1.4	1.4	0.9	0.8	0.5	0.6	0.6	0.5
N = 2,000		1	2	3	4	5	6	7	8	9	10	11	12
	$\text{BP}_S(1)$	n.a.	n.a.	36.4	33.5	33.6	32.8	32.9	32.5	33.8	33.8	34.6	34.7
	$\text{BP}_S(2)$	n.a.	n.a.	36.0	31.9	28.1	29.0	28.6	29.2	29.3	28.5	28.8	27.7
	$\text{LBM}_S(1)$	n.a.	n.a.	36.4	33.6	33.6	32.9	32.9	32.5	33.9	33.8	34.6	34.7
	$\text{LBM}_S(2)$	n.a.	n.a.	36.0	31.9	28.1	29.1	28.7	29.2	29.5	28.5	28.8	27.9
	$\text{BP}_W(1)$	5.4	5.4	4.7	3.8	3.3	2.9	3.3	3.0	3.0	3.5	2.8	2.9
	$\text{BP}_W(2)$	6.0	5.1	4.6	3.4	3.2	3.8	3.5	3.5	3.3	3.0	2.7	2.3
	$\text{LBM}_W(1)$	5.4	5.4	4.7	3.9	3.3	2.9	3.3	3.1	3.0	3.5	2.9	2.9
	$\text{LBM}_W(2)$	6.0	5.1	4.6	3.4	3.2	3.8	3.5	3.5	3.3	3.0	2.8	2.3
	$\text{BP}_{SN}(1)$	4.2	4.3	4.4	4.3	5.3	5.2	4.5	4.2	4.2	3.7	3.8	3.1
	$\text{BP}_{SN}(2)$	4.4	4.7	5.2	4.6	3.9	4.1	3.6	3.3	3.4	3.3	3.3	3.2
	$\text{LBM}_{SN}(1)$	4.2	4.3	4.4	4.3	5.3	5.2	4.5	4.2	4.3	3.7	3.9	3.2
	$\text{LBM}_{SN}(2)$	4.4	4.7	5.2	4.6	3.9	4.1	3.6	3.3	3.4	3.3	3.3	3.2
	$\text{BP}_S$	n.a.	n.a.	53.9	49.5	47.3	45.0	44.6	47.1	46.8	47.0	46.9	44.5
	$\text{LBM}_S$	n.a.	n.a.	53.9	49.6	47.4	45.1	44.6	47.2	47.2	47.2	47.2	44.6
	$\text{BP}_W$	5.5	4.6	3.0	2.9	2.6	2.6	2.8	2.4	2.2	1.6	1.9	2.0
	$\text{LBM}_W$	5.5	4.7	3.0	2.9	2.6	2.6	2.8	2.4	2.2	1.6	1.9	2.0
	$\text{BP}_{SN}$	4.5	3.7	4.6	3.7	3.8	3.1	3.0	3.0	3.3	1.7	1.8	1.4
$\text{LBM}_{SN}$	4.5	3.7	4.6	3.7	3.8	3.1	3.0	3.0	3.3	1.7	1.8	1.4	

Table 4.10 – Empirical power (in %) of the modified and standard versions of the LBM and BP tests for the null hypothesis of a periodic noise against a  $\text{PARMA}_2(1, 1)$  given by (4.33) with the parameter  $\alpha_0 = (0.8, 0.9, 0.0, 0.0)'$ . Model I with innovation  $\epsilon_t \sim \mathcal{N}(0, 1)$ ; Model II with innovation  $\epsilon_t \sim (4.34)$  and Model III with innovation  $\epsilon_t \sim (4.35)$ . The nominal asymptotic level of the tests is  $\alpha = 5\%$ .

Model	Tests	Lag $m$											
		1	2	3	4	5	6	7	8	9	10	11	12
I	BP <sub>S</sub> (1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	BP <sub>S</sub> (2)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	LBM <sub>S</sub> (1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	LBM <sub>S</sub> (2)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	BP <sub>W</sub> (1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	BP <sub>W</sub> (2)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	LBM <sub>W</sub> (1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	LBM <sub>W</sub> (2)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	BP <sub>SN</sub> (1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	BP <sub>SN</sub> (2)	99.9	100.0	100.0	99.9	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	LBM <sub>SN</sub> (1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	LBM <sub>SN</sub> (2)	99.9	100.0	100.0	99.9	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	BP <sub>S</sub>	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	LBM <sub>S</sub>	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	BP <sub>W</sub>	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	LBM <sub>W</sub>	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	BP <sub>SN</sub>	99.8	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	LBM <sub>SN</sub>	99.8	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
II	BP <sub>S</sub> (1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	BP <sub>S</sub> (2)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	LBM <sub>S</sub> (1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	LBM <sub>S</sub> (2)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	BP <sub>W</sub> (1)	100.0	100.0	100.0	99.9	99.9	99.8	99.9	99.8	99.8	99.9	99.9	99.9
	BP <sub>W</sub> (2)	100.0	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
	LBM <sub>W</sub> (1)	100.0	100.0	100.0	99.9	99.9	99.8	99.9	99.8	99.8	99.9	99.9	99.9
	LBM <sub>W</sub> (2)	100.0	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
	BP <sub>SN</sub> (1)	99.1	98.8	98.8	99.0	98.9	98.8	98.7	98.5	98.6	98.1	98.3	98.2
	BP <sub>SN</sub> (2)	99.1	98.8	98.9	98.8	98.9	98.8	98.2	98.3	98.4	98.1	97.9	97.7
	LBM <sub>SN</sub> (1)	99.1	98.8	98.7	99.0	98.7	98.7	98.6	98.5	98.5	98.1	98.3	98.2
	LBM <sub>SN</sub> (2)	99.1	98.8	98.9	98.7	98.9	98.8	98.2	98.2	98.4	98.1	97.9	97.7
	BP <sub>S</sub>	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	LBM <sub>S</sub>	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	BP <sub>W</sub>	99.9	99.9	99.9	99.9	99.8	99.7	99.8	99.8	99.8	99.8	99.8	99.8
	LBM <sub>W</sub>	99.9	99.9	99.9	99.9	99.8	99.7	99.8	99.8	99.8	99.8	99.8	99.8
	BP <sub>SN</sub>	97.6	98.4	98.8	98.9	98.6	98.5	98.5	98.3	98.4	98.1	97.5	97.1
	LBM <sub>SN</sub>	97.7	98.4	98.8	98.9	98.5	98.5	98.4	98.3	98.2	98.1	97.3	97.0
III	BP <sub>S</sub> (1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	BP <sub>S</sub> (2)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	LBM <sub>S</sub> (1)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	LBM <sub>S</sub> (2)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	BP <sub>W</sub> (1)	98.9	97.9	97.3	96.8	96.6	96.0	95.8	95.4	95.1	94.7	94.3	93.7
	BP <sub>W</sub> (2)	98.6	97.6	97.1	97.2	96.6	95.9	95.8	95.3	94.8	94.4	93.8	93.7
	LBM <sub>W</sub> (1)	98.9	97.9	97.3	96.8	96.6	96.1	95.8	95.4	95.1	94.7	94.3	93.7
	LBM <sub>W</sub> (2)	98.6	97.7	97.1	97.2	96.6	95.9	95.9	95.3	94.8	94.5	93.8	93.7
	BP <sub>SN</sub> (1)	84.6	80.3	78.1	77.8	76.3	74.4	72.0	69.8	66.5	64.5	63.4	62.0
	BP <sub>SN</sub> (2)	84.7	81.2	77.8	76.4	73.9	72.5	71.3	69.3	67.9	66.9	65.2	63.4
	LBM <sub>SN</sub> (1)	84.6	80.2	78.0	77.6	76.5	73.8	71.4	69.5	66.3	64.2	62.9	61.7
	LBM <sub>SN</sub> (2)	84.7	81.3	77.9	76.1	73.4	72.1	71.2	69.1	67.3	66.7	65.1	62.8
	BP <sub>S</sub>	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	LBM <sub>S</sub>	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	BP <sub>W</sub>	98.2	96.8	96.4	96.5	95.9	95.2	94.8	94.9	94.2	93.5	92.7	92.4
	LBM <sub>W</sub>	98.2	96.8	96.4	96.5	95.9	95.3	94.8	94.9	94.3	93.5	92.8	92.4
	BP <sub>SN</sub>	76.4	76.0	73.5	71.2	68.7	64.4	60.3	56.7	51.1	46.8	41.2	37.4
	LBM <sub>SN</sub>	76.6	76.0	73.8	71.1	68.8	63.7	59.9	56.2	50.1	46.0	40.6	37.1

Table 4.11 – Empirical power (in %) of the modified and standard versions of the LBM and BP tests for the null hypothesis of a  $\text{PARMA}_2(1, 0)$  against a  $\text{PARMA}_2(1, 1)$  given by (4.33) with the parameter  $\alpha_0 = (0.8, 0.9, 0.5, 0.45)'$ . Model I with innovation  $\epsilon_t \sim \mathcal{N}(0, 1)$ ; Model II with innovation  $\epsilon_t \sim (4.34)$  and Model III with innovation  $\epsilon_t \sim (4.35)$ . The nominal asymptotic level of the tests is  $\alpha = 5\%$ .

Model	Tests	Lag $m$															
		1	2	3	4	5	6	7	8	9	10	11	12				
I	$\text{BP}_S(1)$	n.a.	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$\text{BP}_S(2)$	n.a.	90.1	87.4	84.8	82.4	80.1	77.4	74.9	73.3	69.6	67.7	66.1				
	$\text{LBM}_S(1)$	n.a.	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$\text{LBM}_S(2)$	n.a.	90.1	87.4	84.8	82.5	80.2	77.4	75.0	73.5	70.0	67.9	66.3				
	$\text{BP}_W(1)$	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$\text{BP}_W(2)$	72.8	79.5	79.5	77.0	73.7	71.2	67.1	64.3	60.2	55.6	52.7	50.8				
	$\text{LBM}_W(1)$	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$\text{LBM}_W(2)$	72.8	79.5	79.5	77.1	73.7	71.4	67.2	64.4	60.4	55.8	53.1	51.2				
	$\text{BP}_{SN}(1)$	97.9	94.4	93.5	90.9	90.6	90.0	89.5	88.7	87.8	88.4	87.9	88.0				
	$\text{BP}_{SN}(2)$	54.1	54.2	51.1	50.0	45.3	43.8	42.4	41.5	39.0	37.1	37.3	35.3				
	$\text{LBM}_{SN}(1)$	97.9	94.4	93.5	90.9	90.7	90.1	89.6	88.8	87.9	88.2	87.9	88.1				
	$\text{LBM}_{SN}(2)$	54.1	54.2	51.1	50.0	45.3	43.8	42.4	41.5	39.4	37.2	37.4	35.6				
	$\text{BP}_S$	n.a.	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$\text{LBM}_S$	n.a.	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$\text{BP}_W$	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$\text{LBM}_W$	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$\text{BP}_{SN}$	96.0	95.6	94.9	94.8	94.1	94.6	94.4	93.7	92.4	91.7	90.6	90.1				
	$\text{LBM}_{SN}$	96.0	95.6	94.9	94.8	94.1	94.6	94.6	93.7	92.5	91.7	90.6	90.2				
II	$\text{BP}_S(1)$	n.a.	100.0	100.0	100.0	99.9	100.0	100.0	100.0	100.0	99.9	99.9	99.9				
	$\text{BP}_S(2)$	n.a.	94.4	93.9	92.7	91.8	89.7	89.7	86.8	84.9	84.0	82.8	81.9				
	$\text{LBM}_S(1)$	n.a.	100.0	100.0	100.0	99.9	100.0	100.0	100.0	100.0	99.9	99.9	99.9				
	$\text{LBM}_S(2)$	n.a.	94.4	93.9	92.7	91.8	89.7	89.7	86.8	85.0	84.0	83.0	82.0				
	$\text{BP}_W(1)$	97.7	99.2	99.1	99.0	99.0	99.0	98.9	98.8	98.7	98.5	98.5	98.2				
	$\text{BP}_W(2)$	68.9	80.1	81.2	79.8	78.1	76.2	73.1	70.6	68.9	66.2	64.3	61.5				
	$\text{LBM}_W(1)$	97.7	99.2	99.1	99.0	99.0	99.0	98.9	98.8	98.7	98.5	98.5	98.2				
	$\text{LBM}_W(2)$	68.9	80.1	81.2	79.9	78.1	76.5	73.2	70.7	69.3	66.7	64.6	61.8				
	$\text{BP}_{SN}(1)$	83.8	85.5	82.5	80.6	78.8	76.9	75.6	74.0	71.3	69.3	68.6	65.2				
	$\text{BP}_{SN}(2)$	48.6	48.5	49.3	46.3	44.8	41.9	39.5	35.9	34.3	32.5	30.5	28.9				
	$\text{LBM}_{SN}(1)$	83.8	85.5	82.5	80.6	78.8	76.9	75.5	74.1	71.5	69.5	68.8	65.3				
	$\text{LBM}_{SN}(2)$	48.6	48.5	49.3	46.4	45.1	41.9	39.6	36.1	34.5	32.7	30.7	29.1				
	$\text{BP}_S$	n.a.	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$\text{LBM}_S$	n.a.	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$\text{BP}_W$	97.9	99.2	99.2	99.2	99.2	99.1	99.2	99.0	98.8	99.0	98.9	98.9				
	$\text{LBM}_W$	97.9	99.2	99.2	99.2	99.2	99.1	99.2	99.0	99.0	99.0	98.9	98.9				
	$\text{BP}_{SN}$	82.1	87.2	87.3	85.4	84.7	82.9	81.8	80.5	77.9	75.0	72.8	70.6				
	$\text{LBM}_{SN}$	82.2	87.2	87.5	85.4	84.7	83.0	82.0	80.5	78.0	75.0	73.1	70.7				
III	$\text{BP}_S(1)$	n.a.	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$\text{BP}_S(2)$	n.a.	96.3	96.7	96.3	96.1	95.2	94.6	93.3	92.7	92.1	91.7	90.5				
	$\text{LBM}_S(1)$	n.a.	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$\text{LBM}_S(2)$	n.a.	96.3	96.7	96.3	96.1	95.2	94.6	93.3	92.7	92.1	91.7	90.5				
	$\text{BP}_W(1)$	96.2	97.8	97.7	96.8	96.2	95.5	94.9	94.0	92.6	91.3	90.6	89.9				
	$\text{BP}_W(2)$	72.8	78.6	79.3	75.7	71.7	66.9	64.3	60.8	57.2	54.8	51.9	48.6				
	$\text{LBM}_W(1)$	96.2	97.8	97.8	96.8	96.2	95.5	95.0	94.0	92.9	91.5	90.6	90.0				
	$\text{LBM}_W(2)$	72.8	78.8	79.3	75.7	71.9	67.0	64.4	61.1	57.2	54.9	52.0	48.9				
	$\text{BP}_{SN}(1)$	82.7	77.4	75.5	71.5	69.1	65.1	64.2	60.7	58.7	56.4	52.7	49.8				
	$\text{BP}_{SN}(2)$	53.9	48.4	49.1	42.6	40.4	35.8	32.9	29.9	27.1	26.4	25.1	22.0				
	$\text{LBM}_{SN}(1)$	82.7	77.4	75.5	71.5	69.0	65.1	64.2	60.8	58.7	56.6	53.1	49.9				
	$\text{LBM}_{SN}(2)$	53.9	48.5	49.2	42.6	40.5	35.8	33.0	29.9	27.2	26.4	25.2	22.1				
	$\text{BP}_S$	n.a.	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$\text{LBM}_S$	n.a.	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$\text{BP}_W$	97.0	98.0	97.5	97.3	95.9	94.9	93.4	92.5	91.8	89.9	89.6	88.3				
	$\text{LBM}_W$	97.0	98.0	97.5	97.3	95.9	94.9	93.4	92.6	91.8	90.1	89.6	88.4				
	$\text{BP}_{SN}$	80.0	81.4	79.1	76.4	73.7	67.2	63.0	59.6	53.8	50.7	46.0	41.2				
	$\text{LBM}_{SN}$	80.0	81.5	79.2	76.5	73.7	67.5	63.0	59.8	54.0	51.0	46.3	41.4				

## 4.5 Adequacy of weak PARMA models for real datasets

We now consider an application to the daily log returns (also simply called the returns) of four stock market indices (closing values): CAC 40 (Paris), DAX (Frankfurt), Nikkei (Osaka) and SP500 (New York). The returns are defined by  $r_t = 100 \log(p_t/p_{t-1})$  where  $p_t$  denotes the price index of the corresponding index at time  $t$ . The observations cover the period from January 4, 1999 to November 20, 2020. The data can be downloaded from the website Yahoo Finance: <http://fr.finance.yahoo.com/>.

In Financial Econometric the returns are often assumed to be a weak white noise (though they are not generally independent sequences). In view of the so-called volatility clustering it is well known that the strong white noise assumption is generally not adequate for these series (see for instance [10, 13, 22, 33]). The squares of the returns have often second-order moments close to those of an ARMA(1, 1) (which is compatible with a GARCH(1, 1) model for the returns). We will test if these hypotheses remind valid in the case of periodic models by fitting weak PARMA models on the returns and on their squares.

A day-of-the-week seasonality property of the stock markets returns series was largely investigated in the literature by [1, 4, 7, 20, 23, 40], to name a few. Most of these studies focus on the description of day-of-the-week seasonality in returns and volatility. In particular the so-called Monday effect in the finance literature was observed in many stock markets. Therefore we focus on the dynamics of the squared returns  $(r_t^2)_{t \geq 1}$  process.

Because of the presence of holidays many weeks comprise less than five observations. We removed the entire weeks when there was less than five data. The effective number of observations used for each index is given in Tables 4.15 and 4.16 and the periodicity is then  $T = 5$ .

As above-mentioned it is well known that daily observed stock returns tend to show the day-of-the-week effects. Table 4.12 displays the sample correlations between  $r_{5n+1}$  (Monday) and  $r_{5n+v-h}$  (other days for  $v = 1, \dots, T = 5$ ); between  $r_{5n+1}^2$  and  $r_{5n+v-h}^2$  for the four indices considered. From Table 4.12 we observe that: a) the returns on Mondays generally display small correlations with those of other days while the squared returns on Mondays are generally strongly correlated; b) the squared returns on Mondays are also positively correlated with those of the other days; c) the squared returns on Mondays are frequently strongly correlated with those of the preceding Fridays (except for Nikkei index).

First we apply portmanteau tests on each series of daily returns for checking the hypothesis that the returns constitute a periodic white noise. In this section we only present the results on the Ljung-Box-McLeod tests since they are very close to those of the Box-Pierce tests. Tables 4.13 and 4.14 display the  $p$ -values and the statistics (for the self-normalized versions) of the standard and modified LBM tests for the mean corrected returns of each index. For the CAC 40, DAX and SP500 series, the periodic strong white noise hypothesis is rejected by the standard global LBM test and even frequently for a specified season  $v \in \{1, \dots, 5\}$  at the nominal level  $\alpha = 5\%$ . This is not surprising because as above-mentioned the standard test required the iid assumption and, in particular in view of the so-called volatility clustering, it is well known that the strong white noise model is not adequate for these series (see for instance [10, 20, 22]). For the Nikkei index the periodic strong white noise assumption is not rejected. But since the class of periodic strong noises is a subset of the class of periodic weak noises, these results show that the standard inference based on the assumption of a periodic strong noise can be misleading (see for instance [20]). By contrast, the periodic weak white noise hypothesis is not rejected for the four indices by all the global proposed tests, even for a specified season  $v$ . To summarize, the outputs of Tables 4.13 and 4.14 are in accordance with the common belief that these series are not strong white noises but could be weak white noises.

Next, turning to the dynamics of the squared returns by fitting a following weak PARMA<sub>5</sub>(1, 1) model. Denoting by  $(X_{5n+v})_{n \geq 0}$  the mean corrected series of the squared returns  $(r_{5n+v}^2)_{n \geq 0}$ , we

fit the model

$$X_{5n+v} - \phi(v)X_{5n+v-1} = \epsilon_{5n+v} - \theta(v)\epsilon_{5n+v-1}. \quad (4.36)$$

To check the stationary properties of  $X_{5n+v}$  it is convenient to consider the solution to the characteristic equation of the autoregressive part of Equation (4.36), which with our notations in Section 4.2 can be shown to be equal to

$$\det(\Phi(z)) = \det(\Phi_0 - \Phi_1 z) = 1 - \phi(1)\phi(2)\phi(3)\phi(4)\phi(5)z = 0 \text{ for } |z| > 1 \quad (4.37)$$

where

$$\Phi_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -\phi(2) & 1 & 0 & 0 & 0 \\ 0 & -\phi(3) & 1 & 0 & 0 \\ 0 & 0 & -\phi(4) & 1 & 0 \\ 0 & 0 & 0 & -\phi(5) & 1 \end{pmatrix} \text{ and } \Phi_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & \phi(1) \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

To show that  $X_{5n+v}$  is invertible, it is convenient to consider the solution to the characteristic equation of the moving average part of Equation (4.36). The same previous result holds for the invertible Model (4.36) with  $\phi(v)$  replaced by  $\theta(v)$  in (4.37).

Table 4.15 (resp. Table 4.16) presents the QLS estimated parameters of Model (4.36), their  $p$ -values (in parentheses) and their estimated standard errors (into brackets, under weak assumption on  $\epsilon_{5n+v}$ ) of the squared returns of the CAC 40 and DAX indices (resp. the SP500 and Nikkei indices). As expected all the estimated parameters are significant at any reasonable levels, except:  $\hat{\mu}_4$  for the CAC 40 index;  $\hat{\mu}_2$  (at 1% and 5% levels) and  $\hat{\mu}_4$  for the DAX index;  $\hat{\mu}_2$ ,  $\hat{\mu}_3$  and  $\hat{\mu}_4$  for the SP500 index. For all indices the mean  $\hat{\mu}_1$  on Monday is positive and significant, it seems more possibly to talk of a global Monday effect. For the other days the means  $\hat{\mu}_2$ ,  $\hat{\mu}_3$ ,  $\hat{\mu}_4$  and  $\hat{\mu}_5$  are all negatives. Therefore Tuesday, Wednesday, Thursday and Friday seem the bad days for the Nikkei index since they are significant. Tuesday, Wednesday and Friday seem also bad days for the CAC 40 and DAX (at 5% level) indices. Finally Friday seems a particularly bad day for the SP500 index.

For all indices the autoregressive coefficients  $\hat{\phi}(v)$  are all positive and significant for all days. These coefficients are greater than one on Monday (for CAC 40, DAX and SP500), on Wednesday (for CAC 40, SP500 and Nikkei) and on Friday for Nikkei index. We also observe that the coefficients  $\hat{\phi}(v)$  are the biggest on Monday (for CAC 40, DAX and SP500) and on Wednesday (for Nikkei). Furthermore, with these three indices (CAC 40, DAX and SP500) and the period considered, it is probably more appropriate to talk of a Monday effect. By contrast for the Nikkei index, it is probably more appropriate to talk of a Wednesday effect rather than a Monday effect. Additionally, Tables 4.15 and 4.16 show that there is evidence that the estimated noise standard deviations (estimated volatility) for all indices is considerably greater on Monday than the other days (Tuesday, Wednesday, Thursday and Friday), for which it is smaller and almost constant.

Note that from Tables 4.15 and 4.16 and for all series, the product of the estimated coefficients  $\hat{\phi}(v)$  (resp.  $\hat{\theta}(v)$ ) are smaller than one. Thus in view of (4.37)  $|z| > 1$  for all series (the smallest  $z$  verify  $|z| = 1.126 > 1$ ). So we think that the assumption (A1) is satisfied and thus our asymptotic normality theorem on the residual autocorrelations can be applied. After estimating the model (4.36), the next important step in the modeling consists in checking if the estimated model fits satisfactorily the data. Thus under the null hypothesis that the model has been correctly identified, the residuals are approximately a white noise.

We thus apply portmanteau tests to the residuals of Model (4.36) for each series. Tables 4.17 and 4.18 display the  $p$ -values and the statistics (for the self-normalized versions) of the standard and modified LBM tests for the mean corrected returns of each index. From these tables we draw the conclusion that the strong PARMA<sub>5</sub>(1, 1) model is rejected by the standard

global LBM test and even for a specified season  $v$  at the nominal level  $\alpha = 5\%$ . By contrast a weak  $\text{PARMA}_5(1, 1)$  model is not rejected. Note that for the first and second-order structures we found for the returns considered, namely a periodic weak white noise for the returns and a weak  $\text{PARMA}_5(1, 1)$  model for the squares of the returns, are compatible with a  $\text{PGARCH}(1, 1)$  model.

Figures 4.1, 4.2, 4.3 and 4.4 display the residual autocorrelations and their 5% significance limits under the strong  $\text{PARMA}_5(1, 1)$  and weak  $\text{PARMA}_5(1, 1)$  assumptions. In view of Figures 4.1, 4.2, 4.3 and 4.4, the diagnostic checking of residuals does not indicate any inadequacy for the proposed tests. All of the sample autocorrelations should lie between the bands (at 95%) shown as dashed lines (green color) and solid lines (red color) for the modified tests, while the horizontal dotted (blue color) for standard test indicate that strong  $\text{PARMA}_5(1, 1)$  is not adequate. Figures 4.1, 4.2, 4.3 and 4.4 confirm the conclusions drawn from Tables 4.17 and 4.18.

To conclude our empirical investigations, a comparison of the four indices indicates that Nikkei is systematically more volatile over two days of the week (Monday and Friday) than the other three. By contrast the DAX index is systematically more volatile on Tuesday and Thursday than the other three. Finally SP500 index is the most volatile on Wednesday.



4.5. ADEQUACY OF WEAK PARMA MODELS FOR REAL DATASETS

Table 4.12 – Sample correlations between  $r_{5n+1}$  (Monday) and  $r_{5n+v-h}$  (other days); between  $r_{5n+1}^2$  and  $r_{5n+v-h}^2$  for the four indices considered.

Index	Correlations	Lag $h$									
		1	2	3	4	5	6	7	8	9	10
CAC 40	$\hat{\rho}(r_{5n+1}, r_{5n+1-h})$	-0.047	-0.032	-0.011	-0.005	-0.004	0.053	-0.014	0.045	0.016	-0.011
	$\hat{\rho}(r_{5n+1}, r_{5n+2-h})$	0.003	-0.059	-0.016	-0.018	-0.040	0.002	-0.036	-0.013	-0.058	-0.007
	$\hat{\rho}(r_{5n+1}, r_{5n+3-h})$	-0.013	0.050	0.045	-0.012	0.005	-0.008	0.024	-0.022	-0.016	-0.051
	$\hat{\rho}(r_{5n+1}, r_{5n+4-h})$	0.010	0.068	-0.047	0.004	0.005	-0.024	-0.033	0.033	0.029	0.038
	$\hat{\rho}(r_{5n+1}, r_{5n+5-h})$	-0.026	0.023	0.025	-0.008	0.015	-0.005	-0.041	0.006	-0.046	0.055
	$\hat{\rho}(r_{5n+1}^2, r_{5n+1-h}^2)$	0.220	0.139	0.076	0.098	0.040	0.132	0.114	0.090	0.097	0.039
	$\hat{\rho}(r_{5n+1}^2, r_{5n+2-h}^2)$	0.153	0.204	0.116	0.162	0.098	0.065	0.100	0.096	0.030	0.076
	$\hat{\rho}(r_{5n+1}^2, r_{5n+3-h}^2)$	0.311	0.325	0.237	0.194	0.121	0.112	0.044	0.042	0.045	0.072
	$\hat{\rho}(r_{5n+1}^2, r_{5n+4-h}^2)$	0.102	0.109	0.093	0.085	0.051	0.109	0.041	0.042	0.029	0.058
	$\hat{\rho}(r_{5n+1}^2, r_{5n+5-h}^2)$	0.220	0.090	0.057	0.034	0.030	0.018	0.080	0.093	0.082	0.029
DAX	$\hat{\rho}(r_{5n+1}, r_{5n+1-h})$	-0.038	-0.029	-0.028	0.013	0.003	0.024	-0.015	0.037	0.026	0.003
	$\hat{\rho}(r_{5n+1}, r_{5n+2-h})$	0.008	-0.022	-0.039	-0.080	-0.036	-0.017	-0.007	-0.004	-0.035	-0.023
	$\hat{\rho}(r_{5n+1}, r_{5n+3-h})$	-0.044	0.091	0.083	0.019	-0.031	0.033	0.050	-0.022	-0.020	0.007
	$\hat{\rho}(r_{5n+1}, r_{5n+4-h})$	0.037	0.076	-0.009	-0.017	-0.005	-0.030	0.008	0.060	0.019	0.017
	$\hat{\rho}(r_{5n+1}, r_{5n+5-h})$	-0.053	0.032	0.043	0.031	-0.018	-0.045	0.007	-0.007	-0.035	0.041
	$\hat{\rho}(r_{5n+1}^2, r_{5n+1-h}^2)$	0.115	0.091	0.050	0.073	0.035	0.056	0.053	0.057	0.035	0.078
	$\hat{\rho}(r_{5n+1}^2, r_{5n+2-h}^2)$	0.121	0.264	0.130	0.120	0.083	0.051	0.085	0.019	0.007	0.042
	$\hat{\rho}(r_{5n+1}^2, r_{5n+3-h}^2)$	0.241	0.101	0.119	0.115	0.154	0.094	0.080	0.072	0.119	0.063
	$\hat{\rho}(r_{5n+1}^2, r_{5n+4-h}^2)$	0.077	0.089	0.063	0.048	0.047	0.057	0.049	0.024	0.066	0.037
	$\hat{\rho}(r_{5n+1}^2, r_{5n+5-h}^2)$	0.174	0.095	0.082	0.056	0.069	0.013	0.073	0.082	0.110	0.063
SP500	$\hat{\rho}(r_{5n+1}, r_{5n+1-h})$	-0.019	-0.035	-0.057	-0.027	0.010	-0.012	0.064	0.027	0.071	-0.059
	$\hat{\rho}(r_{5n+1}, r_{5n+2-h})$	0.037	0.011	0.049	-0.071	0.046	-0.002	-0.023	0.022	-0.030	-0.013
	$\hat{\rho}(r_{5n+1}, r_{5n+3-h})$	0.006	0.062	-0.023	-0.029	0.045	0.047	0.077	0.023	0.019	0.020
	$\hat{\rho}(r_{5n+1}, r_{5n+4-h})$	0.040	-0.084	-0.017	0.050	-0.049	-0.076	-0.014	0.002	0.058	-0.006
	$\hat{\rho}(r_{5n+1}, r_{5n+5-h})$	0.046	0.027	0.014	-0.080	0.042	0.047	-0.007	-0.044	0.020	0.015
	$\hat{\rho}(r_{5n+1}^2, r_{5n+1-h}^2)$	0.167	0.252	0.097	0.143	0.093	0.025	0.112	0.064	0.066	0.031
	$\hat{\rho}(r_{5n+1}^2, r_{5n+2-h}^2)$	0.247	0.462	0.147	0.233	0.080	0.119	0.080	0.100	0.059	0.060
	$\hat{\rho}(r_{5n+1}^2, r_{5n+3-h}^2)$	0.204	0.237	0.229	0.150	0.170	0.084	0.074	0.045	0.040	0.037
	$\hat{\rho}(r_{5n+1}^2, r_{5n+4-h}^2)$	0.425	0.195	0.193	0.198	0.280	0.173	0.163	0.070	0.060	0.054
	$\hat{\rho}(r_{5n+1}^2, r_{5n+5-h}^2)$	0.097	0.126	0.194	0.160	0.230	0.157	0.162	0.076	0.104	0.040
Nikkei	$\hat{\rho}(r_{5n+1}, r_{5n+1-h})$	-0.071	-0.010	0.045	0.004	0.015	-0.028	0.006	-0.016	0.025	0.005
	$\hat{\rho}(r_{5n+1}, r_{5n+2-h})$	-0.004	-0.037	-0.003	-0.030	-0.002	-0.052	0.055	-0.034	0.009	0.016
	$\hat{\rho}(r_{5n+1}, r_{5n+3-h})$	0.059	-0.030	0.047	-0.003	-0.002	-0.011	0.007	0.037	-0.020	-0.018
	$\hat{\rho}(r_{5n+1}, r_{5n+4-h})$	0.032	0.054	-0.066	0.028	-0.012	-0.023	-0.006	-0.022	-0.028	-0.028
	$\hat{\rho}(r_{5n+1}, r_{5n+5-h})$	0.000	0.004	-0.025	0.049	0.001	0.031	0.032	-0.005	-0.003	0.011
	$\hat{\rho}(r_{5n+1}^2, r_{5n+1-h}^2)$	0.069	0.038	0.018	0.023	0.054	0.059	0.043	0.018	0.015	0.046
	$\hat{\rho}(r_{5n+1}^2, r_{5n+2-h}^2)$	0.051	0.044	0.030	0.042	0.016	0.020	0.029	0.007	0.003	0.015
	$\hat{\rho}(r_{5n+1}^2, r_{5n+3-h}^2)$	0.114	0.096	0.049	0.004	0.050	0.016	0.040	0.041	0.026	0.045
	$\hat{\rho}(r_{5n+1}^2, r_{5n+4-h}^2)$	0.194	0.109	0.102	0.032	0.032	0.015	-0.002	0.033	0.008	0.120
	$\hat{\rho}(r_{5n+1}^2, r_{5n+5-h}^2)$	0.071	0.028	0.053	0.039	0.047	0.007	0.010	0.027	0.019	0.036

4.5. ADEQUACY OF WEAK PARMA MODELS FOR REAL DATASETS

Table 4.13 – Modified and standard versions of portmanteau tests to check the null hypothesis that the returns follow a periodic white noise, based on  $m$  squared residuals autocovariances.

Index	Tests	Lag $m$											
		1	2	3	4	5	6	7	8	9	10	11	12
CAC 40	LBM <sub>S</sub> (1)	0.083	0.113	0.210	0.335	0.468	0.210	0.277	0.178	0.225	0.286	0.351	0.431
	LBM <sub>S</sub> (2)	0.001	0.000	0.000	0.001	0.002	0.003	0.005	0.009	0.010	0.014	0.014	0.016
	LBM <sub>S</sub> (3)	0.617	0.806	0.508	0.675	0.051	0.068	0.079	0.104	0.116	0.147	0.189	0.083
	LBM <sub>S</sub> (4)	0.504	0.570	0.276	0.424	0.516	0.508	0.253	0.248	0.310	0.393	0.461	0.436
	LBM <sub>S</sub> (5)	0.423	0.674	0.324	0.079	0.071	0.075	0.010	0.018	0.031	0.047	0.066	0.016
	LBM <sub>W</sub> (1)	0.263	0.435	0.570	0.710	0.797	0.641	0.695	0.619	0.664	0.711	0.755	0.817
	LBM <sub>W</sub> (2)	0.005	0.007	0.007	0.015	0.039	0.054	0.085	0.111	0.129	0.170	0.159	0.176
	LBM <sub>W</sub> (3)	0.782	0.905	0.766	0.872	0.374	0.414	0.436	0.501	0.523	0.591	0.654	0.522
	LBM <sub>W</sub> (4)	0.631	0.760	0.607	0.745	0.833	0.840	0.671	0.675	0.740	0.814	0.848	0.823
	LBM <sub>W</sub> (5)	0.604	0.842	0.678	0.433	0.428	0.419	0.236	0.295	0.345	0.389	0.446	0.293
	LBM <sub>SN</sub> (1)	8.8	22.3	32.5	33.1	61.2	61.5	62.3	86.8	253.2	262.0	295.6	333.8
	LBM <sub>SN</sub> (2)	75.0	152.0	156.6	370.1	406.7	410.1	445.7	905.0	906.0	907.0	1071.4	1208.8
	LBM <sub>SN</sub> (3)	0.9	10.6	10.9	16.2	126.6	136.7	137.9	158.9	178.7	189.4	203.6	216.5
	LBM <sub>SN</sub> (4)	1.3	11.6	49.7	57.6	60.7	150.3	375.1	382.4	383.3	383.3	430.5	462.0
	LBM <sub>SN</sub> (5)	2.7	31.5	55.8	127.8	166.0	187.3	188.4	327.8	403.3	544.2	548.1	699.1
	LBM <sub>S</sub>	0.011	0.016	0.005	0.009	0.003	0.002	0.001	0.001	0.002	0.007	0.015	0.004
	LBM <sub>W</sub>	0.247	0.426	0.420	0.513	0.472	0.456	0.359	0.404	0.494	0.610	0.668	0.572
	LBM <sub>SN</sub>	201.6	367.2	741.2	1334.8	1660.1	2363.3	2425.7	3318.8	4241.9	5046.5	5889.5	6504.8
DAX	LBM <sub>S</sub> (1)	0.166	0.217	0.247	0.360	0.498	0.524	0.602	0.499	0.506	0.600	0.367	0.368
	LBM <sub>S</sub> (2)	0.016	0.014	0.033	0.062	0.067	0.016	0.026	0.010	0.016	0.002	0.003	0.004
	LBM <sub>S</sub> (3)	0.298	0.408	0.393	0.328	0.147	0.217	0.298	0.194	0.266	0.127	0.075	0.085
	LBM <sub>S</sub> (4)	0.205	0.005	0.008	0.009	0.001	0.002	0.002	0.004	0.005	0.009	0.003	0.003
	LBM <sub>S</sub> (5)	0.256	0.148	0.236	0.295	0.052	0.087	0.070	0.089	0.128	0.090	0.103	0.018
	LBM <sub>W</sub> (1)	0.350	0.496	0.546	0.674	0.773	0.806	0.849	0.800	0.799	0.859	0.751	0.773
	LBM <sub>W</sub> (2)	0.048	0.123	0.175	0.273	0.329	0.205	0.271	0.198	0.230	0.116	0.137	0.199
	LBM <sub>W</sub> (3)	0.555	0.685	0.738	0.749	0.643	0.733	0.798	0.725	0.787	0.690	0.628	0.650
	LBM <sub>W</sub> (4)	0.437	0.140	0.190	0.237	0.177	0.210	0.227	0.253	0.279	0.322	0.247	0.246
	LBM <sub>W</sub> (5)	0.438	0.437	0.560	0.651	0.356	0.420	0.410	0.443	0.503	0.446	0.475	0.283
	LBM <sub>SN</sub> (1)	13.7	14.0	66.0	117.0	117.3	121.1	121.4	148.5	180.7	181.0	209.0	252.7
	LBM <sub>SN</sub> (2)	18.9	40.8	43.7	62.6	62.6	63.0	113.3	116.1	117.2	179.0	182.7	187.7
	LBM <sub>SN</sub> (3)	1.9	16.6	55.4	89.4	137.6	172.1	189.5	218.1	218.2	357.6	412.9	462.7
	LBM <sub>SN</sub> (4)	1.9	50.2	86.8	140.6	140.9	170.3	171.5	222.7	300.2	316.6	319.5	390.1
	LBM <sub>SN</sub> (5)	4.8	83.5	84.7	194.1	208.5	209.8	259.8	330.4	359.3	392.6	430.2	430.8
	LBM <sub>S</sub>	0.040	0.002	0.006	0.014	0.001	0.001	0.001	0.001	0.002	0.000	0.000	0.000
	LBM <sub>W</sub>	0.407	0.307	0.470	0.622	0.455	0.481	0.558	0.509	0.586	0.479	0.405	0.399
	LBM <sub>SN</sub>	143.5	541.5	736.3	960.5	1644.0	2049.5	2363.4	3694.2	3982.4	5504.6	6356.8	6779.2

4.5. ADEQUACY OF WEAK PARMA MODELS FOR REAL DATASETS

Table 4.14 – Modified and standard versions of portmanteau tests to check the null hypothesis that the returns follow a periodic white noise, based on  $m$  squared residuals autocovariances.

Index	Tests	Lag $m$											
		1	2	3	4	5	6	7	8	9	10	11	12
SP500	LBM <sub>S</sub> (1)	0.496	0.349	0.097	0.124	0.195	0.272	0.075	0.084	0.015	0.005	0.008	0.002
	LBM <sub>S</sub> (2)	0.000	0.000	0.000	0.001	0.002	0.003	0.007	0.012	0.020	0.027	0.028	0.028
	LBM <sub>S</sub> (3)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	LBM <sub>S</sub> (4)	0.960	0.822	0.000	0.000	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000
	LBM <sub>S</sub> (5)	0.224	0.470	0.100	0.116	0.192	0.012	0.007	0.002	0.004	0.003	0.001	0.001
	LBM <sub>W</sub> (1)	0.686	0.765	0.585	0.657	0.732	0.783	0.608	0.638	0.443	0.343	0.381	0.294
	LBM <sub>W</sub> (2)	0.027	0.048	0.157	0.215	0.264	0.340	0.409	0.476	0.544	0.583	0.589	0.607
	LBM <sub>W</sub> (3)	0.152	0.193	0.425	0.363	0.299	0.325	0.331	0.343	0.329	0.316	0.318	0.297
	LBM <sub>W</sub> (4)	0.983	0.948	0.133	0.174	0.247	0.231	0.289	0.252	0.213	0.219	0.148	0.172
	LBM <sub>W</sub> (5)	0.365	0.666	0.315	0.407	0.504	0.152	0.130	0.104	0.127	0.130	0.080	0.103
	LBM <sub>SN</sub> (1)	0.4	34.8	115.4	140.6	155.1	172.6	262.9	267.9	294.9	326.4	334.8	582.8
	LBM <sub>SN</sub> (2)	25.9	37.9	142.6	166.0	185.5	192.0	284.5	295.9	337.4	343.7	375.5	384.6
	LBM <sub>SN</sub> (3)	13.5	19.3	25.9	45.3	45.3	203.9	237.5	243.1	348.0	350.4	354.3	366.3
	LBM <sub>SN</sub> (4)	0.0	1.7	93.4	93.6	93.8	94.9	102.4	140.6	213.8	216.8	233.0	233.1
	LBM <sub>SN</sub> (5)	12.8	14.3	73.2	93.6	153.2	234.0	244.7	300.1	419.2	426.1	441.9	462.7
	LBM <sub>S</sub>	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	LBM <sub>W</sub>	0.149	0.354	0.279	0.361	0.359	0.342	0.324	0.322	0.292	0.284	0.258	0.263
	LBM <sub>SN</sub>	205.3	384.9	658.0	1445.5	1563.3	1787.3	1937.5	2006.4	2458.9	2797.1	3173.7	3848.8
	Nikkei	LBM <sub>S</sub> (1)	0.015	0.048	0.037	0.075	0.119	0.138	0.204	0.262	0.290	0.371	0.441
LBM <sub>S</sub> (2)		0.089	0.222	0.196	0.319	0.219	0.139	0.158	0.205	0.127	0.178	0.208	0.213
LBM <sub>S</sub> (3)		0.571	0.832	0.697	0.837	0.851	0.920	0.952	0.974	0.986	0.993	0.992	0.996
LBM <sub>S</sub> (4)		0.069	0.157	0.265	0.336	0.307	0.404	0.503	0.551	0.622	0.477	0.556	0.567
LBM <sub>S</sub> (5)		0.199	0.412	0.273	0.128	0.121	0.023	0.040	0.022	0.032	0.049	0.067	0.083
LBM <sub>W</sub> (1)		0.045	0.117	0.079	0.137	0.209	0.256	0.355	0.403	0.431	0.525	0.595	0.567
LBM <sub>W</sub> (2)		0.118	0.343	0.438	0.537	0.461	0.396	0.413	0.471	0.376	0.432	0.458	0.462
LBM <sub>W</sub> (3)		0.652	0.898	0.805	0.897	0.910	0.956	0.972	0.987	0.993	0.997	0.996	0.998
LBM <sub>W</sub> (4)		0.279	0.445	0.574	0.622	0.577	0.639	0.701	0.740	0.778	0.678	0.758	0.755
LBM <sub>W</sub> (5)		0.620	0.769	0.696	0.568	0.573	0.398	0.449	0.373	0.405	0.436	0.460	0.488
LBM <sub>SN</sub> (1)		54.4	54.5	126.8	127.8	221.0	226.9	579.8	605.1	609.4	919.6	923.9	979.8
LBM <sub>SN</sub> (2)		14.0	14.0	85.0	123.7	171.5	176.9	177.6	215.6	240.1	240.2	319.8	320.4
LBM <sub>SN</sub> (3)		2.1	2.4	5.2	10.9	21.9	23.1	25.5	53.4	53.4	71.2	98.1	98.3
LBM <sub>SN</sub> (4)		5.3	6.1	54.0	94.8	100.5	115.2	126.8	127.5	127.7	176.7	324.8	345.9
LBM <sub>SN</sub> (5)		2.7	3.6	18.5	40.3	111.3	133.8	229.6	291.9	292.8	308.1	340.9	499.5
LBM <sub>S</sub>		0.015	0.135	0.096	0.155	0.145	0.069	0.153	0.182	0.211	0.291	0.409	0.444
LBM <sub>W</sub>		0.323	0.609	0.597	0.641	0.630	0.545	0.640	0.671	0.689	0.742	0.809	0.821
LBM <sub>SN</sub>		343.6	426.6	1013.1	1309.4	1602.2	2142.9	2931.9	3909.1	4737.3	5585.1	7404.2	8118.3

Table 4.15 – QLS estimates, their  $p$ -values (in parentheses) and their estimated standard errors (in brackets) of a weak PARMA<sub>5</sub>(1, 1) model fitted to the mean corrected series of the squared returns of the CAC 40 and DAX indices.

Index	CAC 40				DAX			
NT	6830				6740			
Day	$\hat{\mu}_v$	$\hat{\phi}_v$	$\hat{\theta}_v$	$\hat{\sigma}_v$	$\hat{\mu}_v$	$\hat{\phi}_v$	$\hat{\theta}_v$	$\hat{\sigma}_v$
Monday	1.094 (0.000) [0.256]	2.183 (0.000) [0.225]	1.958 (0.000) [0.197]	$8.844 \times 10^{-8}$	1.399 (0.000) [0.335]	2.062 (0.000) [0.315]	1.825 (0.000) [0.279]	$11.964 \times 10^{-8}$
Tuesday	-0.363 (0.001) [0.113]	0.444 (0.004) [0.153]	0.411 (0.006) [0.148]	$3.958 \times 10^{-8}$	-0.279 (0.098) [0.168]	0.954 (0.000) [0.182]	0.929 (0.000) [0.179]	$5.458 \times 10^{-8}$
Wednesday	-0.325 (0.005) [0.116]	1.501 (0.007) [0.552]	1.501 (0.008) [0.567]	$3.648 \times 10^{-8}$	-0.459 (0.000) [0.107]	0.702 (0.000) [0.095]	0.757 (0.000) [0.074]	$3.367 \times 10^{-8}$
Thursday	-0.059 (0.730) [0.171]	0.840 (0.000) [0.219]	0.767 (0.000) [0.198]	$5.979 \times 10^{-8}$	-0.214 (0.224) [0.176]	0.960 (0.000) [0.140]	0.846 (0.000) [0.140]	$6.103 \times 10^{-8}$
Friday	-0.346 (0.007) [0.128]	0.727 (0.000) [0.165]	0.652 (0.000) [0.169]	$4.491 \times 10^{-8}$	-0.447 (0.000) [0.115]	0.664 (0.000) [0.135]	0.595 (0.000) [0.132]	$3.978 \times 10^{-8}$

Table 4.16 – QLS estimates, their  $p$ -values (in parentheses) and their estimated standard errors (in brackets) of a weak PARMA<sub>5</sub>(1, 1) model fitted to the mean corrected series of the squared returns of the SP500 and Nikkei indices.

Index	SP500				Nikkei			
NT	6550				5920			
Day	$\hat{\mu}_v$	$\hat{\phi}_v$	$\hat{\theta}_v$	$\hat{\sigma}_v$	$\hat{\mu}_v$	$\hat{\phi}_v$	$\hat{\theta}_v$	$\hat{\sigma}_v$
Monday	0.718 (0.001) [0.213]	1.900 (0.000) [0.435]	1.613 (0.008) [0.611]	$7.163 \times 10^{-8}$	2.716 (0.000) [0.418]	0.738 (0.000) [0.109]	0.529 (0.000) [0.128]	$14.391 \times 10^{-8}$
Tuesday	-0.149 (0.250) [0.129]	0.871 (0.000) [0.060]	0.747 (0.000) [0.071]	$4.093 \times 10^{-8}$	-0.994 (0.000) [0.138]	0.491 (0.000) [0.093]	0.472 (0.000) [0.093]	$4.690 \times 10^{-8}$
Wednesday	-0.116 (0.502) [0.173]	1.518 (0.000) [0.283]	1.362 (0.000) [0.185]	$4.958 \times 10^{-8}$	-0.571 (0.000) [0.154]	1.669 (0.000) [0.159]	1.551 (0.000) [0.172]	$4.817 \times 10^{-8}$
Thursday	-0.061 (0.620) [0.123]	0.648 (0.000) [0.118]	0.490 (0.000) [0.109]	$3.687 \times 10^{-8}$	-0.646 (0.000) [0.150]	0.900 (0.000) [0.104]	0.654 (0.000) [0.093]	$4.670 \times 10^{-8}$
Friday	-0.391 (0.000) [0.074]	0.497 (0.000) [0.037]	0.402 (0.000) [0.042]	$2.418 \times 10^{-8}$	-0.500 (0.006) [0.182]	1.357 (0.000) [0.241]	1.397 (0.000) [0.363]	$5.360 \times 10^{-8}$

Table 4.17 – Modified and standard versions of portmanteau tests to check the null hypothesis that the squared returns follow a weak  $PARMA_5(1, 1)$  models, based on  $m$  squared residuals autocovariances.

Index	Tests	Lag $m$												
		1	2	3	4	5	6	7	8	9	10	11	12	
CAC 40	LBM <sub>S</sub> (1)	n.a.	n.a.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	LBM <sub>S</sub> (2)	n.a.	n.a.	0.063	0.177	0.317	0.474	0.209	0.306	0.195	0.271	0.238	0.247	
	LBM <sub>S</sub> (3)	n.a.	n.a.	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	LBM <sub>S</sub> (4)	n.a.	n.a.	0.034	0.019	0.044	0.078	0.127	0.197	0.280	0.038	0.061	0.080	
	LBM <sub>S</sub> (5)	n.a.	n.a.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	LBM <sub>W</sub> (1)	0.652	0.401	0.275	0.336	0.307	0.307	0.328	0.382	0.386	0.382	0.362	0.372	
	LBM <sub>W</sub> (2)	0.509	0.682	0.725	0.771	0.849	0.892	0.767	0.806	0.804	0.838	0.818	0.806	
	LBM <sub>W</sub> (3)	0.047	0.414	0.460	0.391	0.430	0.392	0.374	0.366	0.369	0.356	0.358	0.347	
	LBM <sub>W</sub> (4)	0.291	0.311	0.550	0.462	0.524	0.517	0.544	0.581	0.590	0.478	0.491	0.488	
	LBM <sub>W</sub> (5)	0.409	0.555	0.260	0.254	0.267	0.240	0.253	0.203	0.215	0.249	0.268	0.274	
	LBM <sub>SN</sub> (1)	1.0	15.1	23.8	23.9	36.6	42.1	49.2	159.3	196.7	277.4	295.0	313.7	
	LBM <sub>SN</sub> (2)	2.6	2.6	6.2	29.7	29.7	35.2	96.3	106.1	137.0	142.3	228.6	250.8	
	LBM <sub>SN</sub> (3)	118.7	128.9	150.7	152.5	153.8	155.9	177.2	182.5	241.4	243.0	244.9	264.7	
	LBM <sub>SN</sub> (4)	3.0	20.7	23.9	58.2	85.7	117.0	122.5	126.4	126.8	129.7	134.4	403.5	
	LBM <sub>SN</sub> (5)	4.9	8.1	15.9	19.4	75.1	113.2	113.6	254.1	269.6	643.0	659.9	742.0	
	LBM <sub>S</sub>	n.a.	n.a.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	LBM <sub>W</sub>	0.523	0.558	0.416	0.406	0.414	0.388	0.345	0.329	0.317	0.314	0.332	0.358	
	LBM <sub>SN</sub>	239.6	483.0	641.0	771.3	830.9	1068.7	1300.6	1632.9	2460.4	2814.5	3299.4	6016.9	
DAX	LBM <sub>S</sub> (1)	n.a.	n.a.	0.131	0.319	0.346	0.483	0.627	0.735	0.796	0.776	0.701	0.744	
	LBM <sub>S</sub> (2)	n.a.	n.a.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	LBM <sub>S</sub> (3)	n.a.	n.a.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	LBM <sub>S</sub> (4)	n.a.	n.a.	0.027	0.085	0.112	0.198	0.228	0.175	0.254	0.344	0.415	0.470	
	LBM <sub>S</sub> (5)	n.a.	n.a.	0.002	0.008	0.023	0.010	0.016	0.005	0.008	0.014	0.012	0.000	
	LBM <sub>W</sub> (1)	0.530	0.772	0.380	0.473	0.361	0.426	0.503	0.567	0.573	0.648	0.610	0.716	
	LBM <sub>W</sub> (2)	0.234	0.291	0.070	0.138	0.130	0.157	0.196	0.190	0.187	0.206	0.242	0.255	
	LBM <sub>W</sub> (3)	0.650	0.375	0.391	0.427	0.515	0.548	0.419	0.453	0.401	0.345	0.341	0.355	
	LBM <sub>W</sub> (4)	0.503	0.583	0.479	0.539	0.559	0.584	0.586	0.483	0.526	0.569	0.574	0.576	
	LBM <sub>W</sub> (5)	0.177	0.345	0.315	0.391	0.445	0.358	0.438	0.361	0.366	0.396	0.386	0.272	
	LBM <sub>SN</sub> (1)	1.0	13.5	13.5	14.6	46.1	99.0	123.6	170.0	235.9	258.6	381.2	392.2	
	LBM <sub>SN</sub> (2)	19.7	21.5	80.0	80.4	80.4	112.1	134.8	157.3	159.6	182.2	199.1	204.6	
	LBM <sub>SN</sub> (3)	2.0	25.1	45.4	48.8	98.5	104.5	108.4	109.1	117.2	170.6	178.5	182.8	
	LBM <sub>SN</sub> (4)	1.2	2.5	12.4	19.1	78.0	79.1	88.5	166.2	168.0	169.9	173.3	173.3	
	LBM <sub>SN</sub> (5)	7.8	17.3	31.6	40.0	42.7	171.4	181.6	194.3	194.8	194.8	285.6	514.8	
	LBM <sub>S</sub>	n.a.	n.a.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	LBM <sub>W</sub>	0.428	0.442	0.312	0.369	0.425	0.426	0.387	0.394	0.36	0.332	0.364	0.438	
	LBM <sub>SN</sub>	77.0	105.9	343.8	427.1	548.1	707.6	887.5	1314.5	1542.8	1850.3	2371.7	3368.8	

4.5. ADEQUACY OF WEAK PARMA MODELS FOR REAL DATASETS

Table 4.18 – Modified and standard versions of portmanteau tests to check the null hypothesis that the squared returns follow a weak PARMA<sub>5</sub>(1, 1) models, based on  $m$  squared residuals autocovariances.

Index	Tests	Lag $m$																
		1	2	3	4	5	6	7	8	9	10	11	12					
SP500	LBM <sub>S</sub> (1)	n.a.	n.a.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	LBM <sub>S</sub> (2)	n.a.	n.a.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	LBM <sub>S</sub> (3)	n.a.	n.a.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	LBM <sub>S</sub> (4)	n.a.	n.a.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	LBM <sub>S</sub> (5)	n.a.	n.a.	0.006	0.018	0.010	0.001	0.003	0.004	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.000	
	LBM <sub>W</sub> (1)	0.632	0.378	0.386	0.412	0.455	0.386	0.396	0.377	0.343	0.342	0.365	0.372	0.372	0.372	0.372	0.372	0.372
	LBM <sub>W</sub> (2)	0.097	0.299	0.312	0.298	0.394	0.401	0.384	0.381	0.388	0.402	0.408	0.407	0.407	0.407	0.407	0.407	0.407
	LBM <sub>W</sub> (3)	0.126	0.231	0.270	0.276	0.231	0.257	0.248	0.249	0.257	0.253	0.233	0.240	0.240	0.240	0.240	0.240	0.240
	LBM <sub>W</sub> (4)	0.534	0.489	0.366	0.423	0.407	0.425	0.434	0.425	0.418	0.409	0.412	0.412	0.412	0.412	0.412	0.412	0.412
	LBM <sub>W</sub> (5)	0.733	0.878	0.659	0.713	0.646	0.522	0.545	0.517	0.383	0.456	0.467	0.453	0.453	0.453	0.453	0.453	0.453
	LBM <sub>SN</sub> (1)	0.5	13.4	19.9	20.4	39.2	52.7	116.0	140.1	141.1	142.3	169.4	187.9	187.9	187.9	187.9	187.9	187.9
	LBM <sub>SN</sub> (2)	10.7	14.2	16.2	16.6	27.0	28.6	37.7	46.5	87.8	88.7	91.5	164.3	164.3	164.3	164.3	164.3	164.3
	LBM <sub>SN</sub> (3)	9.9	10.0	15.4	25.8	34.4	43.9	44.7	146.5	147.4	150.7	153.1	192.4	192.4	192.4	192.4	192.4	192.4
	LBM <sub>SN</sub> (4)	1.6	13.7	14.0	44.0	51.3	52.8	72.9	126.5	126.7	134.9	211.9	217.7	217.7	217.7	217.7	217.7	217.7
	LBM <sub>SN</sub> (5)	0.5	3.9	18.1	125.7	136.4	144.2	146.3	159.5	243.3	256.4	257.9	322.6	322.6	322.6	322.6	322.6	322.6
	LBM <sub>S</sub>	n.a.	n.a.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	LBM <sub>W</sub>	0.526	0.345	0.344	0.369	0.400	0.389	0.394	0.388	0.335	0.332	0.352	0.361	0.361	0.361	0.361	0.361	0.361
	LBM <sub>SN</sub>	51.1	234.4	303.2	629.0	882.3	1263.3	2209.1	3376.7	4499.5	4661.4	7230.0	7631.6	7631.6	7631.6	7631.6	7631.6	7631.6
Nikkei	LBM <sub>S</sub> (1)	n.a.	n.a.	0.735	0.925	0.829	0.795	0.853	0.919	0.958	0.932	0.947	0.947	0.947	0.947	0.947	0.947	
	LBM <sub>S</sub> (2)	n.a.	n.a.	0.001	0.002	0.005	0.005	0.009	0.017	0.029	0.047	0.083	0.083	0.083	0.083	0.083	0.083	
	LBM <sub>S</sub> (3)	n.a.	n.a.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	LBM <sub>S</sub> (4)	n.a.	n.a.	0.000	0.000	0.000	0.001	0.002	0.003	0.006	0.004	0.008	0.010	0.010	0.010	0.010	0.010	
	LBM <sub>S</sub> (5)	n.a.	n.a.	0.021	0.062	0.131	0.211	0.179	0.205	0.292	0.283	0.362	0.052	0.052	0.052	0.052	0.052	
	LBM <sub>W</sub> (1)	0.781	0.967	0.976	0.987	0.892	0.863	0.885	0.910	0.931	0.922	0.933	0.949	0.949	0.949	0.949	0.949	0.949
	LBM <sub>W</sub> (2)	0.228	0.256	0.281	0.281	0.292	0.272	0.279	0.289	0.294	0.291	0.303	0.312	0.312	0.312	0.312	0.312	0.312
	LBM <sub>W</sub> (3)	0.586	0.604	0.220	0.192	0.195	0.195	0.186	0.210	0.210	0.207	0.212	0.214	0.214	0.214	0.214	0.214	0.214
	LBM <sub>W</sub> (4)	0.096	0.113	0.476	0.488	0.497	0.501	0.517	0.515	0.519	0.518	0.533	0.519	0.519	0.519	0.519	0.519	0.519
	LBM <sub>W</sub> (5)	0.270	0.634	0.640	0.647	0.672	0.692	0.634	0.648	0.672	0.642	0.666	0.486	0.486	0.486	0.486	0.486	0.486
	LBM <sub>SN</sub> (1)	0.0	0.3	1.4	2.1	10.5	19.6	32.4	32.5	37.3	160.2	213.0	216.1	216.1	216.1	216.1	216.1	216.1
	LBM <sub>SN</sub> (2)	5.4	25.7	34.4	56.7	88.6	88.8	93.3	234.4	299.0	308.9	345.4	459.1	459.1	459.1	459.1	459.1	459.1
	LBM <sub>SN</sub> (3)	1.2	6.3	20.1	214.9	215.6	220.8	283.7	289.1	377.6	533.5	654.6	782.1	782.1	782.1	782.1	782.1	782.1
	LBM <sub>SN</sub> (4)	33.2	53.3	70.1	87.6	99.5	125.9	127.2	127.3	139.5	174.4	182.0	363.1	363.1	363.1	363.1	363.1	363.1
	LBM <sub>SN</sub> (5)	10.9	23.6	23.8	25.6	47.4	50.5	163.3	163.8	267.4	354.6	395.2	516.7	516.7	516.7	516.7	516.7	516.7
	LBM <sub>S</sub>	n.a.	n.a.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	LBM <sub>W</sub>	0.150	0.356	0.406	0.388	0.400	0.404	0.397	0.417	0.425	0.452	0.455	0.440	0.440	0.440	0.440	0.440	0.440
	LBM <sub>SN</sub>	102.2	140.3	1062.8	1962.3	2536.5	3060.9	3719.6	4521.2	5420.0	6002.4	6715.0	8404.3	8404.3	8404.3	8404.3	8404.3	8404.3

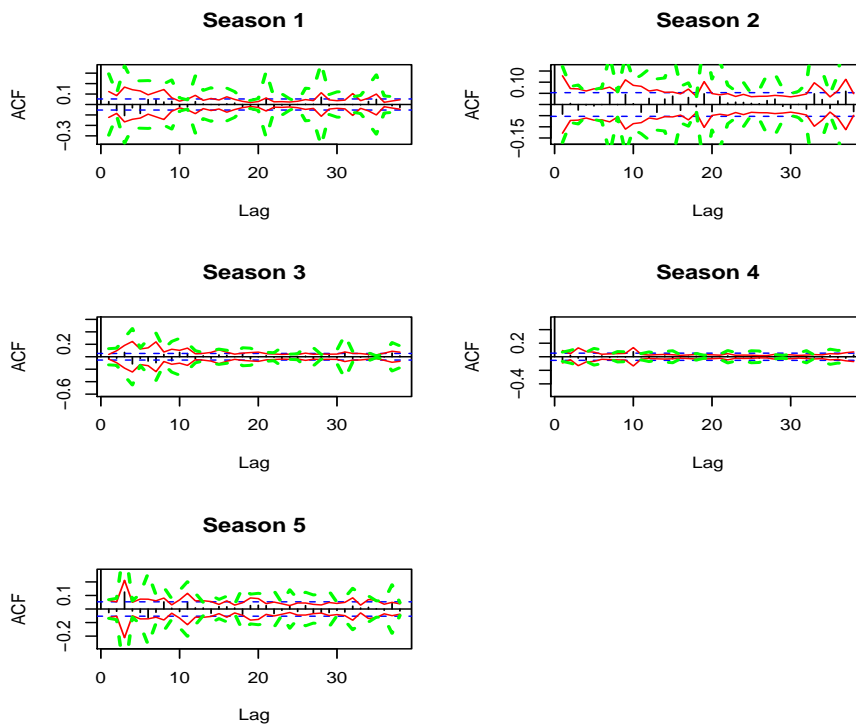


Figure 4.1 – Autocorrelation of the  $\text{PARMA}_5(1, 1)$  residuals for the squares of the CAC 40 returns. The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong  $\text{PARMA}_5(1, 1)$  assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak  $\text{PARMA}_5(1, 1)$  assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 4.3.1. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 4.3.5.



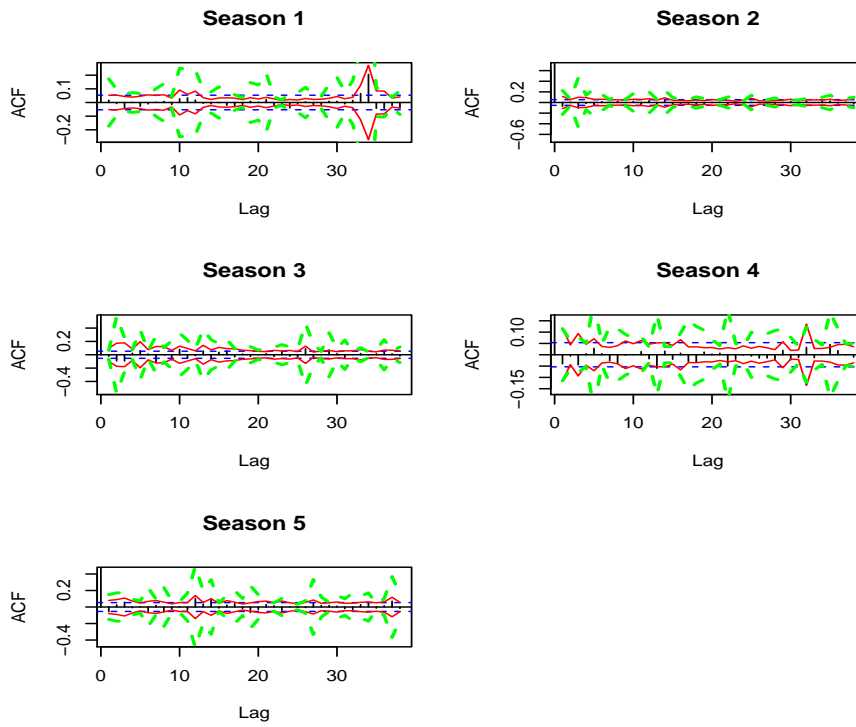


Figure 4.2 – Autocorrelation of the  $\text{PARMA}_5(1,1)$  residuals for the squares of the DAX returns. The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong  $\text{PARMA}_5(1,1)$  assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak  $\text{PARMA}_5(1,1)$  assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 4.3.1. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 4.3.5.

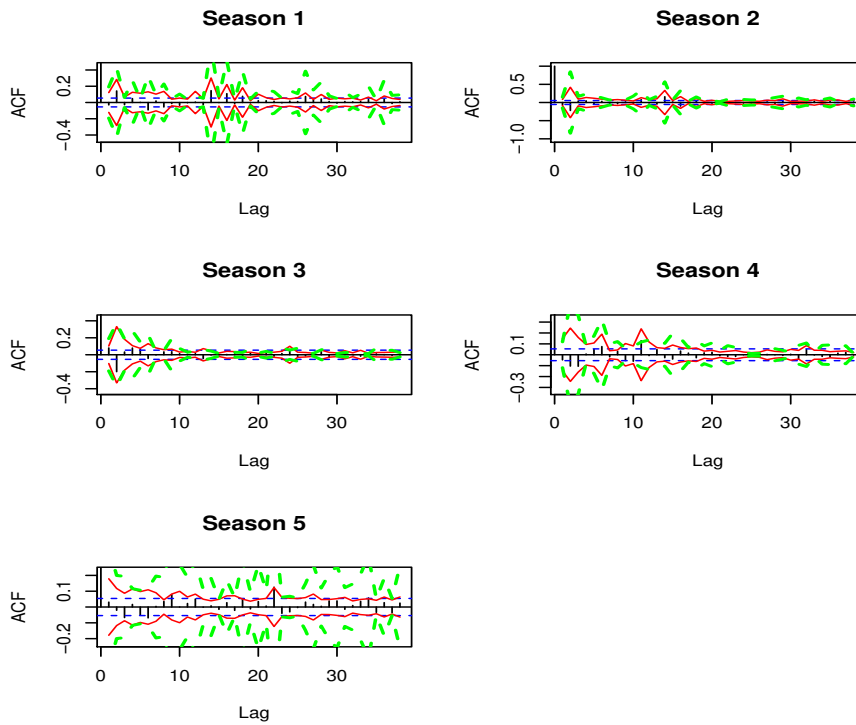


Figure 4.3 – Autocorrelation of the  $\text{PARMA}_5(1, 1)$  residuals for the squares of the SP500 returns. The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong  $\text{PARMA}_5(1, 1)$  assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak  $\text{PARMA}_5(1, 1)$  assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 4.3.1. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 4.3.5.

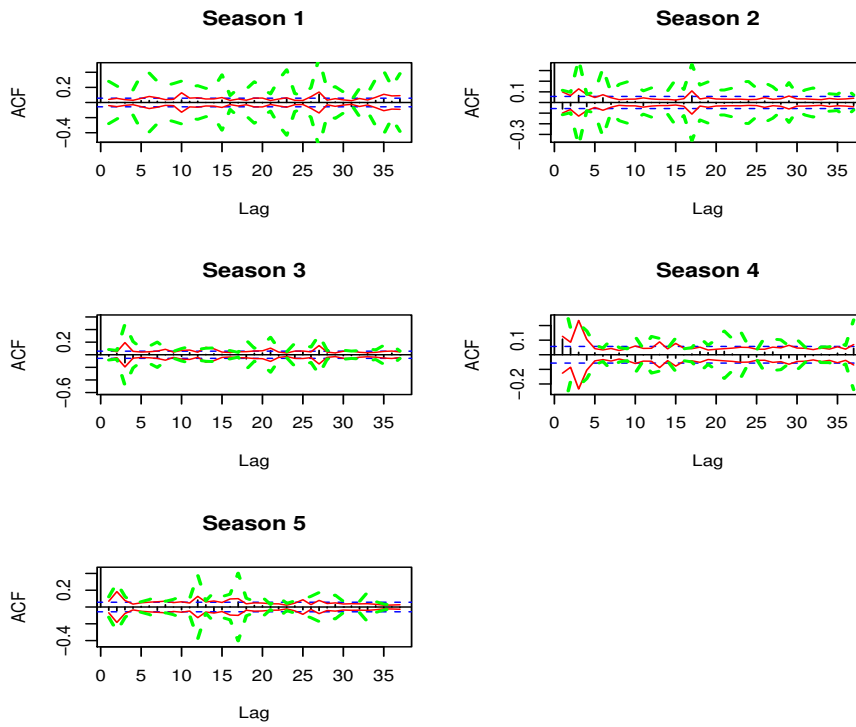


Figure 4.4 – Autocorrelation of the  $\text{PARMA}_5(1,1)$  residuals for the squares of the Nikkei returns. The horizontal dotted lines (blue color) correspond to the 5% significant limits obtained under the strong  $\text{PARMA}_5(1,1)$  assumption. The solid lines (red color) and dashed lines (green color) correspond also to the 5% significant limits under the weak  $\text{PARMA}_5(1,1)$  assumption. The full lines correspond to the asymptotic significance limits for the residual autocorrelations obtained in Theorem 4.3.1. The dashed lines (green color) correspond to the self-normalized asymptotic significance limits for the residual autocorrelations obtained in Theorem 4.3.5.

## 4.6 Proofs

### 4.6.1 Proof of Proposition 4.3.1

First we remark that the asymptotic normality of the joint distribution of  $\sqrt{N}(\hat{\alpha}'_{LS} - \alpha'_0, \gamma'_m(\nu))'$  can be established along the same lines as the proof of Theorem 2 in [20]. The detailed proof is omitted. From (4.14) and (4.21) we have

$$\begin{aligned}\sqrt{N}\begin{pmatrix} \hat{\alpha}_{LS} - \alpha_0 \\ \gamma_m(\nu) \end{pmatrix} &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \begin{pmatrix} -J(\alpha_0, \omega^2)^{-1} \sum_{v=1}^T \omega_v^{-2} \epsilon_{nT+v} \frac{\partial \epsilon_{nT+v}(\alpha_0)}{\partial \alpha} \\ (\epsilon_{nT+v-1}, \dots, \epsilon_{nT+v-m})' \epsilon_{nT+v} \end{pmatrix} + \begin{pmatrix} \mathbf{0}_{\mathbb{P}(1)} \\ \mathbf{0}_m \end{pmatrix} \\ &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} W_{nT+v} + o_{\mathbb{P}}(1),\end{aligned}$$

where  $\mathbf{0}_m$  is the vector of  $\mathbb{R}^{m \times 1}$  with zero components. It is clear that  $W_{nT+v}$  is a measurable function of  $\epsilon_{nT+v}, \epsilon_{nT+v-1}, \dots$ . Thus by using the same arguments as in [20] (see proof of Theorem 2), the central limit theorem (CLT) for strongly mixing processes  $(W_{nT+v})_{n \in \mathbb{Z}}$  of [24] implies that  $(1/\sqrt{N}) \sum_{n=0}^{N-1} W_{nT+v}$  has a limiting normal distribution with mean 0 and covariance matrix  $\Xi(\nu)$ .

In view of (4.21), by applying the CLT for mixing processes we directly obtain

$$\begin{aligned}\Omega^{LS} &= \lim_{N \rightarrow \infty} \text{Var} \left( -\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} J(\alpha_0, \omega^2)^{-1} \sum_{v=1}^T \omega_v^{-2} \epsilon_{nT+v} \frac{\partial \epsilon_{nT+v}(\alpha_0)}{\partial \alpha} \right) \\ &:= J(\alpha_0, \omega^2)^{-1} I(\alpha_0, \omega^2) J(\alpha_0, \omega^2)^{-1},\end{aligned}$$

which gives the first block of the asymptotic covariance matrix of Proposition 4.3.1.

By the stationarity of  $(\epsilon_n)$  and Lebesgue's dominated convergence theorem, we obtain that for  $l, l' \geq 1$

$$\begin{aligned}\text{Cov} \left( \sqrt{N} \gamma(\nu, l), \sqrt{N} \gamma(\nu, l') \right) &= \text{Cov} \left( \frac{1}{\sqrt{N}} \sum_{n=l+1}^N \epsilon_{nT+v} \epsilon_{nT+v-l}, \frac{1}{\sqrt{N}} \sum_{n'=l'+1}^N \epsilon_{n'T+v} \epsilon_{n'T+v-l'} \right) \\ &= \frac{1}{N} \sum_{n=l+1}^N \sum_{n'=l'+1}^N \mathbb{E} (\epsilon_{nT+v} \epsilon_{nT+v-l} \epsilon_{n'T+v} \epsilon_{n'T+v-l'}) \\ &= \frac{1}{N} \sum_{h=-N+1}^{N-1} (N - |h|) \mathbb{E} (\epsilon_{nT+v} \epsilon_{nT+v-l} \epsilon_{(n-h)T+v} \epsilon_{(n-h)T+v-l'}) \\ &\xrightarrow{N \rightarrow \infty} \sum_{h=-\infty}^{\infty} \mathbb{E} (\epsilon_{nT+v} \epsilon_{nT+v-l} \epsilon_{(n-h)T+v} \epsilon_{(n-h)T+v-l'}) := \Gamma_{\nu}(l, l').\end{aligned}$$

We thus have  $\Sigma_{\gamma_m(\nu)} = [\Gamma_{\nu}(l, l')]_{1 \leq l, l' \leq m}$ .

Let  $Y_n = \sum_{v=1}^T \omega_v^{-2} \epsilon_{nT+v} \partial \epsilon_{nT+v}(\alpha_0) / \partial \alpha$ . Finally by the stationarity of  $(\epsilon_n)$  and

$(\epsilon_{nT+v} \partial \epsilon_{nT+v}(\alpha_0) / \partial \alpha)_{n \in \mathbb{Z}}$  we have

$$\begin{aligned}
& \text{Cov} \left( -J(\alpha_0, \omega^2)^{-1} \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} Y_n, \sqrt{N} \Upsilon(v, l) \right) \\
&= \text{Cov} \left( -J(\alpha_0, \omega^2)^{-1} \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \sum_{v=1}^T \omega_v^{-2} \epsilon_{nT+v} \frac{\partial \epsilon_{nT+v}(\alpha_0)}{\partial \alpha}, \frac{1}{\sqrt{N}} \sum_{n'=l}^{N-1} \epsilon_{n'T+v} \epsilon_{n'T+v-l} \right) \\
&= -J(\alpha_0, \omega^2)^{-1} \sum_{v'=1}^T \omega_{v'}^{-2} \frac{1}{N} \sum_{n=0}^{N-1} \sum_{n'=l}^{N-1} \mathbb{E} \left( \epsilon_{nT+v} \frac{\partial \epsilon_{nT+v}(\alpha_0)}{\partial \alpha} \epsilon_{n'T+v} \epsilon_{n'T+v-l} \right) \\
&= -J(\alpha_0, \omega^2)^{-1} \sum_{v=1}^T \omega_v^{-2} \frac{1}{N} \sum_{h=-N+1}^{N-1} (N - |h|) \mathbb{E} \left( \epsilon_{nT+v} \frac{\partial \epsilon_{nT+v}(\alpha_0)}{\partial \alpha} \epsilon_{(n-h)T+v} \epsilon_{(n-h)T+v-l} \right) \\
&\xrightarrow{N \rightarrow \infty} -J(\alpha_0, \omega^2)^{-1} \sum_{v=1}^T \omega_v^{-2} \sum_{h=-\infty}^{\infty} \mathbb{E} \left( \epsilon_{nT+v} \frac{\partial \epsilon_{nT+v}(\alpha_0)}{\partial \alpha} \epsilon_{(n-h)T+v} \epsilon_{(n-h)T+v-l} \right) := \Sigma_{\hat{\alpha}, \Upsilon_m(v)}(\cdot, l),
\end{aligned}$$

by the dominated convergence theorem.

Note that the existence of the above matrices is a consequence of Assumption **(A4)** and of Davydov's inequality (see [17]). The proof is then complete.  $\square$

#### 4.6.2 Proof of Theorem 4.3.1

The joint asymptotic distribution of  $\sqrt{N} \Upsilon_m(v)$  and  $\sqrt{N}(\hat{\alpha}_{LS} - \alpha_0)$  obtained in Proposition 4.3.1 shows that  $\sqrt{N} \hat{\Upsilon}_m(v)$  has a limiting normal distribution with mean zero and covariance matrix

$$\begin{aligned}
\lim_{N \rightarrow \infty} \text{Var} \left( \sqrt{N} \hat{\Upsilon}_m(v) \right) &= \lim_{N \rightarrow \infty} \text{Var} \left( \sqrt{N} \Upsilon_m(v) \right) + \Psi_m(v) \lim_{N \rightarrow \infty} \text{Var} \left( \sqrt{N}(\hat{\alpha}_{LS} - \alpha_0) \right) \Psi_m'(v) \\
&\quad + \Psi_m(v) \lim_{N \rightarrow \infty} \text{Cov} \left( \sqrt{N}(\hat{\alpha}_{LS} - \alpha_0), \sqrt{N} \Upsilon_m(v) \right) \\
&\quad + \lim_{N \rightarrow \infty} \text{Cov} \left( \sqrt{N} \Upsilon_m(v), \sqrt{N}(\hat{\alpha}_{LS} - \alpha_0) \right) \Psi_m'(v) \\
&= \Sigma_{\Upsilon_m(v)} + \Psi_m(v) \Omega^{LS} \Psi_m'(v) + \Psi_m(v) \Sigma_{\hat{\alpha}, \Upsilon_m(v)} + \Sigma'_{\hat{\alpha}, \Upsilon_m(v)} \Psi_m'(v).
\end{aligned}$$

By the stationarity of  $(\epsilon_n)$  and Lebesgue's dominated convergence theorem, we obtain that for  $l, l' \geq 1$

$$\begin{aligned}
\text{Cov} \left( \sqrt{N} \Upsilon(v, l), \sqrt{N} \Upsilon(v', l') \right) &= \text{Cov} \left( \frac{1}{\sqrt{N}} \sum_{n=l+1}^N \epsilon_{nT+v} \epsilon_{nT+v-l}, \frac{1}{\sqrt{N}} \sum_{n'=l'+1}^N \epsilon_{n'T+v'} \epsilon_{n'T+v'-l'} \right) \\
&= \frac{1}{N} \sum_{n=l+1}^N \sum_{n'=l'+1}^N \mathbb{E} \left( \epsilon_{nT+v} \epsilon_{nT+v-l} \epsilon_{n'T+v'} \epsilon_{n'T+v'-l'} \right) \\
&= \frac{1}{N} \sum_{h=-N+1}^{N-1} (N - |h|) \mathbb{E} \left( \epsilon_{nT+v} \epsilon_{nT+v-l} \epsilon_{(n-h)T+v'} \epsilon_{(n-h)T+v'-l'} \right) \\
&\xrightarrow{N \rightarrow \infty} \sum_{h=-\infty}^{\infty} \mathbb{E} \left( \epsilon_{nT+v} \epsilon_{nT+v-l} \epsilon_{(n-h)T+v'} \epsilon_{(n-h)T+v'-l'} \right) := \Gamma_{v, v'}(l, l').
\end{aligned}$$

We thus have  $\Sigma_{\Upsilon_m}(v, v') = [\Gamma_{v, v'}(l, l')]_{1 \leq l, l' \leq m}$ .

Finally by the stationarity of  $(\boldsymbol{\epsilon}_n)$  and  $(\boldsymbol{\epsilon}_{nT+v} \partial \boldsymbol{\epsilon}_{nT+v}(\boldsymbol{\alpha}_0) / \partial \boldsymbol{\alpha})_{n \in \mathbb{Z}}$  we have

$$\begin{aligned}
& \text{Cov}\left(\sqrt{N}(\hat{\boldsymbol{\alpha}}_{\text{LS}} - \boldsymbol{\alpha}_0), \sqrt{N}\boldsymbol{\gamma}(\mathbf{v}', l)\right) \\
&= \text{Cov}\left(-\mathbf{J}(\boldsymbol{\alpha}_0, \omega^2)^{-1} \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \sum_{v=1}^T \omega_v^{-2} \boldsymbol{\epsilon}_{nT+v} \frac{\partial \boldsymbol{\epsilon}_{nT+v}(\boldsymbol{\alpha}_0)}{\partial \boldsymbol{\alpha}}, \frac{1}{\sqrt{N}} \sum_{n'=l}^{N-1} \boldsymbol{\epsilon}_{n'T+v'} \boldsymbol{\epsilon}_{n'T+v'-l}\right) \\
&= -\mathbf{J}(\boldsymbol{\alpha}_0, \omega^2)^{-1} \sum_{v'=1}^T \omega_{v'}^{-2} \frac{1}{N} \sum_{n=0}^{N-1} \sum_{n'=l}^{N-1} \mathbb{E}\left(\boldsymbol{\epsilon}_{nT+v} \frac{\partial \boldsymbol{\epsilon}_{nT+v}(\boldsymbol{\alpha}_0)}{\partial \boldsymbol{\alpha}} \boldsymbol{\epsilon}_{n'T+v'} \boldsymbol{\epsilon}_{n'T+v'-l}\right) \\
&= -\mathbf{J}(\boldsymbol{\alpha}_0, \omega^2)^{-1} \sum_{v=1}^T \omega_v^{-2} \frac{1}{N} \sum_{h=-N+1}^{N-1} (N-|h|) \mathbb{E}\left(\boldsymbol{\epsilon}_{nT+v} \frac{\partial \boldsymbol{\epsilon}_{nT+v}(\boldsymbol{\alpha}_0)}{\partial \boldsymbol{\alpha}} \boldsymbol{\epsilon}_{(n-h)T+v'} \boldsymbol{\epsilon}_{(n-h)T+v'-l}\right) \\
&\xrightarrow{N \rightarrow \infty} -\mathbf{J}(\boldsymbol{\alpha}_0, \omega^2)^{-1} \sum_{v=1}^T \omega_v^{-2} \sum_{h=-\infty}^{\infty} \mathbb{E}\left(\boldsymbol{\epsilon}_{nT+v} \frac{\partial \boldsymbol{\epsilon}_{nT+v}(\boldsymbol{\alpha}_0)}{\partial \boldsymbol{\alpha}} \boldsymbol{\epsilon}_{(n-h)T+v'} \boldsymbol{\epsilon}_{(n-h)T+v'-l}\right) := \boldsymbol{\Sigma}_{\hat{\boldsymbol{\alpha}}, \boldsymbol{\gamma}_m}(\mathbf{v}, \mathbf{v}')(\cdot, l),
\end{aligned}$$

by the dominated convergence theorem. We then deduce that

$$\begin{aligned}
\lim_{N \rightarrow \infty} \text{Cov}\left(\sqrt{N}\hat{\boldsymbol{\gamma}}_m(\mathbf{v}), \sqrt{N}\hat{\boldsymbol{\gamma}}_m(\mathbf{v}')\right) &= \lim_{N \rightarrow \infty} \text{Cov}\left(\sqrt{N}\boldsymbol{\gamma}_m(\mathbf{v}), \sqrt{N}\boldsymbol{\gamma}_m(\mathbf{v}')\right) \\
&\quad + \boldsymbol{\Psi}_m(\mathbf{v}) \lim_{N \rightarrow \infty} \text{Var}\left(\sqrt{N}(\hat{\boldsymbol{\alpha}}_{\text{LS}} - \boldsymbol{\alpha}_0)\right) \boldsymbol{\Psi}'_m(\mathbf{v}') \\
&\quad + \boldsymbol{\Psi}_m(\mathbf{v}) \lim_{N \rightarrow \infty} \text{Cov}\left(\sqrt{N}(\hat{\boldsymbol{\alpha}}_{\text{LS}} - \boldsymbol{\alpha}_0), \sqrt{N}\boldsymbol{\gamma}_m(\mathbf{v}')\right) \\
&\quad + \lim_{N \rightarrow \infty} \text{Cov}\left(\sqrt{N}\boldsymbol{\gamma}_m(\mathbf{v}), \sqrt{N}(\hat{\boldsymbol{\alpha}}_{\text{LS}} - \boldsymbol{\alpha}_0)\right) \boldsymbol{\Psi}'_m(\mathbf{v}') \\
&= \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_m}(\mathbf{v}, \mathbf{v}') + \boldsymbol{\Psi}_m(\mathbf{v}) \boldsymbol{\Omega}^{\text{LS}} \boldsymbol{\Psi}'_m(\mathbf{v}') + \boldsymbol{\Psi}_m(\mathbf{v}) \boldsymbol{\Sigma}_{\hat{\boldsymbol{\alpha}}, \boldsymbol{\gamma}_m}(\mathbf{v}, \mathbf{v}') \\
&\quad + \boldsymbol{\Sigma}'_{\hat{\boldsymbol{\alpha}}, \boldsymbol{\gamma}_m(\mathbf{v})} \boldsymbol{\Psi}'_m(\mathbf{v}') \\
&=: \nabla_{\hat{\boldsymbol{\gamma}}}(\mathbf{v}, \mathbf{v}').
\end{aligned}$$

We come back to the vector  $\hat{\boldsymbol{\rho}}_m(\mathbf{v})$  and observe that from (4.17), we have  $\sqrt{N}(\hat{\boldsymbol{\gamma}}(\mathbf{v}, 0) - \boldsymbol{\gamma}(\mathbf{v}, 0)) = o_{\mathbb{P}}(1)$ . Applying the CLT for mixing processes (see [24]) to the process  $(\boldsymbol{\epsilon}_t^2)_t$ , we obtain

$$\sqrt{N}(\hat{\sigma}_v^2 - \sigma_v^2) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} (\boldsymbol{\epsilon}_{nT+v}^2 - \mathbb{E}[\boldsymbol{\epsilon}_{nT+v}^2]) + o_{\mathbb{P}}(1) \xrightarrow[N \rightarrow \infty]{\text{in law}} \mathcal{N}\left(0, \sum_{h=-\infty}^{\infty} \text{Cov}(\boldsymbol{\epsilon}_v^2, \boldsymbol{\epsilon}_{hT+v}^2)\right).$$

So we have  $\sqrt{N}(\hat{\sigma}_v^2 - \sigma_v^2) = o_{\mathbb{P}}(1)$  and  $\sqrt{N}(\boldsymbol{\gamma}(\mathbf{v}, 0) - \sigma_v^2) = o_{\mathbb{P}}(1)$ . Using the  $\delta$ -method we deduce that  $\sqrt{N}(\sqrt{\hat{\sigma}_v^2} - \sqrt{\sigma_v^2}) = o_{\mathbb{P}}(1)$  and  $\sqrt{N}(\sqrt{\boldsymbol{\gamma}(\mathbf{v}, 0)} - \sqrt{\sigma_v^2}) = o_{\mathbb{P}}(1)$ . Now using (4.23) and the ergodic theorem we obtain

$$\mathbf{N}\left(\frac{\hat{\boldsymbol{\gamma}}(\mathbf{v}, h)}{\sqrt{\hat{\boldsymbol{\gamma}}(\mathbf{v}, 0)\hat{\boldsymbol{\gamma}}(\mathbf{v}-h, 0)}} - \frac{\hat{\boldsymbol{\gamma}}(\mathbf{v}, h)}{\sigma_v \sigma_{v-h}}\right) = \sqrt{N}\hat{\boldsymbol{\gamma}}(\mathbf{v}, h) \frac{\sqrt{N}(\sigma_v \sigma_{v-h} - \sqrt{\hat{\boldsymbol{\gamma}}(\mathbf{v}, 0)\hat{\boldsymbol{\gamma}}(\mathbf{v}-h, 0)})}{\sigma_v \sigma_{v-h} \hat{\boldsymbol{\gamma}}(\mathbf{v}, 0)} = o_{\mathbb{P}}(1),$$

which means  $\sqrt{N}\hat{\boldsymbol{\rho}}(\mathbf{v}, h) = \sqrt{N}\hat{\boldsymbol{\gamma}}(\mathbf{v}, h) / (\sigma_v \sigma_{v-h}) + o_{\mathbb{P}}(N^{-1/2})$ . For  $h = 1, \dots, m$ , it follows that

$$\sqrt{N}\hat{\boldsymbol{\rho}}_m(\mathbf{v}) = \mathbf{D}_m^{-1}(\mathbf{v}) \left(\sqrt{N}\hat{\boldsymbol{\gamma}}(\mathbf{v}, 1), \dots, \sqrt{N}\hat{\boldsymbol{\gamma}}(\mathbf{v}, m)\right)' + o_{\mathbb{P}}(1) = \sqrt{N}\mathbf{D}_m^{-1}(\mathbf{v})\hat{\boldsymbol{\gamma}}_m(\mathbf{v}) + o_{\mathbb{P}}(1), \quad (4.38)$$

where  $\mathbf{D}_m(\mathbf{v}) = \text{Diag}(\sigma_v \sigma_{v-1}, \dots, \sigma_v \sigma_{v-m})$ . Thus from (4.38) the asymptotic distribution of the residual autocorrelations  $\sqrt{N}\hat{\boldsymbol{\rho}}_m(\mathbf{v})$  depends on the distribution of  $\sqrt{N}\hat{\boldsymbol{\gamma}}_m(\mathbf{v})$ . Consequently we have

$$\lim_{N \rightarrow \infty} \text{Var}\left(\sqrt{N}\hat{\boldsymbol{\rho}}_m(\mathbf{v})\right) = \lim_{N \rightarrow \infty} \text{Var}\left(\mathbf{D}_m^{-1}(\mathbf{v})\sqrt{N}\hat{\boldsymbol{\gamma}}_m(\mathbf{v})\right) = \mathbf{D}_m^{-1}(\mathbf{v}) \nabla_{\hat{\boldsymbol{\gamma}}}(\mathbf{v}, \mathbf{v}) \mathbf{D}_m^{-1}(\mathbf{v}).$$

Observe that from (4.38) we deduce

$$\begin{aligned}\sqrt{N}\hat{\rho}_{mT} &= \left( D_m^{-1}(\mathbf{v})\sqrt{N}\hat{\gamma}_m(1), \dots, D_m^{-1}(\mathbf{v})\sqrt{N}\hat{\gamma}_m(T) \right)' + o_{\mathbb{P}}(1) \\ &= D_{mT}^{-1}\sqrt{N}\hat{\gamma}_{mT} + o_{\mathbb{P}}(1),\end{aligned}\quad (4.39)$$

where  $D_{mT} = \text{Diag}(\sigma_1\sigma_{1-1}, \dots, \sigma_1\sigma_{1-m}, \dots, \sigma_T\sigma_{T-1}, \dots, \sigma_T\sigma_{T-m})$ .

Thus from (4.39) the asymptotic distribution of the residual autocorrelations  $\sqrt{N}\hat{\rho}_{mT}$  depends on the distribution of  $\sqrt{N}\hat{\gamma}_{mT}$ . Consequently we have

$$\lim_{N \rightarrow \infty} \text{Var} \left( \sqrt{N}\hat{\rho}_m \right) = \lim_{N \rightarrow \infty} \text{Var} \left( D_{mT}^{-1}\sqrt{N}\hat{\gamma}_{mT} \right) = D_{mT}^{-1} \nabla_{\hat{\gamma}} D_{mT}^{-1} =: \nabla_{\hat{\rho}}.$$

The proof is completed.  $\square$

### 4.6.3 Proof of Proposition 4.3.2

The following proofs are quite technical and are adaptations of the arguments used in [13].

To prove the invertibility of the normalized matrix  $C_m(\mathbf{v}, \mathbf{v})$  we need to introduce the following notation. Let  $S_n^i(\mathbf{v})$  be the  $i$ -th component of the vector  $S_n(\mathbf{v}) = \sum_{j=0}^n (\Lambda(\mathbf{v})W_{jT+\mathbf{v}} - \gamma_m(\mathbf{v})) \in \mathbb{R}^m$ . We remark that

$$S_{n-1}^i(\mathbf{v}) = S_n^i(\mathbf{v}) - \sum_{k=1}^{(p+q)T} \delta_{i,k}(\mathbf{v}) \omega_{\mathbf{v}}^{-2} \epsilon_{nT+\mathbf{v}} \frac{\partial}{\partial \alpha_k} \epsilon_{nT+\mathbf{v}}(\alpha_0) - \epsilon_{nT+\mathbf{v}} \epsilon_{nT+\mathbf{v}-i} + \gamma(\mathbf{v}, i), \quad (4.40)$$

where  $\delta_{i,k}(\mathbf{v})$  is the  $(i, k)$ -th entry of the  $m \times (p+q)T$  matrix  $G(\mathbf{v}) := -\Psi_m(\mathbf{v})J(\alpha_0, \omega^2)^{-1}$ .

If the matrix  $C_m(\mathbf{v}, \mathbf{v})$  is not invertible, there exists some real constants  $c_1, \dots, c_m$  not all equal to zero, such that we have

$$\sum_{i=1}^m \sum_{j=1}^m c_j [C_m(\mathbf{v}, \mathbf{v})]_{ij} c_i = \frac{1}{N^2} \sum_{n=0}^{N-1} \sum_{i=1}^m \sum_{j=1}^m c_j S_n^j(\mathbf{v}) S_n^i(\mathbf{v}) c_i = \frac{1}{N^2} \sum_{n=0}^{N-1} \left( \sum_{i=1}^m c_i S_n^i(\mathbf{v}) \right)^2 = 0,$$

which implies that  $\sum_{i=1}^m c_i S_n^i(\mathbf{v}) = 0$  for all  $n \geq 0$ .

Then by (4.40) it would imply that

$$\sum_{i=1}^m \sum_{k=1}^{(p+q)T} c_i \delta_{i,k}(\mathbf{v}) \omega_{\mathbf{v}}^{-2} \epsilon_{nT+\mathbf{v}} \frac{\partial}{\partial \alpha_k} \epsilon_{nT+\mathbf{v}}(\alpha_0) + \sum_{i=1}^m c_i \epsilon_{nT+\mathbf{v}} \epsilon_{nT+\mathbf{v}-i} = \sum_{i=1}^m c_i \gamma(\mathbf{v}, i). \quad (4.41)$$

By the ergodic Theorem, we also have  $\sum_{i=1}^m c_i \gamma(\mathbf{v}, i) \rightarrow 0$  almost-surely as  $N$  goes to infinity.

Consequently replacing this convergence in (4.41) implies that for all  $n \geq 0$

$$\sum_{i=1}^m \sum_{k=1}^{(p+q)T} c_i \sum_{\mathbf{v}=1}^T \delta_{i,k}(\mathbf{v}) \omega_{\mathbf{v}}^{-2} \epsilon_{nT+\mathbf{v}} \frac{\partial}{\partial \alpha_k} \epsilon_{nT+\mathbf{v}}(\alpha_0) + \sum_{i=1}^m c_i \epsilon_{nT+\mathbf{v}} \epsilon_{nT+\mathbf{v}-i} = 0, \quad \text{a.s.}$$

Note that from (4.6) and under (A1) we have

$$\epsilon_n(\alpha_0) = \Phi^{-1}(\mathbf{L})\Theta(\mathbf{L})\mathbf{X}_n. \quad (4.42)$$

In view of (4.42) we deduce

$$\frac{\partial \epsilon_n(\alpha_0)}{\partial \alpha_i} = \sum_{\ell=1}^{\infty} d_{\ell}' \epsilon_{n-\ell}(\alpha_0), \quad (4.43)$$



where the sequence of matrices  $d_\ell = d_\ell(i)$  is such that  $\|d_\ell\| \rightarrow 0$  at a geometric rate as  $\ell \rightarrow \infty$  (see for instance [20, Lemma 7.11]). Using (4.43), the  $v$ -th component of  $\partial \epsilon_n(\alpha_0)/\partial \alpha_i$  is of the form:

$$\frac{\partial}{\partial \alpha_i} \epsilon_{nT+v}(\alpha_0) = \sum_{\ell=1}^{\infty} \sum_{j=1}^T d'_\ell(v, j) \epsilon_{nT+v-\ell}(\alpha_0) \quad (4.44)$$

where  $|d'_\ell(v, j)| = |d'_\ell(i, v, j)| \leq K\tau^\ell$  uniformly in  $(i, v, j) \in \{1, \dots, (p+q)T\} \times \{1, \dots, T\}^2$  and  $\tau$  is a constant belonging to  $[0, 1)$  with  $K > 0$ . In view of (4.44) it yields that

$$\epsilon_{nT+v} \left\{ \sum_{\ell \geq 1} \left( \sum_{i=1}^m \sum_{k=1}^{(p+q)T} c_i \delta_{i,k}(v) \sum_{j=1}^T d'_\ell(v, j) \epsilon_{nT+v-\ell} \omega_v^{-2} \right) \epsilon_{nT+v-\ell} + \sum_{\ell=1}^m c_\ell \epsilon_{nT+v-\ell} \right\} = 0, \text{ a.s.}$$

Or equivalently,

$$\begin{aligned} \epsilon_{nT+v} \left\{ \sum_{\ell=1}^m \left( \sum_{i=1}^m \sum_{k=1}^{(p+q)T} c_i \delta_{i,k}(v) \sum_{j=1}^T d'_\ell(v, j) \omega_v^{-2} + c_\ell \right) \epsilon_{nT+v-\ell} \right. \\ \left. + \sum_{\ell \geq m+1} \left( \sum_{i=1}^m \sum_{k=1}^{(p+q)T} c_i \delta_{i,k}(v) \sum_{j=1}^T d'_\ell(v, j) \omega_v^{-2} \right) \epsilon_{nT+v-\ell} \right\} = 0, \text{ a.s.} \end{aligned}$$

Thanks to Assumption (A5),  $(\epsilon_n)_{n \in \mathbb{Z}}$  has a positive density in some neighborhood of zero and then  $\epsilon_{nT+v} \neq 0$  almost-surely. Hence we obtain

$$\begin{aligned} \sum_{\ell=1}^m \left( \sum_{i=1}^m \sum_{k=1}^{(p+q)T} c_i \sum_{v=1}^T \delta_{i,k}(v) \sum_{j=1}^T d'_\ell(v, j) \omega_v^{-2} + c_\ell \right) \epsilon_{nT+v-\ell} \\ + \sum_{\ell \geq m+1} \left( \sum_{i=1}^m \sum_{k=1}^{(p+q)T} c_i \sum_{v=1}^T \delta_{i,k}(v) \sum_{j=1}^T d'_\ell(v, j) \omega_v^{-2} \right) \epsilon_{nT+v-\ell} = 0, \text{ a.s.} \end{aligned}$$

Since the variance of the linear periodic innovation process is not equal to zero, we deduce that

$$\begin{cases} \sum_{i=1}^m \sum_{k=1}^{(p+q)T} c_i \delta_{i,k}(v) \sum_{j=1}^T d'_\ell(v, j) \omega_v^{-2} + c_\ell = 0 & \text{for all } \ell \in \{1, \dots, m\} \\ \sum_{i=1}^m \sum_{k=1}^{(p+q)T} c_i \delta_{i,k}(v) \sum_{j=1}^T d'_\ell(v, j) \omega_v^{-2} = 0 & \text{for all } \ell \in \{m+1, \dots\}. \end{cases}$$

Then we would have  $c_1 = \dots = c_m = 0$  which is impossible. Thus we have a contradiction and the matrix  $C_m(v, v) \in \mathbb{R}^{m \times m}$  is non singular.

The proof of the matrix  $C_{mT} \in \mathbb{R}^{mT \times mT}$  follows the same lines that of  $C_m(v, v)$ .  $\square$

The proofs of Theorems 4.3.4 and 4.3.5 are similar to that given by [13] in the weak ARMA case.

#### 4.6.4 Proof of Theorem 4.3.4

We recall that the Skorokhod space  $\mathbb{D}^\ell[0,1]$  is the set of  $\mathbb{R}^\ell$ -valued functions on  $[0,1]$  which are right-continuous and has left limits everywhere. It is endowed with the Skorokhod topology and the weak convergence on  $\mathbb{D}^\ell[0,1]$  is mentioned by  $\xrightarrow{\mathbb{D}^\ell}$ .

The proof is divided in two steps.

**Functional central limit theorem for  $(\Lambda(\nu)W_{nT+\nu})_{n \in \mathbb{Z}}$** 

In view of (4.19) and (4.22) we deduce that

$$\begin{aligned} \sqrt{N}\hat{\gamma}_m(\nu) &= \sqrt{N}\gamma_m(\nu) + \sqrt{N}\Psi_m(\nu)(\hat{\alpha}_{LS} - \alpha_0) + o_{\mathbb{P}}(1) \\ &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} W_{2,nT+\nu} + \Psi_m(\nu) \left( \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} W_{1,nT+\nu} + o_{\mathbb{P}}(1) \right) + o_{\mathbb{P}}(1) \\ &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \Lambda(\nu)W_{nT+\nu} + o_{\mathbb{P}}(1). \end{aligned} \quad (4.45)$$

Now it is clear that the asymptotic behaviour of  $\hat{\gamma}_m(\nu)$  is related to the limit distribution of  $(W_{nT+\nu})_{n \in \mathbb{Z}} = (W'_{1,nT+\nu}, W'_{2,nT+\nu})'_{n \in \mathbb{Z}}$ . Our first goal is to show that there exists a lower triangular matrix  $\Pi(\nu)$  with nonnegative diagonal entries such that

$$\frac{1}{\sqrt{N}} \sum_{n=0}^{\lfloor Nr \rfloor - 1} \Lambda(\nu)W_{nT+\nu} \xrightarrow[N \rightarrow \infty]{\mathbb{D}^m} \left( \Pi(\nu)\Pi'(\nu) \right)^{1/2} B_m(r), \quad (4.46)$$

where  $(B_m(r))_{r \geq 0}$  is a  $m$ -dimensional standard Brownian motion. Let  $t = nT + \nu$  and using (4.44),  $W_t$  can be rewritten as

$$W_t = \left( - \left\{ \sum_{\ell=1}^{\infty} \sum_{j=1}^T d'_{\ell}(1, \nu, j) \epsilon_t \epsilon_{t-i}, \dots, \sum_{\ell=1}^{\infty} \sum_{j=1}^T d'_{\ell}((p+q)T, \nu, j) \epsilon_t \epsilon_{t-i} \right\}' J(\alpha_0, \omega^2)^{-1}', \epsilon_t \epsilon_{t-1}, \dots, \epsilon_t \epsilon_{t-m} \right)'.$$

The non-correlation between  $\epsilon_t$ 's implies that the process  $(W_t)_t$  of  $\mathbb{R}^{(p+q)T+m}$  is centered. In order to apply the functional central limit theorem for strongly mixing process (see [24]), we need to identify the asymptotic covariance matrix in the classical central limit theorem for the sequence  $(W_t)_t$ . It is proved in Proposition 4.3.1 that

$$\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} W_{nT+\nu} \xrightarrow[N \rightarrow \infty]{\text{in law}} \mathcal{N}(0, \Xi(\nu) := 2\pi f_W(0)),$$

where  $f_W(0)$  is the spectral density of the stationary process  $(W_{nT+\nu})_{n \in \mathbb{Z}}$  evaluated at frequency 0. The existence of the matrix  $\Xi(\nu)$  is a consequence of Assumption **(A4)** and of Davydov's inequality (see [17]). Since the matrix  $\Xi(\nu)$  is positive definite it can be factorized as  $\Xi(\nu) = Y(\nu)Y'(\nu)$ , where the  $(pT+qT+m) \times (pT+qT+m)$  lower triangular matrix  $Y(\nu)$  has nonnegative diagonal entries. Therefore we have

$$\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \Lambda(\nu)W_{nT+\nu} \xrightarrow[n \rightarrow \infty]{\text{in law}} \mathcal{N}\left(0, \Lambda(\nu)\Xi(\nu)\Lambda'(\nu)\right),$$

and the new variance matrix can also be factorized as

$$\Lambda(\nu)\Xi(\nu)\Lambda'(\nu) = (\Lambda(\nu)Y(\nu))(\Lambda(\nu)Y(\nu))' := \Pi(\nu)\Pi'(\nu),$$

where  $\Pi(\nu) \in \mathbb{R}^{m \times (pT+qT)}$ . Thus

$$\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} (\Pi(\nu)\Pi'(\nu))^{-1/2} \Lambda(\nu)W_{nT+\nu} \xrightarrow[N \rightarrow \infty]{\text{in law}} \mathcal{N}(0, I_m),$$

where  $(\Pi(\nu)\Pi'(\nu))^{-1/2}$  is the Moore-Penrose inverse (see [36], p. 36) of  $(\Pi(\nu)\Pi'(\nu))^{1/2}$ .

Using the same arguments as in the proof of Theorem 2 in [20] the asymptotic distribution of  $N^{-1/2} \sum_{n=0}^{N-1} W_{nT+v}$  when  $N$  tends to infinity is obtained by introducing the random vector  $W_t^k := W_{nT+v}^k$  defined for any strictly positive integer  $k$  by

$$W_t^k = \left( - \left\{ \sum_{\ell=1}^k \sum_{j=1}^T d'_\ell(1, v, j) \epsilon_t \epsilon_{t-i}, \dots, \sum_{\ell=1}^k \sum_{j=1}^T d'_\ell((p+q)T, v, j) \epsilon_t \epsilon_{t-i} \right\}' J(\alpha_0, \omega^2)^{-1'}, \epsilon_t \epsilon_{t-1}, \dots, \epsilon_t \epsilon_{t-m} \right)'.$$

Since  $W_{nT+v}^k$  depends on a finite number of values of the periodic noise process  $(\epsilon_{nT+v})_{n \in \mathbb{Z}}$ . It also satisfies a mixing property (see Theorem 14.1 in [16], p. 210). Then applying the central limit theorem for strongly mixing process of [24] shows that its asymptotic distribution is normal with zero mean and variance matrix  $\Xi_k(v, v)$  that converges when  $k$  tends to infinity to  $\Xi(v)$ . More precisely we have

$$\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} W_{nT+v}^k \xrightarrow[N \rightarrow \infty]{\text{in law}} \mathcal{N}(0, \Xi_k(v)).$$

The above arguments also apply to matrix  $\Xi_k(v)$  with some matrix  $\Pi_k(v)$  which is defined analogously as  $\Pi(v)$ . Consequently we obtain

$$\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \Lambda(v) W_{nT+v}^k \xrightarrow[n \rightarrow \infty]{\text{in law}} \mathcal{N}\left(0, \Lambda(v) \Xi_k(v) \Lambda'(v)\right)$$

and we also have  $N^{-1/2} \sum_{n=0}^{N-1} (\Pi_k(v) \Pi'_k(v))^{-1/2} \Lambda(v) W_{nT+v}^k \xrightarrow[N \rightarrow \infty]{\text{in law}} \mathcal{N}(0, I_m)$ .

Now we are able to apply the functional central limit theorem (see [24]) and we obtain that

$$\frac{1}{\sqrt{N}} \sum_{n=0}^{\lfloor Nr \rfloor - 1} (\Pi_k(v) \Pi'_k(v))^{-1/2} \Lambda(v) W_{nT+v}^k \xrightarrow[N \rightarrow \infty]{\mathbb{D}^m} B_m(r).$$

Since for all  $n \in \{0, \dots, \lfloor nr \rfloor - 1\}$  we write

$$\begin{aligned} (\Pi(v) \Pi'(v))^{-1/2} \Lambda(v) W_{nT+v}^k &= \left( (\Pi(v) \Pi'(v))^{-1/2} - (\Pi_k(v) \Pi'_k(v))^{-1/2} \right) \Lambda(v) W_{nT+v}^k \\ &\quad + (\Pi_k(v) \Pi'_k(v))^{-1/2} \Lambda(v) W_{nT+v}^k, \end{aligned}$$

we obtain the following weak convergence on  $\mathbb{D}^m[0, 1]$ :

$$\frac{1}{\sqrt{N}} \sum_{n=0}^{\lfloor Nr \rfloor - 1} (\Pi(v) \Pi'(v))^{-1/2} \Lambda(v) W_{nT+v}^k \xrightarrow[N \rightarrow \infty]{\mathbb{D}^m} B_m(r).$$

In order to conclude that (4.46) is true, it remains to observe that uniformly with respect to  $N$

$$Y_N^k(r) := \frac{1}{\sqrt{N}} \sum_{n=0}^{\lfloor Nr \rfloor - 1} (\Pi(v) \Pi'(v))^{-1/2} \Lambda(v) Z_{nT+v}^k \xrightarrow[k \rightarrow \infty]{\mathbb{D}^m} 0, \quad (4.47)$$

where the random vector  $Z_t^k := Z_{nT+v}^k$  is defined by

$$Z_t^k = \left( - \left\{ \sum_{\ell=k+1}^{\infty} \sum_{j=1}^T d'_\ell(1, v, j) \epsilon_t \epsilon_{t-i}, \dots, \sum_{\ell=k+1}^{\infty} \sum_{j=1}^T d'_\ell((p+q)T, v, j) \epsilon_t \epsilon_{t-i} \right\}' J(\alpha_0, \omega^2)^{-1'}, \epsilon_t \epsilon_{t-1}, \dots, \epsilon_t \epsilon_{t-m} \right)'.$$

Using the same arguments as those used in the proof of Theorem 2 in [20] we obtain

$$\sup_N \text{Var} \left( \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} Z_{nT+v}^k \right) \xrightarrow[k \rightarrow \infty]{} 0$$

and since  $\lfloor Nr \rfloor \leq N$  we have

$$\sup_{0 \leq r \leq 1} \sup_N \left\{ \left\| Y_N^k(r) \right\| \right\} \xrightarrow[k \rightarrow \infty]{} 0.$$

Thus (4.47) is true and the proof of (4.46) is achieved.

### Limit theorem

To conclude the prove of Theorem 4.3.4, we follow the arguments developed in [13]. Note that the previous step ensures us that Assumption 1 in [33] is satisfied for the sequence  $(\Lambda(\mathbf{v})W_{nT+\mathbf{v}})_{n \geq 0}$ . Firstly from (4.46) we deduce that

$$\begin{aligned} \frac{1}{\sqrt{N}} S_{\lfloor Nr \rfloor}(\mathbf{v}) &= \frac{1}{\sqrt{N}} \sum_{n=0}^{\lfloor Nr \rfloor - 1} \Lambda(\mathbf{v}) W_{nT+\mathbf{v}} - \frac{\lfloor Nr \rfloor}{N} \left( \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \Lambda(\mathbf{v}) W_{nT+\mathbf{v}} \right) \\ &\xrightarrow[N \rightarrow \infty]{\mathbb{D}^m} (\Pi(\mathbf{v})\Pi'(\mathbf{v}))^{1/2} \mathbf{B}_m(r) - r(\Pi(\mathbf{v})\Pi'(\mathbf{v}))^{1/2} \mathbf{B}_m(1). \end{aligned} \quad (4.48)$$

Observe now that the continuous mapping theorem implies

$$\begin{aligned} C_m(\mathbf{v}, \mathbf{v}) &= \frac{1}{N} \sum_{n=0}^{N-1} \left( \frac{1}{\sqrt{N}} S_n \right) \left( \frac{1}{\sqrt{N}} S_n \right)' \\ &\xrightarrow[N \rightarrow \infty]{\mathbb{D}^m} (\Pi(\mathbf{v})\Pi'(\mathbf{v}))^{1/2} \left[ \int_0^1 \{ \mathbf{B}_m(r) - r\mathbf{B}_m(1) \} \{ \mathbf{B}_m(r) - r\mathbf{B}_m(1) \}' dr \right] (\Pi(\mathbf{v})\Pi'(\mathbf{v}))^{1/2} \\ &= (\Pi(\mathbf{v})\Pi'(\mathbf{v}))^{1/2} \mathbf{V}_m (\Pi(\mathbf{v})\Pi'(\mathbf{v}))^{1/2}. \end{aligned}$$

Using (4.45), (4.48) and again the continuous mapping theorem on the Skorokhod space, one finally obtains

$$\begin{aligned} N\hat{\Upsilon}'_m(\mathbf{v})C_m^{-1}(\mathbf{v}, \mathbf{v})\hat{\Upsilon}_m(\mathbf{v}) &\xrightarrow[N \rightarrow \infty]{\mathbb{D}^m} \left\{ (\Pi(\mathbf{v})\Pi'(\mathbf{v}))^{1/2} \mathbf{B}_m(1) \right\}' \left\{ (\Pi(\mathbf{v})\Pi'(\mathbf{v}))^{1/2} \mathbf{V}_m (\Pi(\mathbf{v})\Pi'(\mathbf{v}))^{1/2} \right\}^{-1} \\ &\quad \times \left\{ (\Pi(\mathbf{v})\Pi'(\mathbf{v}))^{1/2} \mathbf{B}_m(1) \right\} \\ &= \mathbf{B}'_m(1) \mathbf{V}_m^{-1} \mathbf{B}_m(1) := \mathcal{U}_m. \end{aligned}$$

Consequently from (4.38) it follows that

$$N\hat{\rho}'_m(\mathbf{v})\mathbf{D}_m(\mathbf{v})C_m^{-1}(\mathbf{v}, \mathbf{v})\mathbf{D}_m(\mathbf{v})\hat{\rho}_m(\mathbf{v}) \xrightarrow[N \rightarrow \infty]{\mathbb{D}^m} \mathcal{U}_m,$$

which completes the proof of the convergence of  $\sqrt{N}\hat{\rho}_m(\mathbf{v})$ .

The proof of the convergence of  $\sqrt{N}\hat{\rho}_{mT}$  follows the same lines that of  $\sqrt{N}\hat{\rho}_m(\mathbf{v})$  which completes the proof of Theorem 4.3.4.  $\square$

### 4.6.5 Proof of Theorem 4.3.5

We write  $\hat{C}_m(\mathbf{v}, \mathbf{v}) = C_m(\mathbf{v}, \mathbf{v}) + \Upsilon_N(\mathbf{v}, \mathbf{v})$  where  $\Upsilon_N(\mathbf{v}, \mathbf{v}) = N^{-2} \sum_{n=0}^{N-1} (S_n(\mathbf{v})S'_n(\mathbf{v}) - \hat{S}_n(\mathbf{v})\hat{S}'_n(\mathbf{v}))$ . Now observe that there are three kinds of entries in the matrix  $\Upsilon_N$ . The first one is a sum composed of

$$\mathbf{v}_{nT+\mathbf{v}}^{k,k'} = \epsilon_{nT+\mathbf{v}}^2(\alpha_0) \epsilon_{nT+\mathbf{v}-k}(\alpha_0) \epsilon_{nT+\mathbf{v}-k'}(\alpha_0) - e_{nT+\mathbf{v}}^2(\hat{\alpha}_{LS}) e_{nT+\mathbf{v}-k}(\hat{\alpha}_{LS}) e_{nT+\mathbf{v}-k'}(\hat{\alpha}_{LS})$$

for  $(k, k') \in \{1, \dots, m\}^2$ . By [20, Lemma 7.11] and the consistency of  $\hat{\alpha}_{LS}$ , we have  $\mathbf{v}_{nT+\mathbf{v}}^{k,k'} = o_{\mathbb{P}}(1)$  almost surely. The two last kinds of entries of  $\Upsilon_N(\mathbf{v}, \mathbf{v})$  come from the following quantities for  $i, j \in \{1, \dots, T(p+q)\}$  and  $k \in \{1, \dots, m\}$

$$\begin{aligned} \tilde{\mathbf{v}}_{nT+\mathbf{v}}^{k,i} &= \epsilon_{nT+\mathbf{v}}^2(\alpha_0) \epsilon_{nT+\mathbf{v}-k}(\alpha_0) \frac{\partial \epsilon_{nT+\mathbf{v}}(\alpha_0)}{\partial \alpha_i} - e_{nT+\mathbf{v}}^2(\hat{\alpha}_{LS}) e_{nT+\mathbf{v}-k}(\hat{\alpha}_{LS}) \frac{\partial e_{nT+\mathbf{v}}(\hat{\alpha}_{LS})}{\partial \alpha_i}, \\ \tilde{\mathbf{v}}_{nT+\mathbf{v}}^{i,j} &= \epsilon_{nT+\mathbf{v}}^2(\alpha_0) \frac{\partial \epsilon_{nT+\mathbf{v}}(\alpha_0)}{\partial \alpha_i} \frac{\partial \epsilon_{nT+\mathbf{v}}(\alpha_0)}{\partial \alpha_j} - e_{nT+\mathbf{v}}^2(\hat{\alpha}_{LS}) \frac{\partial e_{nT+\mathbf{v}}(\hat{\alpha}_{LS})}{\partial \alpha_i} \frac{\partial e_{nT+\mathbf{v}}(\hat{\alpha}_{LS})}{\partial \alpha_j} \end{aligned}$$

and they also satisfy  $\tilde{v}_{nT+v}^{k,i} + \tilde{v}_{nT+v}^{i,j} = o_{\mathbb{P}}(1)$  almost surely. Consequently,  $Y_N(\mathbf{v}, \mathbf{v}) = o_{\mathbb{P}}(1)$  almost surely as  $N$  goes to infinity. Thus one may find a matrix  $Y_N^*(\mathbf{v}, \mathbf{v})$ , that tends to the null matrix almost surely, such that

$$\begin{aligned} N \hat{Y}'_m(\mathbf{v}) \hat{C}_m^{-1}(\mathbf{v}, \mathbf{v}) \hat{Y}_m(\mathbf{v}) &= N \hat{Y}'_m(\mathbf{v}) (C_m(\mathbf{v}, \mathbf{v}) + Y_N(\mathbf{v}, \mathbf{v}))^{-1} \hat{Y}_m(\mathbf{v}) \\ &= N \hat{Y}'_m(\mathbf{v}) C_m^{-1}(\mathbf{v}, \mathbf{v}) \hat{Y}_m(\mathbf{v}) + N \hat{Y}'_m(\mathbf{v}) Y_N^*(\mathbf{v}, \mathbf{v}) \hat{Y}_m(\mathbf{v}). \end{aligned}$$

Thanks to the arguments developed in the proof of Theorem 4.3.4,  $N \hat{Y}'_m(\mathbf{v}) \hat{Y}_m$  converges in distribution. So  $N \hat{Y}'_m(\mathbf{v}) Y_N^*(\mathbf{v}, \mathbf{v}) \hat{Y}_m(\mathbf{v})$  tends to zero in distribution, hence in probability. Then  $N \hat{Y}'_m(\mathbf{v}) \hat{C}_m^{-1}(\mathbf{v}, \mathbf{v}) \hat{Y}_m(\mathbf{v})$  and  $N \hat{Y}'_m(\mathbf{v}) C_m^{-1}(\mathbf{v}, \mathbf{v}) \hat{Y}_m(\mathbf{v})$  have the same limit in distribution and the result is proved.

Turning to the global normalization matrix we write  $\hat{C}_{mT} = C_{mT} + Y_N$  where  $Y_N = N^{-2} \sum_{n=0}^{N-1} (\mathbf{S}_n \mathbf{S}'_n - \hat{\mathbf{S}}_n \hat{\mathbf{S}}'_n)$ . The proof of the convergence of  $\hat{C}_{mT}$  to  $C_{mT}$  follows the same lines that of  $\hat{C}_m(\mathbf{v}, \mathbf{v})$  to  $C_m(\mathbf{v}, \mathbf{v})$  which completes the proof of Theorem 4.3.5.  $\square$

## Bibliography

- [1] Aknouche, A., Demmouche, N., Dimitrakopoulos, S., and Touche, N. (2020). Bayesian analysis of periodic asymmetric power *garch* models. *Stud. Nonlinear Dyn. Econom.*, 24(4):1–24. [124](#)
- [2] Akutowicz, E. J. (1957). On an explicit formula in linear least squares prediction. *Math. Scand.*, 5:261–266. [107](#)
- [3] Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3):817–858. [106](#), [108](#)
- [4] Balaban, E., Bayar, A., and Kan, Ö. B. (2001). Stock returns, seasonality and asymmetric conditional volatility in world equity markets. *Applied Economics Letters*, 8(4):263–268. [124](#)
- [5] Basawa, I. V. and Lund, R. (2001). Large sample properties of parameter estimates for periodic ARMA models. *J. Time Ser. Anal.*, 22(6):651–663. [102](#)
- [6] Berk, K. N. (1974). Consistent autoregressive spectral estimates. *Ann. Statist.*, 2:489–502. Collection of articles dedicated to Jerzy Neyman on his 80th birthday. [107](#), [108](#)
- [7] Bollerslev, T. and Ghysels, E. (1996). Periodic autoregressive conditional heteroscedasticity. *Journal of Business & Economic Statistics*, 14(2):139–151. [99](#), [124](#)
- [8] Boubacar Mainassara, Y. (2011). Multivariate portmanteau test for structural VARMA models with uncorrelated but non-independent error terms. *J. Statist. Plann. Inference*, 141(8):2961–2975. [101](#)
- [9] Boubacar Maïnassara, Y. (2012). Selection of weak VARMA models by modified Akaike’s information criteria. *J. Time Series Anal.*, 33(1):121–130. [101](#)
- [10] Boubacar Mainassara, Y., Carbon, M., and Francq, C. (2012). Computing and estimating information matrices of weak ARMA models. *Comput. Statist. Data Anal.*, 56(2):345–361. [107](#), [124](#)
- [11] Boubacar Mainassara, Y. and Francq, C. (2011). Estimating structural VARMA models with uncorrelated but non-independent error terms. *J. Multivariate Anal.*, 102(3):496–505. [101](#), [107](#)
- [12] Boubacar Maïnassara, Y. and Kokonendji, C. C. (2016). Modified Schwarz and Hannan-Quinn information criteria for weak VARMA models. *Stat. Inference Stoch. Process.*, 19(2):199–217. [101](#)
- [13] Boubacar Maïnassara, Y. and Saussereau, B. (2018). Diagnostic checking in multivariate arma models with dependent errors using normalized residual autocorrelations. *J. Amer. Statist. Assoc.*, 113(524):1813–1827. [100](#), [101](#), [108](#), [109](#), [124](#), [140](#), [141](#), [144](#)
- [14] Box, G. E. P. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J. Amer. Statist. Assoc.*, 65:1509–1526. [99](#), [104](#)
- [15] Brockwell, P. J. and Davis, R. A. (1991). *Time series: theory and methods*. Springer Series in Statistics. Springer-Verlag, New York, second edition. [101](#), [106](#)
- [16] Davidson, J. (1994). *Stochastic limit theory*. Advanced Texts in Econometrics. The Clarendon Press, Oxford University Press, New York. An introduction for econometricians. [143](#)

- [17] Davydov, J. A. (1968). The convergence of distributions which are generated by stationary random processes. *Teor. Veroyatnost. i Primenen.*, 13:730–737. [138](#), [142](#)
- [18] den Haan, W. J. and Levin, A. T. (1997). A practitioner’s guide to robust covariance matrix estimation. In *Robust inference*, volume 15 of *Handbook of Statist.*, pages 299–342. North-Holland, Amsterdam. [107](#), [108](#)
- [19] Duchesne, P. and Lafaye de Micheaux, P. (2013). Distributions for residual autocovariances in parsimonious periodic vector autoregressive models with applications. *J. Time Series Anal.*, 34(4):496–507. [99](#)
- [20] Francq, C., Roy, R., and Saidi, A. (2011). Asymptotic properties of weighted least squares estimation in weak PARMA models. *J. Time Series Anal.*, 32(6):699–723. [98](#), [99](#), [100](#), [101](#), [102](#), [105](#), [124](#), [137](#), [141](#), [143](#), [144](#)
- [21] Francq, C., Roy, R., and Zakoïan, J.-M. (2005). Diagnostic checking in ARMA models with uncorrelated errors. *J. Amer. Statist. Assoc.*, 100(470):532–544. [100](#), [105](#), [107](#)
- [22] Francq, C. and Zakoïan, J.-M. (2019). *GARCH Models: Structure, Statistical Inference and Financial Applications*. Wiley. [124](#)
- [23] Franses, P. H. and Paap, R. (2000). Modelling day-of-the-week seasonality in the s&p 500 index. *Applied Financial Economics*, 10(5):483–488. [124](#)
- [24] Herrndorf, N. (1984). A functional central limit theorem for weakly dependent sequences of random variables. *Ann. Probab.*, 12(1):141–153. [137](#), [139](#), [142](#), [143](#)
- [25] Hipel, K. and McLeod, A. I. (1994). *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier, Amsterdam. [98](#), [103](#), [104](#)
- [26] Holan, S. H., Lund, R., and Davis, G. (2010). The ARMA alphabet soup: a tour of ARMA model variants. *Stat. Surv.*, 4:232–274. [98](#)
- [27] Imhof, J. P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika*, 48:419–426. [108](#)
- [28] Jones, R. H. and Brelsford, W. M. (1967). Time series with periodic structure. *Biometrika*, 54:403–408. [98](#)
- [29] Katayama, N. (2012). Chi-squared portmanteau tests for structural VARMA models with uncorrelated errors. *J. Time Series Anal.*, 33(6):863–872. [101](#)
- [30] Kuan, C.-M. and Lee, W.-M. (2006). Robust M tests without consistent estimation of the asymptotic covariance matrix. *J. Amer. Statist. Assoc.*, 101(475):1264–1275. [108](#)
- [31] Li, W. K. (2004). *Diagnostic checks in time series*. Boca Raton, Florida: Chapman and Hall. [99](#)
- [32] Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2):pp. 297–303. [99](#), [104](#)
- [33] Lobato, I. N. (2001). Testing that a dependent process is uncorrelated. *J. Amer. Statist. Assoc.*, 96(455):1066–1076. [108](#), [109](#), [110](#), [124](#), [144](#)
- [34] Lund, R. and Basawa, I. V. (2000). Recursive prediction and likelihood evaluation for periodic ARMA models. *J. Time Ser. Anal.*, 21(1):75–93. [98](#), [100](#)



- [35] Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer-Verlag, Berlin. [101](#)
- [36] Magnus, J. R. and Neudecker, H. (1999). *Matrix differential calculus with applications in statistics and econometrics*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester. Revised reprint of the 1988 original. [142](#)
- [37] McLeod, A. I. (1994). Diagnostic checking of periodic autoregression models with application. *J. Time Ser. Anal.*, 15(2):221–233. [99](#), [104](#)
- [38] McLeod, A. I. (1995). Addendum: Diagnostic checking of periodic autoregression models with application [J. Time Ser. Anal. **15** (1994), no. 2, 221–233]. *J. Time Ser. Anal.*, 16(6):647–648. [99](#), [104](#)
- [39] Newey, W. K. and West, K. D. (1987). A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708. [106](#), [108](#)
- [40] Regnard, N. and Zakoian, J.-M. (2011). A conditionally heteroskedastic model with time-varying coefficients for daily gas spot prices. *Energy Economics*, 33(6):1240–1251. [124](#)
- [41] Reinsel, G. C. (1997). *Elements of multivariate time series analysis*. Springer Series in Statistics. Springer-Verlag, New York, second edition. [101](#)
- [42] Shao, Q. and Lund, R. (2004). Computation and characterization of autocorrelations and partial autocorrelations in periodic ARMA models. *J. Time Ser. Anal.*, 25(3):359–372. [103](#)
- [43] Shao, X. (2010a). A self-normalized approach to confidence interval construction in time series. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 72(3):343–366. [108](#)
- [44] Shao, X. (2010b). Corrigendum: A self-normalized approach to confidence interval construction in time series. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 72(5):695–696. [108](#)
- [45] Shao, X. (2012). Parametric inference in stationary time series models with dependent errors. *Scand. J. Stat.*, 39(4):772–783. [108](#)
- [46] Shao, X. (2015). Self-normalization for time series: a review of recent developments. *J. Amer. Statist. Assoc.*, 110(512):1797–1817. [108](#)
- [47] Ursu, E. and Duchesne, P. (2009). On modelling and diagnostic checking of vector periodic autoregressive time series models. *J. Time Series Anal.*, 30(1):70–96. [99](#)
- [48] Vecchia, A. V. (1985). Periodic autoregressive-moving average modeling with applications to water resources. *Journal of the American Water Resources Association*, 21(5):721–730. [100](#)



# Estimation and validation of SPARMA models with uncorrelated but non-independent error terms

## Abstract

We are interested in a class of seasonal autoregressive moving average (SARMA) models with periodically varying parameters, so-called seasonal periodic autoregressive moving average models (SPARMA) under the assumption that the errors are uncorrelated but non-independent (*i.e.* weak SPARMA models). Relaxing the independence assumption on the errors considerably extends the range of application of the SPARMA models, and allows one to cover linear representations of general nonlinear processes. These models combines a periodic autoregressive moving average (PARMA) structure and a multiplicative seasonal time series models. Estimation and checking steps are considered. We establish the asymptotic properties of the quasi-generalized least squares (QLS) estimator of these models. Particular attention is given to the estimation of the asymptotic variance matrix of the QLS estimator, which may be very different from that obtained in the standard framework. We derive the asymptotic distributions of residual and normalized residual empirical autocovariances and autocorrelations of these weak SPARMA models. We then deduce the modified portmanteau statistics. We establish the asymptotic behaviour of the proposed statistics, which can be quite different from the usual chi-squared approximation used under independent and identically distributed (iid) assumptions on the noise. We also propose another test based on a self-normalization approach to check the adequacy of SPARMA models. A set of Monte Carlo experiments are presented. **Keywords:** Quasi-generalized least squares, seasonality, goodness-of-fit test, residual autocorrelations, seasonality, self-normalization, weak PARMA models, weak SARMA, weak SPARMA models.

## 5.1 Introduction

Numerous time series encountered in real applications display seasonal behavior. The multiplicative seasonal autoregressive moving average (SARMA), denoted  $SARMA(p, q)(p_T, q_T)$ , time series model is widely used in the modeling of seasonal time series (see Chapter 12 in [25]), where  $T$  denotes the length of the seasonal period.

The class of SARMA has been extensively studied in the literature. In the SARMA model, the parameters do not vary across seasons. Therefore the  $h$  lag correlations in a stationary model do not vary with the season. But we have in practice many time series with a season-dependent correlation structure. Often the application of a filter does not allow as to find a second or-

der stationarity process. On the other hand in many applications, the model parameters must change periodically to adequately describe the series. This has resulted for instance in the study of the Periodic Autoregressive Moving Average (PARMA) models, which is designed to preserve the autocorrelation structure of a time series that varies with the seasons. The periodic models have largely study in the literature (see [28], [41] and [54] among others). See also [40] who propose the use of periodic autocorrelation and periodic partial autocorrelation functions to identify PARMA models. See [50], [43] and [38] for other identification methods. The estimation procedure have also been investigated: for instanec [53] developed a procedure to obtain a maximum likelihood estimator of the parameters of a PARMA model, [44] studied the moment estimation of PARMA models, [4] studied the asymptotic properties of the parameter estimates of a PARMA model, [33] developed an algorithm for a maximum likelihood estimation of PARMA models. Recement [52] developed a robust modelling method to identify and estimate the Periodic Autoregressive (PAR) model to capture the periodic variations in river flow. See also [6], [7] who proposed an estimation procedure in order to obtain a more parsimonious and concise PAR model. For the validation step:[25] proposes a portmanteau test to check the adequacy of a PARMA model. See also [35].

Seasonal and periodic time series models are very different, as the nature of the correlation structure differs in several important respects: purely stochastic in seasonal models and purely deterministic in periodic models (see for instance [51]). Periodic models do not generally try to model explicitly the dependencies between periods (see [5]). Thus the combination of seasonal and periodical series may be worth considering. This is why we present a class of models that will allow us to have the consecutive observation dependency within a period and the observation dependency between periods (season). The proposed model is an extension oh the model introduced by [5]. The Seasonal Periodic Autoregressive Moving Average (SPARMA) model that we introduce enables the SARMA model parameters to vary periodically. This class of model contains the SARMA and PARMA model as a special case (see also [5]).

There are few studies on periodic self-regressive seasonal moving average models. See [5] who investigated stationarity conditions and the asymptotic distribution of least-squares estimators in a particular case of a SPAR. See also [51] who generalized the work of [5] to the case of multivariate SPAR models.

The process  $(X_{nT+v})$  admits a SPARMA representation with period  $T$  ( $T$  is assumed known throughout), if it satisfies the following equation

$$a(L, v)\mathbf{a}(L, v)X_{nT+v} = b(L, v)\mathbf{b}(L, v)\epsilon_{nT+v}, \quad (5.1)$$

where  $L$  stands for the back-shift operator. The non-seasonal polynomials AR and MA are defined respectively by  $a(L, v) = 1 - \sum_{i=1}^p a_{0i}(v)L^i$  and  $b(L, v) = 1 - \sum_{i=1}^q b_{0i}(v)L^i$  while the seasonal polynomials AR and MA are given respectively by  $\mathbf{a}(L, v) = 1 - \sum_{i=1}^P \mathbf{a}_{0i}(v)L^{Ti}$  and  $\mathbf{b}(L, v) = 1 - \sum_{i=1}^Q \mathbf{b}_{0i}(v)L^{Ti}$ . We assume that  $a_{0p}^2 + b_{0q}^2 + \mathbf{a}_{0P}^2 + \mathbf{b}_{0Q}^2 \neq 0$  (by convention  $a_{00} = b_{00} = \mathbf{a}_{00} = \mathbf{b}_{00} = 1$ ). We rewrite the equation (5.1) in the following expanded form:

$$\begin{aligned} X_{nT+v} - \sum_{i=1}^{p_v} a_i(v)X_{nT+v-i} - \sum_{j=1}^{P_v} \mathbf{a}_j(v)X_{(n-j)T+v} + \sum_{i=1}^{p_v} \sum_{j=1}^{P_v} \mathbf{a}_j(v)a_i(v)X_{(n-j)T+v-i} \\ = - \sum_{i=1}^{q_v} b_i(v)\epsilon_{nT+v-i} - \sum_{j=1}^{Q_v} \mathbf{b}_j(v)\epsilon_{(n-j)T+v} + \sum_{i=1}^{q_v} \sum_{j=1}^{Q_v} \mathbf{b}_j(v)b_i\epsilon_{(n-j)T+v-i} + \epsilon_{nT+v} \end{aligned} \quad (5.2)$$

The process  $(X_{nT+v})_n$ , will be denoted as follows

$$\text{SPARMA}_T(p_1, \dots, p_T; q_1, \dots, q_T)(P_1, \dots, P_T; Q_1, \dots, Q_T).$$

The linear periodic innovation process  $(\epsilon_t) = (\epsilon_{nT+v})_{n \in \mathbb{Z}}$  is assumed to be a stationary sequence satisfies:

**(A0):** (i)  $\mathbb{E}(\epsilon_{nT+v}) = 0$ , (ii)  $\text{Var}(\epsilon_{nT+v}) = \sigma_v^2$  and (iii)  $\text{Cov}(\epsilon_{nT+v}, \epsilon_{nT+v-h}) = 0 \quad \forall h, n \in \mathbb{Z}, h \neq 0$ .

**Remarque 5.1.1** *It is sometimes necessary to replace assumption (iii) with the stronger assumption that (iii') the variables  $\epsilon_{nT+v}$  and  $\epsilon_{nT+v-h}$  are independent. We then speak of strong periodic white noise. It is important to note that no independence assumption is made contrary to the weak periodic white noise Assumption (see (A0)).*

An example of weak periodic white noise is the periodic generalized autoregressive conditional heteroscedastic (PGARCH) model (see for instance [9], [24] and [1]). Many others examples can also be find in [21]. In the sequel we called a strong SPARMA model when the periodic white noise in (5.2) is strong and a weak SPARMA model in the contrary case. The distinction between a strong and a weak SPARMA is therefore only a matter of hypothesis about the noise. We deduce the following relationship

$$\{\text{strong SPARMA models}\} \subset \{\text{weak SPARMA models}\}.$$

Relaxing the strong assumption on periodic noise will allow to widen the field of application of weak SPARMA models (see for instance [21] in the weak PARMA models).

## 5.2 Causality, stationarity and invertibility of SPARMA models

In this section we study the fundamental properties of the SPARMA processes. There is no loss of generality in taking that  $p_1 = \dots = p_T = p$ ,  $q_1, \dots = q_T = q$ ,  $P_1 = \dots = P_T = P$  and  $Q_1 = \dots = Q_T = Q$  in (5.2), by adding coefficients equal to zero as in PARMA models (see [33] for more details). In this cas we simply write  $\text{SPARMA}_T(p; q)(P; Q)$  instead of  $\text{SPARMA}_T(p, \dots, p; q, \dots, q)(P, \dots, P; Q, \dots, Q)$ . Thus the  $\text{SPARMA}_T(p; q)(P; Q)$  process  $(X_{nT+v})_{n \in \mathbb{Z}}$  satisfies the following difference equation

$$\begin{aligned} X_{nT+v} - \sum_{i=1}^p a_i(v) X_{nT+v-i} - \sum_{j=1}^P \mathbf{a}_j(v) X_{(n-j)T+v} + \sum_{i=1}^p \sum_{j=1}^P \mathbf{a}_j(v) a_i(v) X_{(n-j)T+v-i} \\ = - \sum_{i=1}^q b_i(v) \epsilon_{nT+v-i} - \sum_{j=1}^Q \mathbf{b}_j(v) \epsilon_{(n-j)T+v} + \sum_{i=1}^q \sum_{j=1}^Q \mathbf{b}_j(v) b_i \epsilon_{(n-j)T+v-i} + \epsilon_{nT+v} \end{aligned} \quad (5.3)$$

introduced by [5]. We extend the results obtained by [5] in order to study the invertibility and stationarity conditions of the SPARMA models. It is well known that a SARMA( $p, q$ )( $P, Q$ ) process can be viewed as a particular case of an ARMA( $p + PT, q + QT$ ) process. Similarly a SPARMA( $p, q$ )( $P, Q$ ) model can be formulated as the following PARMA( $p + PT, q + QT$ ) model:

$$X_{nT+v} - \sum_{i=1}^{p+PT} a_i^*(v) X_{nT+v-i} = - \sum_{i=1}^{q+QT} b_i^*(v) \epsilon_{nT+v-i} + \epsilon_{nT+v} \quad (5.4)$$

where the coefficients are given by

$$a_i^*(v) = \begin{cases} a_i(v) & \text{for } 1 \leq i \leq p \\ 0 & \text{for } p < i < T \end{cases} \quad a_{Tj}^*(v) = \mathbf{a}_j(v) \text{ for } 1 \leq j \leq P; \quad a_{Tj+i}^*(v) = -a_i(v) \mathbf{a}_j(v)$$

for  $1 \leq i \leq p$ ,  $1 \leq j \leq P$  and similarly

$$b_i^*(v) = \begin{cases} b_i(v) & \text{for } 1 \leq i \leq q \\ 0 & \text{for } q < i < T \end{cases} \quad b_{Tj}^*(v) = \mathbf{b}_j(v) \text{ for } 1 \leq j \leq Q \text{ and } b_{Tj+i}^*(v) = -b_{0i}(v) \mathbf{b}_j(v)$$

for  $1 \leq i \leq q$ ,  $1 \leq j \leq Q$ .

We denote by  $[a]$  the integer part of the real  $a$ . Let  $p^* = [(p + PT - 1)/T] + 1$  and  $q^* = [(q + QT - 1)/T] + 1$ .

As in [55] the difference equations (5.4) can be written in the following T-dimensional vector ARMA (VARMA) form:

$$A_0 \mathbf{X}_n - \sum_{k=1}^{p^*} A_k \mathbf{X}_{n-k} = B_0 \boldsymbol{\epsilon}_n - \sum_{l=1}^{q^*} B_l \boldsymbol{\epsilon}_{n-l} \quad (5.5)$$

where

$$\mathbf{X}_n = (X_{nT+1}, \dots, X_{nT+T}), \quad \boldsymbol{\epsilon}_n = (\epsilon_{nT+1}, \dots, \epsilon_{nT+T}). \quad (5.6)$$

The  $(T \times T)$ -matrices coefficients  $A_k, k = 0, \dots, p^*$  and  $B_l, l = 0, \dots, q^*$  are defined by

$$(A_0)_{i,j} = \begin{cases} 1 & \text{Si } i = j \\ 0 & \text{Si } i < j \\ -a_{i-j}^*(i) & \text{Si } i > j \end{cases} \quad (B_0)_{i,j} = \begin{cases} 1 & \text{Si } i = j \\ 0 & \text{Si } i < j \\ -b_{i-j}^*(i) & \text{Si } i > j \end{cases}$$

with  $(A_k)_{i,j} = a_{kT+i-j}^*(i), k = 1, \dots, p^*$  and  $(B_l)_{i,j} = b_{lT+i-j}^*(i), l = 1, \dots, q^*$ . The covariance matrix of the T-dimensional white noise  $\boldsymbol{\epsilon}_n$  is  $\Sigma_{\boldsymbol{\epsilon}} = \text{Diag}(\sigma_1^2, \dots, \sigma_T^2) > 0$ . We denote by L the back-shift operator such that  $L^k \boldsymbol{\epsilon}_n = \boldsymbol{\epsilon}_{n-k}$ . It is important to note that the lag operator L operates on the cycle index  $n$ . When it acts on the time index  $t = nT + v$  of the periodic process  $(\epsilon_t)$  it gives  $L^k \epsilon_{nT+v} = \epsilon_{nT+v-k}$ . Equation (5.5) can be written more compactly as

$$A(L)\mathbf{X}_n = B(L)\boldsymbol{\epsilon}_n \quad (5.7)$$

where  $A(z) = A_0 - \sum_{i=1}^{p^*} A_i z^i$  and  $B(z) = B_0 - \sum_{i=1}^{q^*} B_i z^i$ . From (5.6) we can in principle deduce the properties of weak SPARMA parameters estimation, identification and validation from existing results on parameters estimation, identification and validation of the weak VARMA models (see for instance [10, 11, 14, 15, 16, 29]). Therefore we have preferred to work in the univariate SPARMA setting for various reasons. In particular the results obtained directly in terms of the univariate SPARMA representation are more directly usable because fewer parameters are involved and their estimation is easier as in the weak PARMA case (see [22] for more details).

We denote by  $\det(M)$  the determinant of a matrix M. We assume that the SPARMA $_T(p; q)(P; Q)$  process  $(X_{nT+v})_{n \in \mathbb{Z}}$  corresponds to stable and invertible representations, namely

**(A1):** we have  $\det(A(z)) \det(B(z)) \neq 0$  for all  $|z| \leq 1$ . Furthermore as in [14] we assume that the weak VARMA model (5.7) is identifiable (see also [18, 34, 42]).

**Example 5.2.1** We now examine the causality and stationarity conditions on the following SPARMA(1, 1)(1, 1) model:

$$X_{nT+v} = a_1(v)X_{nT+v-1} + \mathbf{a}_1(v)X_{(n-1)T+v} + \mathbf{a}_1(v)a_1(v)X_{(n-1)T+v-1} - b_1(v)\epsilon_{nT+v-1} - \mathbf{b}_1(v)\epsilon_{(n-1)T+v} + \mathbf{b}_1(v)b_1(v)\epsilon_{(n-1)T+v-1} + \epsilon_{nT+v}.$$

This SPARMA(1, 1)(1, 1) model admits the following PARMA(1 + T, 1 + T) representation:

$$X_{nT+v} - \sum_{i=1}^{1+T} a_i^*(v)X_{nT+v-i} = - \sum_{i=1}^{1+T} b_i^*(v)\epsilon_{nT+v-i} + \epsilon_{nT+v}.$$

Finally we deduce the following VARMA(2, 2) model:

$$A_0 \mathbf{X}_n = A_1 \mathbf{X}_{n-1} + A_2 \mathbf{X}_{n-2} - B_1 \boldsymbol{\epsilon}_{n-1} - B_2 \boldsymbol{\epsilon}_{n-2} + B_0 \boldsymbol{\epsilon}_n$$

where the matrices  $A_0, A_1, A_2, B_0, B_1$  and  $B_2$  are given for simplify for  $T = 4$  by

$$A_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -a_1^*(2) & 1 & 0 & 0 \\ -a_2^*(3) & -a_1^*(3) & 1 & 0 \\ -a_3^*(4) & -a_2^*(4) & -a_1^*(4) & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -a_1(2) & 1 & 0 & 0 \\ 0 & -a_1(3) & 1 & 0 \\ 0 & 0 & -a_1(4) & 1 \end{pmatrix}$$

$$A_1 = \begin{pmatrix} a_4^*(1) & a_3^*(1) & a_2^*(1) & a_1^*(1) \\ a_5^*(2) & a_4^*(2) & a_3^*(2) & a_2^*(2) \\ a_6^*(3) & a_5^*(3) & a_4^*(3) & a_3^*(3) \\ a_7^*(4) & a_6^*(4) & a_5^*(4) & a_4^*(4) \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1(1) & 0 & 0 & a_1(1) \\ -\mathbf{a}_1(2)a_1(2) & \mathbf{a}_1(2) & 0 & 0 \\ 0 & -\mathbf{a}_1(3)a_1(3) & \mathbf{a}_1(3) & 0 \\ 0 & 0 & -\mathbf{a}_1(4)a_1(4) & \mathbf{a}_1(4) \end{pmatrix}$$

$$A_2 = \begin{pmatrix} a_8^*(1) & a_7^*(1) & a_6^*(1) & a_5^*(1) \\ a_9^*(2) & a_8^*(2) & a_7^*(2) & a_6^*(2) \\ a_{10}^*(3) & a_9^*(3) & a_8^*(3) & a_7^*(3) \\ a_{11}^*(4) & a_{10}^*(4) & a_9^*(4) & a_8^*(4) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -\mathbf{a}_1(1)a_1(1) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$A_0 - A_1 z - A_2 z^2 = \begin{pmatrix} 1 - \mathbf{a}_1(1)z & 0 & 0 & \mathbf{a}_1(1)a_1(1)z^2 - a_1(1)z \\ \mathbf{a}_1(2)a_1(2)z - a_1(2)z & 1 - \mathbf{a}_1(2)z & 0 & 0 \\ 0 & \mathbf{a}_1(3)a_1(3)z - a_1(3)z & 1 - \mathbf{a}_1(3)z & 0 \\ 0 & 0 & \mathbf{a}_1(4)a_1(4)z - a_1(4)z & 1 - \mathbf{a}_1(4)z \end{pmatrix}$$

we have  $\det(A_0 - A_1 z - A_2 z^2) = \prod_{v=1}^T (1 - \mathbf{a}_1(v)z) (1 - \prod_{v=1}^T a_1(v)z) \neq 0$

and

$$B_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -b_1^*(2) & 1 & 0 & 0 \\ -b_2^*(3) & -b_1^*(3) & 1 & 0 \\ -b_3^*(4) & -b_2^*(4) & -b_1^*(4) & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -b_1(2) & 1 & 0 & 0 \\ 0 & -b_1(3) & 1 & 0 \\ 0 & 0 & -b_1(4) & 1 \end{pmatrix}$$

$$B_1 = \begin{pmatrix} b_4^*(1) & b_3^*(1) & b_2^*(1) & b_1^*(1) \\ b_5^*(2) & b_4^*(2) & b_3^*(2) & b_2^*(2) \\ b_6^*(3) & b_5^*(3) & b_4^*(3) & b_3^*(3) \\ b_7^*(4) & b_6^*(4) & b_5^*(4) & b_4^*(4) \end{pmatrix} = \begin{pmatrix} \mathbf{b}_1(1) & 0 & 0 & b_1(1) \\ -\mathbf{b}_1(2)b_1(2) & \mathbf{b}_1(2) & 0 & 0 \\ 0 & -\mathbf{b}_1(3)b_1(3) & \mathbf{b}_1(3) & 0 \\ 0 & 0 & -\mathbf{b}_1(4)b_1(4) & \mathbf{b}_1(4) \end{pmatrix}$$

$$B_2 = \begin{pmatrix} b_8^*(1) & b_7^*(1) & b_6^*(1) & b_5^*(1) \\ b_9^*(2) & b_8^*(2) & b_7^*(2) & b_6^*(2) \\ b_{10}^*(3) & b_9^*(3) & b_8^*(3) & b_7^*(3) \\ b_{11}^*(4) & b_{10}^*(4) & b_9^*(4) & b_8^*(4) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \mathbf{b}_1(1)b_1(1) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

We have  $\det(B_0 - B_1 z - B_2 z^2) = \prod_{v=1}^T (1 - \mathbf{b}_1(v)z) (1 - \prod_{v=1}^T b_1(v)z) \neq 0$ .

## 5.3 Estimating weak SPARMA models

### 5.3.1 Notations and assumptions

For notation, let  $a(v) = (a_1(v), \dots, a_p(v))'$  and  $b(v) = (b_1(v), \dots, b_q(v))'$  respectively denote the non seasonal vectors of AR and MA parameters for a specified season  $v$ . Similarly we denote by  $\mathbf{a}(v) = (\mathbf{a}_1(v), \dots, \mathbf{a}_p(v))'$  and  $\mathbf{b}(v) = (\mathbf{b}_1(v), \dots, \mathbf{b}_q(v))'$  respectively the seasonal vectors of autoregressive and moving average parameters for a specified season  $v$ .

The  $T(p + q + P + Q)$ -dimensional collection of all SPARMA parameters is also denoted by

$$\alpha = (a(1)', \dots, a(T)', b(1)', \dots, b(T)', \mathbf{a}(1)', \dots, \mathbf{a}(T)', \mathbf{b}(1)', \dots, \mathbf{b}(T)')' \in \mathbb{R}^{k_0 T}$$

where  $k_0 = (p + q + P + Q)$  is the number of parameters in each season. The periodic white noise variance  $\sigma^2 := (\sigma_1^2, \dots, \sigma_T^2)'$  will be treated as a nuisance parameter.

Let  $X_1, \dots, X_{NT}$  be a data sample from the causal and invertible SPARMA model (5.3) with the true parameter value  $\alpha = \alpha_0$  and  $\sigma^2 = \sigma_0^2$ . The sample contains  $N$  full periods of data which are

indexed from 0 to  $N-1$ . Indeed when  $0 \leq n \leq N-1$  and  $1 \leq v \leq T$ ,  $nT+v$  goes from 1 to  $NT$ . The unknown parameter of interest  $\alpha_0$  is supposed to belong to the parameter space

$$\Delta = \left\{ \alpha = (a(1)', \dots, a(T)', b(1)', \dots, b(T)', \mathbf{a}(1)', \dots, \mathbf{a}(T)', \mathbf{b}(1)', \dots, \mathbf{b}(T)')' \right. \\ \left. \text{such that (A1) is verified} \right\}.$$

For  $\alpha \in \Delta$ , let  $\epsilon_{nT+v}(\alpha)$  be the periodically second-order stationary solution of

$$\epsilon_{nT+v}(\alpha) = X_{nT+v} - \sum_{i=1}^p a_i X_{nT+v-i} - \sum_{j=1}^P \mathbf{a}_j X_{nT+v-Tj} + \sum_{i=1}^p \sum_{j=1}^P a_i \mathbf{a}_j X_{nT+v-Tj-i} \\ + \sum_{i=1}^q b_i \epsilon_{nT+v-i}(\alpha) + \sum_{j=1}^Q \mathbf{b}_j \epsilon_{nT+v-Tj}(\alpha) - \sum_{i=1}^q \sum_{j=1}^Q b_i \mathbf{b}_j \epsilon_{nT+v-Tj-i}(\alpha). \quad (5.8)$$

The variable  $\epsilon_{nT+v}(\alpha)$  can be approximated, for  $0 < nT+v \leq NT$ , by  $e_{nT+v}(\alpha)$  defined recursively by

$$e_{nT+v}(\alpha) = X_{nT+v} - \sum_{i=1}^p a_i X_{nT+v-i} - \sum_{j=1}^P \mathbf{a}_j X_{nT+v-Tj} + \sum_{i=1}^p \sum_{j=1}^P a_i \mathbf{a}_j X_{nT+v-Tj-i} \\ + \sum_{i=1}^q b_i e_{nT+v-i}(\alpha) + \sum_{j=1}^Q \mathbf{b}_j e_{nT+v-Tj}(\alpha) - \sum_{i=1}^q \sum_{j=1}^Q b_i \mathbf{b}_j e_{nT+v-Tj-i}(\alpha), \quad (5.9)$$

where the unknown initial values are set to zero:  $e_0(\alpha) = \dots = e_{1-q-TQ}(\alpha) = X_0 = \dots = X_{1-p-TP} = 0$ . As showed by [21, Lemma 7] these initial values are asymptotically negligible uniformly in  $\alpha$ .

Let  $\delta > 0$  chosen such that  $\alpha_0$  belongs to the compact set

$$\Delta_\delta = \left\{ \alpha = (a(1)', \dots, a(T)', b(1)', \dots, b(T)', \mathbf{a}(1)', \dots, \mathbf{a}(T)', \mathbf{b}(1)', \dots, \mathbf{b}(T)')' \right. \\ \left. \text{such that the zeros of } \det((A(z))) \text{ and those of } \det((B(z))) \text{ have modulus } \geq 1 + \delta \right\}.$$

The random variable  $\hat{\alpha}_{OLS}$  is called the ordinary least squares (OLS) estimator of  $\alpha$  if it satisfies, almost surely,

$$S_N(\hat{\alpha}_{OLS}) = \min_{\alpha \in \Delta_\delta} S_N(\alpha) \quad \text{where} \quad S_N(\alpha) = \frac{1}{2N} \sum_{n=0}^{N-1} \sum_{v=1}^T e_{nT+v}^2(\alpha).$$

Because of the presence of heteroskedastic innovation, [5] has shown that the OLS estimator is inefficient compared to weighted least squares estimators in the particular case of  $\text{SPARMA}_T(1, 0)(1, 0)$ . Therefore in this paper we have chosen to consider the quasi-generalised least squares (QLS) estimator. We first define the generalized least squares (GLS) estimator proposed by [5] which is optimal among weighted least squares estimators:

$$Q_N^{\sigma^2}(\hat{\alpha}_{GLS}) = \min_{\alpha \in \Delta_\delta} Q_N^{\sigma^2}(\alpha) \quad \text{where} \quad Q_N^{\sigma^2}(\alpha) = \frac{1}{2N} \sum_{n=0}^{N-1} \sum_{v=1}^T \sigma_v^{-2} e_{nT+v}^2(\alpha). \quad (5.10)$$

This estimator assumes that the parameter  $\sigma^2 = (\sigma_1^2, \dots, \sigma_T^2)'$  is known, but in practice it has to be estimated. Hence a possible consistent estimator of  $\sigma_v^2$  is given by

$$\hat{\sigma}_v^2 = \frac{1}{N} \sum_{n=0}^{N-1} e_{nT+v}^2(\hat{\alpha}_{OLS}).$$

For any consistent estimator  $\hat{\sigma}^2 = (\hat{\sigma}_1^2, \dots, \hat{\sigma}_T^2)'$  of  $\sigma^2$ , the QLS estimator of  $\alpha_0$  is defined by:

$$\hat{\alpha}_{QLS} = \hat{\alpha}_{GLS}^{\hat{\sigma}^2}. \quad (5.11)$$

To establish the consistency of the QLS estimator (QLSE), an additional assumption is needed.



**(A2):** The  $T$ -dimensional white noise  $\{\boldsymbol{\epsilon}_n = (\epsilon_{nT+1}, \dots, \epsilon_{nT+T})', n \in \mathbb{Z}\}$  is ergodic and strictly stationary.

For the asymptotic normality of the QLSE, additional assumptions are also required. It is necessary to assume that  $\alpha_0$  is not on the boundary of the parameter space  $\Delta_\delta$ .

**(A3):** We have  $\alpha_0 \in \overset{\circ}{\Delta}_\delta$ , where  $\overset{\circ}{\Delta}_\delta$  denotes the interior of  $\Delta_\delta$ .

To control the serial dependence of the stationary process  $(\boldsymbol{\epsilon}_n)$ , we introduce the strong mixing coefficients  $\alpha_\epsilon(h)$  defined by

$$\alpha_\epsilon(h) = \sup_{A \in \mathcal{F}_{-\infty}^n, B \in \mathcal{F}_{n+h}^{+\infty}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|,$$

where  $\mathcal{F}_{-\infty}^n = \sigma(\boldsymbol{\epsilon}_u, u \leq n)$  and  $\mathcal{F}_{n+h}^{+\infty} = \sigma(\boldsymbol{\epsilon}_u, u \geq n+h)$ . We use  $\|\cdot\|$  to denote the Euclidian norm of a vector. We will make an integrability assumption on the moment of the noise and a summability condition on the strong mixing coefficients  $(\alpha_\epsilon(k))_{k \geq 0}$ .

**(A4):** We have  $\mathbb{E}\|\boldsymbol{\epsilon}_n\|^{4+2\kappa} < \infty$  and  $\sum_{k=0}^{\infty} \{\alpha_\epsilon(k)\}^{\frac{\kappa}{2+\kappa}} < \infty$  for some  $\kappa > 0$ .

In all the sequel we denote by  $\xrightarrow{d}$  the convergence in distribution. The symbol  $o_{\mathbb{P}}(1)$  is used for a sequence of random variables that converges to zero in probability.

### 5.3.2 Asymptotic properties

**Theorem 5.3.1** *Suppose that  $X_{nT+v}$  is a SPARMA process. Let  $\hat{\alpha}_{\text{QLS}}$  be the QLS estimator defined in (5.11). Then under Assumptions (A0), (A1) and (A2), we have:  $\hat{\alpha}_{\text{QLS}} \rightarrow \alpha_0$  a.s. as  $N \rightarrow \infty$ .*

**Theorem 5.3.2** *Under Assumptions of Theorem 5.3.1, (A3) and (A4), we have:*

$$\sqrt{n}(\hat{\alpha}_{\text{QLS}} - \alpha_0) \xrightarrow{d} \mathcal{N}(0, \Omega := (J(\alpha_0, \hat{\sigma}^2))^{-1} I(\alpha_0, \hat{\sigma}^2) (J(\alpha_0, \hat{\sigma}^2))^{-1})$$

where

$$J(\alpha_0, \hat{\sigma}^2) = \sum_{v=1}^T \hat{\sigma}_v^{-2} \mathbb{E} \left[ \left( \frac{\partial \epsilon_{nT+v}(\alpha)}{\partial \alpha} \right) \left( \frac{\partial \epsilon_{nT+v}(\alpha)}{\partial \alpha} \right)' \right]_{\alpha=\alpha_0}, \quad (5.12)$$

$$I(\alpha_0, \hat{\sigma}^2) = \sum_{v=1}^T \sum_{v'=1}^T \hat{\sigma}_v^{-2} \hat{\sigma}_{v'}^{-2} \sum_{k=-\infty}^{\infty} \mathbb{E} \left[ \epsilon_v(\alpha) \epsilon_{kT+v'}(\alpha) \left( \frac{\partial \epsilon_v(\alpha)}{\partial \alpha} \right) \left( \frac{\partial \epsilon_{kT+v'}(\alpha)}{\partial \alpha} \right)' \right]_{\alpha=\alpha_0}. \quad (5.13)$$

**Remarque 5.3.1** *In the strong SPARMA case, i.e. when (A2) is replaced by the assumption that  $(\epsilon_{nT+v})_{n \in \mathbb{Z}}$  is iid, we have  $I(\alpha_0, \hat{\sigma}^2) = J(\alpha_0, \hat{\sigma}^2)$ , so that the covariance matrix in the strong case is  $\Omega_s := (J(\alpha_0, \hat{\sigma}^2))^{-1}$ . This is the results obtained by [5] in the SPAR(1, 1) case.*

*In the general case we have  $I(\alpha_0, \hat{\sigma}^2) \neq J(\alpha_0, \hat{\sigma}^2)$  and the matrix expression  $I(\alpha_0, \hat{\sigma}^2)$  is in general much more complicated in the weak case than in the strong case. As a consequence the ready-made software used to fit SPARMA do not provide a correct estimation for weak SPARMA processes.*

## 5.4 Estimating the asymptotic variance matrix

Let  $\hat{\epsilon}_{nT+v} = \hat{e}_{nT+v} = e_{nT+v}(\hat{\alpha})$  be the least squares periodic residuals when  $p + q + P + Q > 0$ . We use (5.9) to notice that we have  $\hat{e}_t = 0$  for  $t := nT + v \leq 0$  and  $t > NT$ . By (5.3), for  $n = 0, 1, \dots, N-1$ , it holds that

$$\begin{aligned} \hat{e}_{nT+v} &= X_{nT+v} - \sum_{i=1}^p \hat{a}_i X_{nT+v-i} - \sum_{j=1}^P \hat{\mathbf{a}}_j X_{nT+v-Tj} + \sum_{i=1}^p \sum_{j=1}^P \hat{a}_i \hat{\mathbf{a}}_j X_{nT+v-Tj-i} \\ &\quad + \sum_{i=1}^q \hat{b}_i e_{nT+v-i} + \sum_{j=1}^Q \hat{\mathbf{b}}_j \hat{e}_{nT+v-Tj} - \sum_{i=1}^q \sum_{j=1}^Q \hat{b}_i \hat{\mathbf{b}}_j \hat{e}_{nT+v-Tj-i}, \end{aligned}$$

for  $1 \leq t \leq NT$ , with  $\hat{X}_t = 0$  for  $t \leq 0$  and  $\hat{X}_t = X_t$  for  $t \geq 1$ .

Theorem 5.3.2 can be used to obtain confidence intervals and significance tests for the parameters. The asymptotic covariance  $\Omega$  must however be estimated for statistical inference.

### 5.4.1 Estimation of the matrix $J(\alpha_0, \hat{\sigma}^2)$

The matrix  $J(\alpha_0, \hat{\sigma}^2)$  can be estimated by its empirical counterpart  $J_N(\hat{\alpha}, \hat{\sigma}^2)$ :

$$J_N(\hat{\alpha}, \hat{\sigma}^2) = \sum_{v=1}^T \sigma_v^{-2} \frac{1}{N} \sum_{n=0}^{N-1} \left\{ \frac{\partial}{\partial \alpha} e_{nT+v}(\hat{\alpha}) \right\} \left\{ \frac{\partial}{\partial \alpha} e_{nT+v}(\hat{\alpha}) \right\}'. \quad (5.14)$$

In the standard strong SPARMA case  $\hat{\Omega}_S = (J_N(\hat{\alpha}, \hat{\sigma}^2))^{-1}$  is a strongly consistent estimator of  $\Omega_S$ . In the general weak SPARMA case this estimator is not consistent when  $I(\alpha_0, \hat{\sigma}^2) \neq J(\alpha_0, \hat{\sigma}^2)$  (see Remark 5.3.1). So we need a consistent estimator of  $I(\alpha_0, \hat{\sigma}^2)$ , defined by (5.13).

### 5.4.2 Estimation of the asymptotic matrix $I(\alpha_0, \hat{\sigma}^2)$

Let

$$\Upsilon_n := \Upsilon_n(\alpha_0) = \sum_{v=1}^T \hat{\sigma}_v^{-2} e_{nT+v}(\alpha_0) \frac{\partial e_{nT+v}(\alpha_0)}{\partial \alpha} \quad (5.15)$$

For any  $\alpha \in \Delta_\delta$ ,  $\Upsilon_n$  is a measurable function of  $\{\epsilon_s, s \leq nT + v\}$ . The estimation of this long-run variance  $I(\alpha_0, \hat{\sigma}^2)$  is more complicated. In the literature, two types of estimators are generally employed: the nonparametric kernel estimator, also called Heteroskedasticity and Autocorrelation Consistent (HAC) estimators (see [3] and [39] for general references, and [23] for an application to testing strong linearity in weak ARMA models) and spectral density estimators (see e.g. [8] and [20] for a general references and [12] for estimating  $I(\alpha_0, \hat{\sigma}^2)$  when  $\alpha$  is not necessarily equal to  $\alpha_0$ ).

In the present paper, we focus on an estimator based on a spectral density form for  $I(\alpha_0, \hat{\sigma}^2)$ .

The stationary process  $(\Upsilon_n)$  admit the Wold decomposition

$$\Upsilon_n = u_n + \sum_{i=1}^{\infty} \Psi_i u_{n-i},$$

where  $(u_n)$  is a  $Tk_0$ -variate weak white noise with covariance matrix  $\Sigma_u := \text{Var}(u_n)$ . Assume that  $\Sigma_u$  is non singular, that  $\sum_{i=1}^{\infty} \|\Psi_i\| < \infty$ , and that  $\det(I_{Tk_0} + \sum_{i=1}^{\infty} \Psi_i z^i) \neq 0$  when  $|z| \leq 1$ . Then  $(\Upsilon_n)$  admits a VAR( $\infty$ ) representation (see [2]) of the form

$$\Phi(L)\Upsilon_n = \Upsilon_n - \sum_{i=1}^{\infty} \Phi_i \Upsilon_{n-i} = u_n \quad (5.16)$$

such that  $\sum_{i=1}^{\infty} \|\Phi_i\| < \infty$  and  $\det\{\Phi(z)\} \neq 0$  for all  $|z| \leq 1$ . Interpreting  $(2\pi)^{-1}I(\theta_0, \hat{\sigma}^2)$  as the spectral density of the stationary process  $(\Upsilon_n)_{n \in \mathbb{Z}}$  evaluated at frequency 0 (see [18], p. 459) of the process (5.15). We then obtain that

$$I(\theta_0, \hat{\sigma}^2) = \Phi^{-1}(1)\Sigma_u\Phi^{-1}(1)'. \quad (5.17)$$

Consider the regression of  $\Upsilon_n$  on  $\Upsilon_{n-1}, \dots, \Upsilon_{n-r}$  defined by

$$\Upsilon_n = \sum_{i=1}^r \Phi_{r,i}(v)\Upsilon_{n-i} + u_{r,n} \quad (5.18)$$



where  $u_{r,n}$  is uncorrelated with  $Y_{n-1}, \dots, Y_{n-r}$ . Since  $Y_n(\alpha_0)$  is not observable, we introduce  $\hat{Y}_n \in \mathbb{R}^{T k_0}$  obtained by replacing  $\epsilon_{nT+v}$  by  $e_{nT+v}$  and  $\alpha_0$  by  $\hat{\alpha}_n$  in (5.15):

$$\hat{Y}_n = \sum_{v=1}^T \hat{\sigma}_v^{-2} e_{nT+v}(\hat{\alpha}_n) \frac{\partial e_{nT+v}(\hat{\alpha}_n)}{\partial \alpha}. \quad (5.19)$$

Let  $\hat{\Phi}_r(z) = I_{T k_0} + \sum_{k=1}^r \hat{\Phi}_{r,k} z^k$ , where  $\hat{\Phi}_{r,1}, \dots, \hat{\Phi}_{r,r}$  denote the coefficients of the least squares regression of  $\hat{Y}_n$  on  $\hat{Y}_{n-1}, \dots, \hat{Y}_{n-r}$ . Let  $\hat{u}_{r,n}$  be the residuals of this regression and  $\hat{\Sigma}_{\hat{u}_r}$  be the empirical variance of  $\hat{u}_{r,0}, \dots, \hat{u}_{r,N-1}$ . The least squares estimator of  $\underline{\Phi}_r = (\Phi_{r,1} \dots \Phi_{r,r})$  and  $\Sigma_{u_r} = \text{Var}(u_{r,n})$  are defined by

$$\hat{\underline{\Phi}}_r = \hat{\Sigma}_{\hat{Y}, \hat{Y}_r} \hat{\Sigma}_{\hat{Y}_r}^{-1} \quad \text{and} \quad \hat{\Sigma}_{\hat{u}_r} = \frac{1}{N} \sum_{n=0}^{N-1} \left( \hat{Y}_n - \hat{\underline{\Phi}}_r \hat{Y}_{r,n} \right) \left( \hat{Y}_n - \hat{\underline{\Phi}}_r \hat{Y}_{r,n} \right)' \quad (5.20)$$

where

$$\hat{Y}_{r,n} = (\hat{Y}'_{n-1} \dots \hat{Y}'_{n-r})', \quad \hat{\Sigma}_{\hat{Y}, \hat{Y}_r} = \frac{1}{N} \sum_{n=0}^{N-1} \hat{Y}_r \hat{Y}'_{r,n}, \quad \hat{\Sigma}_{\hat{Y}_r}^{-1} = \frac{1}{N} \sum_{n=0}^{N-1} \hat{Y}_{r,n} \hat{Y}'_{r,n}$$

with by convention  $\hat{Y}_t = 0$  when  $n \leq 0$ , and assuming  $\hat{\Sigma}_{\hat{Y}_r}$  is non singular (which holds true asymptotically).

The asymptotic study of the estimator of  $I(\alpha_0, \hat{\sigma}^2)$  using the spectral density method is given in the following theorem.

**Theorem 5.4.1** *Under Assumptions (A0), (A1), (A3) and (A4), we assume that the process  $(Y_n)_{n \in \mathbb{Z}}$  defined by (5.15) admits the VAR( $\infty$ ) representation (5.16), where  $\|\Phi_i\| = o(i^{-2})$  as  $i \rightarrow \infty$ , the roots of  $\det(\Phi(z)) = 0$  are outside the unit disk, and  $\Sigma_u = \text{Var}(u_n)$  is non singular. Assume moreover that  $\mathbb{E}\|\epsilon_n\|^{8+4\kappa} < \infty$  for some  $\kappa > 0$ . Then the spectral estimator of  $I(\alpha_0, \hat{\sigma}^2)$ :*

$$\hat{I}_N^{\text{SP}} := \hat{\Phi}_r^{-1}(1) \hat{\Sigma}_{\hat{u}_r} \hat{\Phi}_r^{-1}(1) \xrightarrow{N \rightarrow \infty} I(\alpha_0, \hat{\sigma}^2) = \Phi^{-1}(1) \Sigma_u \Phi^{-1}(1)'$$

in probability when  $r = r(N) \rightarrow \infty$  and  $r^3/N \rightarrow 0$  as  $N \rightarrow \infty$ .

## 5.5 Diagnostic checking in weak SPARMA models

After the estimation phase, the next important step consists in checking if the estimated model fits satisfactorily the data. In this section we derive the limiting distribution of the residual autocorrelations and that of the portmanteau statistics based on the standard and the self-normalized approaches in the framework of weak SPARMA models (5.3).

We denote by

$$\gamma(v, l) = \frac{1}{N} \sum_{n=l+1}^N \epsilon_{nT+v} \epsilon_{nT+v-l} \quad \text{and} \quad \hat{\gamma}(v, l) = \frac{1}{N} \sum_{n=l+1}^N \hat{e}_{nT+v} \hat{e}_{nT+v-l} \quad (5.21)$$

the periodic white noise "empirical" autocovariances and residuals autocovariances. It should be noted that  $\gamma(v, l)$  is not a computable statistic because it depends on the unobserved innovations  $\epsilon_{nT+v} = \epsilon_{nT+v}(\alpha_0)$ . For a fixed integer  $m \geq 1$  we let

$$\gamma_m(v) := (\gamma(v, 1), \dots, \gamma(v, m))' \quad \text{and} \quad \hat{\gamma}_m(v) := (\hat{\gamma}(v, 1), \dots, \hat{\gamma}(v, m))'$$

be the vectors of "empirical" autocovariances and residuals autocovariances for a specified season  $v$ . Similarly the global "empirical" autocovariances and residuals autocovariances across all the seasons are denoted by

$$\gamma_{mT} := (\gamma_m(1), \dots, \gamma_m(T))' \quad \text{and} \quad \hat{\gamma}_{mT} := (\hat{\gamma}_m(1), \dots, \hat{\gamma}_m(T))'$$

Note that the mixing assumption **(A4)** will entail the asymptotic normality of  $\gamma_m(\nu)$  and  $\gamma_{mT}$ .

The theoretical and sample autocorrelations at lag  $l$  are respectively defined by

$$\rho(\nu, l) = \frac{\Gamma(\nu, l)}{\sqrt{\Gamma(\nu, 0)\Gamma(\nu - l, 0)}} \text{ and } \hat{\rho}(\nu, l) = \frac{\hat{\Upsilon}(\nu, l)}{\sqrt{\hat{\Upsilon}(\nu, 0)\hat{\Upsilon}(\nu - l, 0)}} \text{ with } \hat{\Upsilon}(\nu - l, 0) = \frac{1}{N} \sum_{n=1}^N \hat{\epsilon}_{nT+\nu-l}^2,$$

and where  $\Gamma(\nu, l) := \text{Cov}(\epsilon_{nT+\nu}, \epsilon_{nT+\nu-l})$  with  $\Gamma(\nu, 0) := \sigma_\nu^2$ . See for instance [25, 45] for some properties of  $\Gamma(\nu, l)$ . In the sequel we will also need these expressions:

$$\hat{\rho}_m(\nu) = (\hat{\rho}(\nu, 1), \dots, \hat{\rho}(\nu, m))' \text{ and } \hat{\rho}_{mT} = (\hat{\rho}_m(1), \dots, \hat{\rho}_m(T))',$$

where  $\hat{\rho}_m(\nu)$  denotes the vector of the first  $m$  sample autocorrelations for a specified season  $\nu$  and  $\hat{\rho}_{mT}$  is the vector of the global sample autocorrelations across all the seasons.

Based on the residual empirical autocorrelations, [17] have proposed a goodness-of-fit test, the so-called portmanteau test (BP), for strong ARMA models. A modification of their test has been proposed by [31] which is nowadays one of the most popular diagnostic checking tools in ARMA modeling of time series. To test simultaneously whether all residual autocorrelations at lags  $1, \dots, m$  of a SPARMA model are equal to zero for a specified period  $\nu$ , the portmanteau of [17, 31] can be adapted as proposed by [36, 37] (see also [25]). Based on these results [36, 37] suggested the following portmanteau statistics defined by

$$Q_m^{\text{BP}}(\nu) = N \sum_{h=1}^m \hat{\rho}^2(\nu, h) \text{ and } Q_m^{\text{LBM}}(\nu) = N \sum_{h=1}^m \varrho(h, \nu, N, T) \hat{\rho}^2(\nu, h), \quad (5.22)$$

where  $\varrho(h, \nu, N, T)$  is called the Ljung-Box-McLeod (LBM) correction and is defined by

$$\varrho(h, \nu, N, T) = \begin{cases} (N+2)/(N-h/T) & \text{if } h \equiv 0 \pmod{T} \\ N/(N - \lfloor (h - \nu + T)/T \rfloor) & \text{otherwise.} \end{cases}$$

The statistic  $Q_m^{\text{LBM}}(\nu)$  has the same asymptotic chi-squared distribution as  $Q_m^{\text{BP}}(\nu)$  and has the reputation of doing better for small or medium sized sample (see [36, 37]). If the strong SPARMA $_T(p; q)(P, Q)$  model is adequate for a specified period  $\nu$ ,  $Q_m^{\text{BP}}(\nu)$  or  $Q_m^{\text{LBM}}(\nu)$  will be asymptotically distributed as a chi-square variable with  $m - (p + q + P + Q)$  degrees of freedom. Note that the lag  $m$  used in (5.22) could be chosen to be different across the seasons but in most applications it is reasonable to use the same value of  $m$  for all seasons (see for instance [25]).

Across seasons the Portmanteau test statistics are asymptotically independent for  $\nu = 1, \dots, T$ . Consequently for the case of the Portmanteau test statistics in (5.22) an overall check to test if the residuals across all the seasons are periodic white noise is given by

$$Q_m^{\text{BP}} = \sum_{\nu=1}^T Q_m^{\text{BP}}(\nu) \text{ and } Q_m^{\text{LBM}} = \sum_{\nu=1}^T Q_m^{\text{LBM}}(\nu). \quad (5.23)$$

Thus if the strong SPARMA $_T(p; q)(P, Q)$  model is adequate the global statistics (5.23) will also be asymptotically distributed as a chi-squared variable with  $mT - T(p + q + P + Q)$  degrees of freedom (see for instance [25] in the strong PARMA cases).

The statistics (5.23) (resp. (5.22)) are usually used to test the following null hypothesis

**(H0)** :  $(X_{nT+\nu})_{n \in \mathbb{Z}}$  satisfies a SPARMA $_T(p; q)(P, Q)$  representation (resp. for a fixed period  $\nu$ );

against the alternative

**(H1)** :  $(X_{nT+\nu})_{n \in \mathbb{Z}}$  does not admit a SPARMA $_T(p; q)(P, Q)$  representation (resp. for a fixed period  $\nu$ ) or  $(X_{nT+\nu})_{n \in \mathbb{Z}}$  satisfies a SPARMA $_T(p'; q')(P', Q')$  representation (resp. for a specified period  $\nu$ ) with  $p' > p$  or  $q' > q$  or  $P' > P$  or  $Q' > Q$ .

For weak SPARMA $_T(p; q)(P, Q)$  models we will shown that the asymptotic distributions of the statistics defined in (5.22) and (5.23) are no longer chi-squared distributions but a mixture of chi-squared distributions weighted by eigenvalues of the asymptotic covariance matrix of the vector of autocorrelations.

### 5.5.1 Asymptotic distribution of the residual autocorrelations

By a Taylor expansion of  $\sqrt{N}\hat{\gamma}(v, h)$  we have

$$\sqrt{N}\hat{\gamma}(v, h) = \sqrt{N}\gamma(v, h) + \left( \mathbb{E} \left[ \epsilon_{nT+v-h} \frac{\partial}{\partial \alpha'} \epsilon_{nT+v}(\alpha_0) \right] \right) \sqrt{N}(\hat{\alpha}_{LS} - \alpha_0) + o_{\mathbb{P}}(1). \quad (5.24)$$

We now remark that in Equation (5.24),  $\mathbb{E}[\epsilon_{nT+v-h}(\partial \epsilon_{nT+v}(\alpha_0)/\partial \alpha')]$  is the line  $h$  of the following matrix  $\Psi_m(v) \in \mathbb{R}^{m \times (p+q+P+Q)T}$  defined by

$$\Psi_m(v) = \mathbb{E} \left\{ \left( \begin{array}{c} \epsilon_{nT+v-1} \\ \vdots \\ \epsilon_{nT+v-m} \end{array} \right) \frac{\partial \epsilon_{nT+v}}{\partial \alpha'} \right\}. \quad (5.25)$$

So for  $h = 1, \dots, m$ , Equation (5.24) becomes

$$\sqrt{N}\hat{\gamma}_m(v) = \left( \sqrt{N}\hat{\gamma}(v, 1), \dots, \sqrt{N}\hat{\gamma}(v, m) \right)' = \sqrt{N}\gamma_m(v) + \Psi_m(v)\sqrt{N}(\hat{\alpha}_{LS} - \alpha_0) + o_{\mathbb{P}}(1). \quad (5.26)$$

From (5.26) it is clear that the asymptotic distribution of the residual autocovariances  $\sqrt{N}\hat{\gamma}_m(v)$  is related to the asymptotic behavior of  $\sqrt{N}(\hat{\alpha}'_{LS} - \alpha'_0, \gamma'_m(v))'$ .

In view of theorem 5.3.1 and (A3) we have almost surely  $\hat{\alpha} \rightarrow \alpha_0 \in \overset{\circ}{\Theta}$ . Thus  $\partial Q_N^{\sigma^2}(\hat{\alpha}_{LS})/\partial \alpha = 0$  for sufficiently large  $N$  and a Taylor expansion gives

$$\sqrt{N} \frac{\partial}{\partial \alpha} O_N^{\sigma^2}(\alpha_0) + J(\alpha_0, \sigma^2) \sqrt{N}(\hat{\alpha} - \alpha_0) = o_{\mathbb{P}}(1), \text{ where } O_N^{\sigma^2}(\alpha) = \frac{1}{2N} \sum_{n=0}^{N-1} \sum_{v=1}^T \sigma_v^{-2} \epsilon_{nT+v}^2(\alpha). \quad (5.27)$$

Consequently under the assumption that  $J(\alpha_0, \sigma^2)$  is invertible and from (5.27) we deduce that

$$\sqrt{N}(\hat{\alpha}_{LS} - \alpha_0) = -J(\alpha_0, \sigma^2)^{-1} \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \sum_{v=1}^T \sigma_v^{-2} \epsilon_{nT+v} \frac{\partial \epsilon_{nT+v}(\alpha_0)}{\partial \alpha} + o_{\mathbb{P}}(1). \quad (5.28)$$

**Proposition 5.5.1** *Assume that  $p + q + P + Q > 0$ . Under the assumptions (A0), (A1), (A2), (A3) and (A4), the random vector*

$$\sqrt{N} \left( (\hat{\alpha} - \alpha_0)', \gamma'_m(v) \right)'$$

*has a limiting centered normal distribution with covariance matrix*

$$\Xi(v) = \begin{pmatrix} \Omega & \Sigma_{\hat{\alpha}, \gamma_m(v)} \\ \Sigma'_{\hat{\alpha}, \gamma_m(v)} & \Sigma_{\gamma_m(v)} \end{pmatrix} = \sum_{h=-\infty}^{\infty} \mathbb{E} \left[ W_{nT+v} W'_{(n-h)T+v} \right],$$

*where from (5.21) and (5.28) we have*

$$W_{nT+v} = \begin{pmatrix} W_{1, nT+v} \\ W_{2, nT+v} \end{pmatrix} = \begin{pmatrix} -J(\alpha_0, \sigma^2)^{-1} \sum_{v=1}^T \sigma_v^{-2} \epsilon_{nT+v} \frac{\partial \epsilon_{nT+v}(\alpha_0)}{\partial \alpha} \\ (\epsilon_{nT+v-1}, \dots, \epsilon_{nT+v-m})' \epsilon_{nT+v} \end{pmatrix}. \quad (5.29)$$

The following theorem provides the limit distribution of the residual autocovariances and autocorrelations of weak SPARMA models for a specified season  $v$ . We also provide the limit distribution of the global residual autocovariances and autocorrelations of weak SPARMA models across all the seasons. Let

$$D_m(v) = \text{Diag}(\sigma_v \sigma_{v-1}, \dots, \sigma_v \sigma_{v-m})$$

and

$$D_{mT} = \text{Diag}(\sigma_1 \sigma_{1-1}, \dots, \sigma_1 \sigma_{1-m}, \dots, \sigma_T \sigma_{T-1}, \dots, \sigma_T \sigma_{T-m}).$$

**Theorem 5.5.1** *Let  $p + q + P + Q > 0$ . Under the assumptions of Proposition 5.5.1 we have*

$$\sqrt{N}\hat{\gamma}_m(\mathbf{v}) \xrightarrow[N \rightarrow \infty]{d} \mathcal{N}(0, \nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v})) \quad (5.30)$$

where

$$\nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v}) = \Sigma_{\gamma_m(\mathbf{v})} + \Psi_m(\mathbf{v})\Omega\Psi_m'(\mathbf{v}) + \Psi_m(\mathbf{v})\Sigma_{\hat{\alpha}, \gamma_m(\mathbf{v})} + \Sigma_{\hat{\alpha}, \gamma_m(\mathbf{v})}'\Psi_m'(\mathbf{v}).$$

Then:

$$\sqrt{N}\hat{\gamma}_{mT} \xrightarrow[N \rightarrow \infty]{d} \mathcal{N}(0, \nabla_{\hat{\gamma}}) \quad (5.31)$$

where the asymptotic covariance matrix  $\nabla_{\hat{\gamma}}$  is a block matrix, with the asymptotic variances given by  $\nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v})$ , for  $\mathbf{v} = 1, \dots, T$  and the asymptotic covariances given by

$$\nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v}') = \Sigma_{\gamma_m}(\mathbf{v}, \mathbf{v}') + \Psi_m(\mathbf{v})\Omega\Psi_m'(\mathbf{v}') + \Psi_m(\mathbf{v})\Sigma_{\hat{\alpha}, \gamma_m}(\mathbf{v}, \mathbf{v}') + \Sigma_{\hat{\alpha}, \gamma_m(\mathbf{v})}'\Psi_m'(\mathbf{v}'), \quad \mathbf{v} \neq \mathbf{v}'.$$

We also have

$$\sqrt{N}\hat{\rho}_m(\mathbf{v}) \xrightarrow[N \rightarrow \infty]{d} \mathcal{N}(0, \nabla_{\hat{\rho}}(\mathbf{v}, \mathbf{v})) \quad \text{where} \quad \nabla_{\hat{\rho}}(\mathbf{v}, \mathbf{v}) = D_m^{-1}(\mathbf{v})\nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v})D_m^{-1}(\mathbf{v}) \quad (5.32)$$

and

$$\sqrt{N}\hat{\rho}_{mT} \xrightarrow[N \rightarrow \infty]{d} \mathcal{N}(0, \nabla_{\hat{\rho}}) \quad \text{where} \quad \nabla_{\hat{\rho}} = D_{mT}^{-1}\nabla_{\hat{\gamma}}D_{mT}^{-1}. \quad (5.33)$$

The asymptotic variance matrices  $\nabla_{\hat{\gamma}}(\mathbf{v}, \mathbf{v})$  and  $\nabla_{\hat{\rho}}(\mathbf{v}, \mathbf{v})$  depend on the unknown matrices  $\Xi(\mathbf{v})$ ,  $\Psi_m(\mathbf{v})$  and the scalar  $\sigma_{v-h}$  for  $h = 1, \dots, m$ . Matrix  $\Psi_m(\mathbf{v})$  and  $\sigma_{v-h}$  can be estimated by its empirical counterpart, respectively by

$$\hat{\Psi}_m(\mathbf{v}) = \frac{1}{N} \sum_{n=0}^{N-1} \left\{ (\hat{\epsilon}_{nT+v-1}, \dots, \hat{\epsilon}_{nT+v-m})' \frac{\partial \hat{\epsilon}_{nT+v}}{\partial \alpha'} \right\} \quad \text{and} \quad \hat{\sigma}_{v-h} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} \hat{\epsilon}_{nT+v-h}^2}.$$

We use a parametric method to estimate the long-run variance  $\Xi(\mathbf{v})$  as in Section 5.4.2. We interpret  $(2\pi)^{-1}\Xi(\mathbf{v})$  as the spectral density of the multivariate stationary periodic process  $(W_{nT+v})_{n \in \mathbb{Z}} = (W'_{1, nT+v}, W'_{2, nT+v})'_{n \in \mathbb{Z}}$  evaluated at frequency 0. We obtain the estimation of the long-run variance  $\Xi(\mathbf{v})$  by replacing  $Y_n$  by  $(W_{nT+v})_{n \in \mathbb{Z}}$  in Section 5.4.2 and we have

$$\hat{\Xi}^{\text{SP}}(\mathbf{v}) := \hat{\Phi}_{v,r}^{-1}(1)\hat{\Sigma}_{\hat{u}_r}(\mathbf{v})\hat{\Phi}_{v,r}^{-1'}(1) \rightarrow \Xi(\mathbf{v}) := \Phi_v^{-1}(1)\Sigma_u(\mathbf{v})\Phi_v^{-1'}(1)$$

in probability when  $r = r(N) \rightarrow \infty$  and  $r^3/N \rightarrow 0$  as  $N \rightarrow \infty$ .

From Theorem 5.5.1 we can deduce the following result which gives the exact limiting distribution of the standard portmanteau statistics (5.22) and (5.23) under general assumptions on the innovation process of the fitted SPARMA model.

**Theorem 5.5.2** *Under Assumptions of Theorem 5.5.1 and (H0), the statistics  $Q_m^{\text{LB}}(\mathbf{v})$  and  $Q_m^{\text{LBM}}(\mathbf{v})$  defined in (5.22) converge in distribution, as  $N \rightarrow \infty$ , to*

$$Z_m^{\mathbf{v}}(\xi_m^{\mathbf{v}}) = \sum_{i=1}^m \xi_{i,m}^{\mathbf{v}} Z_i^2$$

where  $\xi_m^{\mathbf{v}} = (\xi_{1,m}^{\mathbf{v}}, \dots, \xi_{m,m}^{\mathbf{v}})'$  is the vector of the eigenvalues of the matrix  $\nabla_{\hat{\rho}}(\mathbf{v}, \mathbf{v})$  given in (5.32) and  $Z_1, \dots, Z_m$  are independent  $\mathcal{N}(0, 1)$  variables.

The asymptotic distribution of the global portmanteau test statistics (that takes into account all the seasons)  $Q_m^{\text{LB}}$  and  $Q_m^{\text{LBM}}$  defined in (5.23) is also a weighted sum of chi-square random variables:

$$Z_m(\xi_m) = \sum_{i=1}^{mT} \xi_{i,mT} Z_i^2$$

where  $\xi_{mT} = (\xi_{1,mT}, \dots, \xi_{mT,mT})'$  denotes the vector of the eigenvalues of the matrix  $\nabla_{\hat{\rho}}$  given in (5.33).

As mentioned below, Theorem 5.5.2 shows that: for the asymptotic distribution of (5.22) and (5.23) the  $\chi_{m-(p+q+P+Q)}^2$  and  $\chi_{mT-(p+q+P+Q)T}^2$  approximations are no longer valid in the framework of weak SPARMA( $p; q$ )( $P, Q$ ) models. The true asymptotic distributions depend on the periodic nuisance parameters involving  $\sigma_{v-h}$  for  $h = 1, \dots, m$  and the elements of  $\Xi(v)$ . Consequently in order to obtain the asymptotic distribution of the portmanteau statistics (5.22) (resp. (5.23)) under weak assumptions on the periodic noise, one needs a consistent estimator of the asymptotic covariance matrix  $\nabla_{\hat{\rho}}(v, v)$  (resp.  $\nabla_{\hat{\rho}}$ ). We let  $\hat{\nabla}_{\hat{\rho}}(v, v)$  (resp.  $\hat{\nabla}_{\hat{\rho}}$ ) the matrix obtained by replacing  $\Xi(v)$  by  $\hat{\Xi}(v)$ ,  $\Psi_m(v)$  by  $\hat{\Psi}_m(v)$  and  $\sigma_{v-h}$  by  $\hat{\sigma}_{v-h}$  in  $\nabla_{\hat{\rho}}(v, v)$  (resp.  $\nabla_{\hat{\rho}}$ ). Denote by  $\hat{\xi}_m^v = (\hat{\xi}_{1,m}^v, \dots, \hat{\xi}_{m,m}^v)'$  (resp.  $\hat{\xi}_m = (\hat{\xi}_{1,mT}, \dots, \hat{\xi}_{mT,mT})'$ ) the vector of the eigenvalues of  $\hat{\nabla}_{\hat{\rho}}(v, v)$  (resp.  $\hat{\nabla}_{\hat{\rho}}$ ). At the asymptotic level  $\alpha$  and for a specified season  $v$ , the LBM test (resp. the BP test) consists in rejecting the adequacy of the weak SPARMA<sub>T</sub>( $p; q$ )( $P, Q$ ) model when

$$\lim_{N \rightarrow \infty} \mathbb{P}(Q_m^{\text{BP}}(v) > S_m^v(1 - \alpha)) = \lim_{N \rightarrow \infty} \mathbb{P}(Q_m^{\text{LBM}}(v) > S_m^v(1 - \alpha)) = \alpha,$$

where  $S_m^v(1 - \alpha)$  is such that  $\mathbb{P}\{Z_m^v(\hat{\xi}_m^v) > S_m^v(1 - \alpha)\} = \alpha$ . We emphasize the fact that the proposed modified versions of the BP and LBM statistics are more difficult to implement because their critical values have to be computed from the data while the critical values of the standard method are simply deduced from a  $\chi^2$ -table. We shall evaluate the  $p$ -values

$$\mathbb{P}\{Z_m(\hat{\xi}_m^v) > Q_m^{\text{BP}}(v)\} \text{ and } \mathbb{P}\{Z_m(\hat{\xi}_m^v) > Q_m^{\text{LBM}}(v)\}, \text{ with } Z_m(\hat{\xi}_m^v) = \sum_{i=1}^m \hat{\xi}_{i,m}^v Z_i^2,$$

by means of the Imhof algorithm (see [27]) or other exact methods.

The test procedures for  $Q_m^{\text{BP}}$  and  $Q_m^{\text{LBM}}$  defined in (5.23) are similar but they are based on the  $mT$  empirical eigenvalues of  $\hat{\nabla}_{\hat{\rho}}$ .

## 5.5.2 Self-normalized asymptotic distribution of the residual autocorrelations

The nonparametric kernel estimator (see [3, 39]) used to estimate the matrix  $\Xi(v)$  causes serious difficulties regarding the choice of the sequence of weights. The parametric approach based on an autoregressive estimate of the spectral density of  $(W_{nT+v})_{n \in \mathbb{Z}}$  studied in [8, 20] is also facing the problem of choosing the truncation parameter. So the choice of the order of truncation is often crucial and difficult.

In this section, we propose an alternative method where we do not estimate an asymptotic covariance matrix which is an extension to the results obtained by [16]. It is based on a self-normalization approach to construct a test-statistic which is asymptotically distribution-free under the null hypothesis. This approach has been studied by [16] in the weak ARMA case by proposing new portmanteau statistics. In this case the critical values are not computed from the data since they are tabulated by [32]. In some sense this method is finally closer to the standard method in which the critical values are simply deduced from a  $\mathcal{X}^2$ -table. The idea comes from [32] and has been already extended by [16], [30], [46], [47] and [48] to name a few in more general frameworks. See also [49] for a review on some recent developments on the inference of time series data using the self-normalized approach.

Let  $(B_K(r))_{r \geq 0}$  be a  $K$ -dimensional Brownian motion starting from 0. For  $K \geq 1$ , we denote  $\mathcal{U}_K$  the random variable defined by

$$\mathcal{U}_K = B_K'(1) V_K^{-1} B_K(1) \text{ where } V_K = \int_0^1 (B_K(r) - r B_K(1)) (B_K(r) - r B_K(1))' dr.$$

We denote  $\Lambda(v)$  the matrix in  $\mathbb{R}^{m \times (T(p+q+P+Q)+m)}$  defined in block formed by  $\Lambda(v) = (\Psi_m(v) | I_m)$  where  $I_k$  is the identity matrix of order  $k$ . In view of (5.26) and (5.28) we deduce that  $\sqrt{N} \hat{\gamma}_m(v) = N^{-1/2} \sum_{n=1}^{N-1} \Lambda(v) W_{nT+v} + o_{\mathbb{P}}(1)$ .

Contrary to Subsection 5.5.1 we do not rely on the classical method that would consist in estimating the asymptotic covariance matrix of  $(\Lambda(v)W_{nT+v})_{n \in \mathbb{Z}}$ . We need to apply the functional central limit theorem holds for the process  $(W_{nT+v})_{n \in \mathbb{Z}}$  (see Lemma 1 in [32]).

Finally we define the normalization matrix  $C_m(v, v) \in \mathbb{R}^{m \times m}$  for a specified season  $v$  by

$$C_m(v, v) = \frac{1}{N^2} \sum_{n=0}^{N-1} S_n(v) S_n'(v) \text{ where } S_n(v) = \sum_{j=0}^n (\Lambda(v)W_{jT+v} - \gamma_m(v)). \quad (5.34)$$

Similarly we denote by  $C_{mT} \in \mathbb{R}^{mT \times mT}$  the global normalization block matrix across all the seasons with diagonal block given by (5.34) for  $v = 1, \dots, T$  and

$$C_m(v, v') = \frac{1}{N^2} \sum_{n=0}^{N-1} S_n(v) S_n'(v') \text{ where } S_n(v') = \sum_{j=0}^n (\Lambda(v')W_{jT+v'} - \gamma_m(v')) \text{ for } v \neq v'.$$

To ensure the invertibility of the normalization matrix  $C_m(v, v)$  we need the following technical assumption on the distribution of  $\epsilon_n$ .

**(A5):** The process  $(\epsilon_n)_{n \in \mathbb{Z}}$  has a positive density on some neighborhood of zero.

The following proposition gives the invertibility of the normalization matrices  $C_m(v, v)$  and  $C_{mT}$ .

**Proposition 5.5.2** *Under the assumptions of Theorem 5.5.1 and (A5), the matrices  $C_m(v, v)$  and  $C_{mT}$  are almost surely non singular.*

The following theorem states the asymptotic distributions of the residual autocovariances and autocorrelations of weak SPARMA models for a specified season  $v$ . We also provide the limit distribution of the global residual autocovariances and autocorrelations of weak SPARMA models across all the seasons.

**Theorem 5.5.3** *We assume that  $p + q + P + Q > 0$ . Under Assumptions of Theorem 5.5.1, (A5) and under the null hypothesis (H0) we have for a specified season  $v$*

$$N\hat{\gamma}'_m(v)C_m^{-1}(v, v)\hat{\gamma}_m(v) \xrightarrow[N \rightarrow \infty]{d} \mathcal{U}_m \text{ and } N\hat{\rho}'_m(v)D_m(v)C_m^{-1}(v, v)D_m(v)\hat{\rho}_m(v) \xrightarrow[N \rightarrow \infty]{d} \mathcal{U}_m.$$

For the global residual autocovariances and autocorrelations across all the seasons we also have

$$N\hat{\gamma}'_{mT}C_{mT}^{-1}\hat{\gamma}_{mT} \xrightarrow[N \rightarrow \infty]{d} \mathcal{U}_{mT} \text{ and } N\hat{\rho}'_{mT}D_{mT}C_{mT}^{-1}D_{mT}\hat{\rho}_{mT} \xrightarrow[N \rightarrow \infty]{d} \mathcal{U}_{mT}.$$

Of course the above theorem is useless for practical purpose because the normalization matrix  $C_m(v, v)$ ,  $C_{mT}$  and the parameter  $\sigma_{v-h}$  are not observable. This gap will be fixed below (see Theorem 5.5.4) when one replaces the matrix  $C_m(v, v)$ ,  $C_{mT}$  and the scalar  $\sigma_{v-h}$  by their empirical or observable counterparts. Then we denote

$$\hat{C}_m(v, v) = \frac{1}{N^2} \sum_{n=0}^{N-1} \hat{S}_n(v) \hat{S}_n'(v) \text{ where } \hat{S}_n(v) = \sum_{j=0}^n (\hat{\Lambda}(v)\hat{W}_{jT+v} - \hat{\gamma}_m(v)),$$

with  $\hat{\Lambda}(v) = (\hat{\Psi}_m(v)|I_m)$  and where  $\hat{W}_{nT+v}$  and  $\hat{\sigma}_{v-h}$  are defined in Subsection 5.5.1.

We are able to state the following result which is the applicable counterpart of Theorem 5.5.3.

**Theorem 5.5.4** *Under the assumptions of Theorem 5.5.3, for a specified season  $v$  we have*

$$N\hat{\gamma}'_m(v)\hat{C}_m^{-1}(v, v)\hat{\gamma}_m(v) \xrightarrow[N \rightarrow \infty]{d} \mathcal{U}_m \text{ and } Q_m^{\text{SN}}(v) =: N\hat{\rho}'_m(v)\hat{D}_m(v)\hat{C}_m^{-1}(v, v)\hat{D}_m(v)\hat{\rho}_m(v) \xrightarrow[N \rightarrow \infty]{d} \mathcal{U}_m.$$

We also have for the global seasons

$$N\hat{\gamma}'_{mT}\hat{C}_{mT}^{-1}\hat{\gamma}_{mT} \xrightarrow[N \rightarrow \infty]{d} \mathcal{U}_{mT} \text{ and } Q_m^{\text{SN}} =: N\hat{\rho}'_{mT}\hat{D}_{mT}\hat{C}_{mT}^{-1}\hat{D}_{mT}\hat{\rho}_{mT} \xrightarrow[N \rightarrow \infty]{d} \mathcal{U}_{mT}.$$



Based on the above result we propose a modified version of the LBM statistic for a specified season  $\nu$  when one uses the statistic

$$\tilde{Q}_m^{\text{SN}}(\nu) = N\hat{\rho}'_m(\nu)M_m^{1/2}(\nu, N, T)\hat{D}_m(\nu)\hat{C}_m^{-1}(\nu, \nu)\hat{D}_m(\nu)M_m^{1/2}(\nu, N, T)\hat{\rho}_m(\nu), \quad (5.35)$$

where the matrix  $M_m(\nu, N, T) \in \mathbb{R}^{m \times m}$  is diagonal with  $(\varrho(1, \nu, N, T), \dots, \varrho(m, \nu, N, T))$  as diagonal terms.

Similarly we propose a modified version of the LBM statistic for the global seasons when one uses the following statistic

$$\tilde{Q}_m^{\text{SN}} = N\hat{\rho}'_{mT}M_{mT}^{1/2}\hat{D}_{mT}\hat{C}_{mT}^{-1}\hat{D}_{mT}M_{mT}^{1/2}\hat{\rho}_{mT}, \quad (5.36)$$

where the matrix  $M_{mT} \in \mathbb{R}^{mT \times mT}$  is diagonal with

$$(\varrho(1, 1, N, T), \dots, \varrho(m, 1, N, T), \dots, \varrho(1, T, N, T), \dots, \varrho(m, T, N, T))$$

as diagonal terms.

At the asymptotic level  $\alpha$  and for a specified season  $\nu$ , the self-normalized LBM test (resp. the BP test) consists in rejecting the adequacy of the weak SPARMA $_T(p; q)(P; Q)$  model when

$$\tilde{Q}_m^{\text{SN}}(\nu) > \mathcal{U}_m(1 - \alpha) \quad (\text{resp. } Q_m^{\text{SN}}(\nu) > \mathcal{U}_m(1 - \alpha)),$$

where the critical values  $\mathcal{U}_m(1 - \alpha)$  are tabulated in [32, Table 1]. Similarly for the global test we have

$$\tilde{Q}_m^{\text{SN}} > \mathcal{U}_{mT}(1 - \alpha) \quad (\text{resp. } Q_m^{\text{SN}} > \mathcal{U}_{mT}(1 - \alpha)).$$

## 5.6 Numerical illustration

By means of Monte Carlo experiments, we investigate the finite sample properties of the tests introduced in this paper. The numerical illustrations of this section are made with the free statistical R software (see <https://www.r-project.org/>).

We indicate the conventions that we adopt in the discussion and in the tables. One refers to

- $\text{BP}_s(\nu)$  for the standard Box-Pierce test using the statistic  $Q_m^{\text{BP}}(\nu)$ ;
- $\text{LBM}_s(\nu)$  for the standard Ljung-Box-McLeod test using the statistic  $Q_m^{\text{LBM}}(\nu)$ ;
- $\text{BP}_{\text{SN}}(\nu)$  for the modified test using the statistic  $Q_m^{\text{SN}}(\nu)$  with the values of the quantiles of  $\mathcal{U}_m$  simulated in Table 1 of [32];
- $\text{LBM}_{\text{SN}}(\nu)$  for the modified test using the statistic  $\tilde{Q}_m^{\text{SN}}(\nu)$  with the values of the quantiles of  $\mathcal{U}_m$ ;
- $\text{BP}_w(\nu)$  for the modified Box-Pierce test (proposed in Section 5.5.1) using the statistic  $Q_m^{\text{BP}}(\nu)$ ;
- $\text{LBM}_w(\nu)$  for the modified Ljung-Box-McLeod test (see Section 5.5.1) using the statistic  $Q_m^{\text{LBM}}(\nu)$ ;
- $\text{BP}_s$  for the standard global Box-Pierce test using the statistic  $Q_m^{\text{BP}}$ ;
- $\text{LBM}_s$  for the standard global Ljung-Box-McLeod test using the statistic  $Q_m^{\text{LBM}}$ ;
- $\text{BP}_{\text{SN}}$  for the modified global test using the statistic  $Q_m^{\text{SN}}$  with the values of the quantiles of  $\mathcal{U}_{mT}$ ;
- $\text{LBM}_{\text{SN}}$  for the modified global test using the statistic  $\tilde{Q}_m^{\text{SN}}$  with the values of the quantiles of  $\mathcal{U}_{mT}$ ;



- $\text{BP}_W$  for the modified global Box-Pierce test (proposed in Section 5.5.1) using the statistic  $Q_m^{\text{BP}}$ ;
- $\text{LBM}_W$  for the modified global Ljung-Box-McLeod test (see Section 5.5.1) using the statistic  $Q_m^{\text{LBM}}$ .

We will see in the tables that the numerical results using the Ljung-Box-McLeod tests are very close to those of the Box-Pierce tests. Nevertheless they are still presented here for the sake of completeness.

We study numerically the behavior of the quasi-generalized least squares (QLS) estimator for a  $\text{SPARMA}_2(1, 1)(1, 1)$  models of the form

$$\begin{aligned} X_{2n+v} &= a_1(v)X_{2n+v-1} + \mathbf{a}_1(v)X_{2(n-1)+v} + \mathbf{a}_1(v)a_1(v)X_{2(n-1)+v-1} - b_1(v)\epsilon_{2n+v-1} \\ &\quad - \mathbf{b}_1(v)\epsilon_{2(n-1)+v} + \mathbf{b}_1(v)b_1(v)\epsilon_{2(n-1)+v-1} + \epsilon_{2n+v}, \quad v = 1, 2 \end{aligned} \quad (5.37)$$

where the unknown parameter is  $\alpha_0 = (a_1(1), a_1(2), b_1(1), b_1(2), \mathbf{a}_1(1), \mathbf{a}_1(2), \mathbf{b}_1(1), \mathbf{b}_1(2))'$ . We take  $\alpha_0 = (0.8, 0.9, -0.5, -0.45, 0.6, 0.5, -0.7, -0.6)'$ . Three different periodic noises are considered. First we assume that in (5.37) the innovation process  $(\epsilon_{2n+v})_{n \in \mathbb{Z}}$  is an iid centered Gaussian process with variance  $\sigma_v^2 = \mathbb{E}(\epsilon_{2n+v}^2)$  where  $\sigma_1 = 0.9$  and  $\sigma_2 = 1.5$  which corresponds to the strong SPARMA case. For the weak SPARMA cases, we consider that in (5.37) the innovation process  $(\epsilon_{2n+v})_{n \in \mathbb{Z}}$  follows firstly a periodic ARCH(1) given by

$$\epsilon_{2n+v} = \sqrt{\omega(v) + \alpha(v)\epsilon_{2n+v-1}^2} \eta_{2n+v}, \quad v = 1, 2 \quad (5.38)$$

with  $(\omega(1), \omega(2)) = (0.2, 0.4)$ ,  $(\alpha(1), \alpha(2)) = (0.4, 0.45)$  and where  $(\eta_{2n+v})_{n \in \mathbb{Z}}$  is a sequence of iid centered Gaussian random variables with variance 1. Secondly a noise defined by a following PGARCH(1, 1)

$$\begin{cases} \epsilon_{2n+v} = \sqrt{h_{2n+v}} \eta_{2n+v} \\ h_{2n+v} = \omega(v) + \alpha(v)\epsilon_{2n+v-1}^2 + \beta(v)h_{2n+v-1}, \quad v = 1, 2 \end{cases} \quad (5.39)$$

where  $(\omega(1), \omega(2)) = (0.2, 0.4)$ ,  $(\alpha(1), \alpha(2)) = (0.1, 0.15)$  and  $(\beta(1), \beta(2)) = (0.73, 0.72)$ .

### 5.6.1 Empirical size

For each of these three different periodic white noises 1,000 independent replications of length  $(N + 100) \times 2$ . These sequences were plugged into the model (5.37) yielding 1,000 independent replications of the periodic process  $(X_{2n+v})_n$  of length  $(N + 100) \times 2$ . Initial values were set to zero and in order to achieve periodic stationarity the first 200 observations were dropped. For each replication of size  $NT = N \times 2$  we use the QLS method to estimate the corresponding coefficient  $\alpha_0$ . After estimating Model (5.37) we apply portmanteau test to the residuals for different values of  $m \in \{1, \dots, 12\}$ , where  $m$  is the number of autocorrelations used in the portmanteau test statistic. In the experiment we considered two values of  $N$ : 400 and 2,000. The nominal asymptotic level considered of the tests is  $\alpha = 5\%$ . Note that for this level, the empirical relative frequency of rejection size over the 1,000 independent replications should vary within the confidences intervals [3.6%, 6.4%] with probability 95% and [3.3%, 6.9%] with probability 99%.

#### Strong $\text{SPARMA}_2(1, 1)(1, 1)$ model case

Table 5.1 displays relative frequencies (in %) of rejection (over the 1,000 replications) of the null hypothesis ( $\mathbf{H}_0$ ) that the data generating process (DGP for short) follows a strong  $\text{SPARMA}_2(1, 1)(1, 1)$  given by (5.37). Note that the empirical size is not available (n.a.) for the standard BP and LBM tests because these tests are not applicable to  $m \leq 4$  contrary to the proposed tests. As expected the observed relative rejection frequency of the standard tests is very far from the

nominal  $\alpha = 5\%$  for small  $m$ . The results are worse for  $N = 2,000$  than for  $N = 400$ . This is in accordance with the results in the literature on the strong SPARMA models. The theory that the asymptotic distributions of (5.22) and (5.23), namely,  $\chi_{m-(p+q=P+Q)}^2$  and  $\chi_{2m-2(p+q+P+Q)}^2$  approximations are better for larger  $m$  is confirmed. In contrast, the proposed tests well control the error of first kind, even when  $m$  is small (for global or a fixed season  $v$ ).

From this example we draw the conclusion that the proposed versions are preferable to the standard ones, when the number  $m$  of autocorrelations used is small.

#### **Weak SPARMA<sub>2</sub>(1, 1)(1, 1) model case**

We now repeat the same experiment on Model (5.37) by assuming that the innovation process  $(\epsilon_{2n+v})_n$  follows first (5.38) and secondly (5.39).

Tables 5.2 and 5.3 display relative frequencies (in %) of rejection (over the 1,000 replications) of the null hypothesis (**H0**) that the DGP follows two weak SPARMA<sub>2</sub>(1, 1)(1, 1) given by (5.37)–(5.38) and (5.37)–(5.39). Note also that the empirical size is not available (n.a.) for the standard BP and LBM tests because these tests are not applicable to  $m \leq 4$ . As expected, Tables 5.2 and 5.3 show that the standard tests poorly perform to assess the adequacy of these weak SPARMA models. The (global or specified season  $v$ ) observed relative rejection frequencies of the standard tests are definitely outside the significant limits. Thus the standard tests reject very often the true weak SPARMA<sub>2</sub>(1, 1)(1, 1) models. By contrast, the error of first kind is globally well controlled by the proposed tests when  $N$  increases. We draw the conclusion that for these particular weak SPARMA models the proposed tests are clearly preferable to the standard ones.

Table 5.1 – Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a strong SPARMA<sub>2</sub>(1, 1)(1, 1) given by (5.37) where the innovation  $\epsilon_{nT+v} \sim \mathcal{N}(0, \sigma_v^2)$ .

Length N	Tests	Lag $m$											
		1	2	3	4	5	6	7	8	9	10	11	12
N = 400	BP <sub>S</sub> (1)	n.a.	n.a.	n.a.	n.a.	32.6	23.6	19.1	16.5	15.6	15.1	13.9	14.0
	BP <sub>S</sub> (2)	n.a.	n.a.	n.a.	n.a.	32.4	22.5	18.5	15.9	14.7	13.6	12.9	12.1
	LBM <sub>S</sub> (1)	n.a.	n.a.	n.a.	n.a.	32.8	24.0	19.7	16.6	15.8	15.7	14.5	14.6
	LBM <sub>S</sub> (2)	n.a.	n.a.	n.a.	n.a.	32.7	23.0	19.1	16.1	14.8	14.1	13.2	12.6
	BP <sub>W</sub> (1)	4.4	4.7	4.2	3.9	3.3	3.1	2.4	3.0	2.7	3.2	2.8	2.5
	BP <sub>W</sub> (2)	5.3	4.2	5.6	5.2	4.4	3.8	4.0	3.6	3.4	3.0	2.7	2.1
	LBM <sub>W</sub> (1)	4.4	4.8	4.2	3.9	3.4	3.1	2.6	3.1	2.8	3.3	3.2	2.5
	LBM <sub>W</sub> (2)	5.3	4.4	5.6	5.2	4.4	3.8	4.0	3.6	3.7	3.2	2.9	2.3
	BP <sub>SN</sub> (1)	4.7	4.0	4.4	4.7	5.3	4.5	5.4	5.8	6.8	6.3	6.2	6.0
	BP <sub>SN</sub> (2)	4.7	4.4	4.7	5.2	4.2	5.3	4.6	4.3	5.0	4.6	5.9	5.5
	LBM <sub>SN</sub> (1)	4.7	4.0	4.7	4.9	5.5	4.7	5.5	5.8	6.9	6.4	6.3	6.2
	LBM <sub>SN</sub> (2)	4.7	4.5	4.7	5.3	4.3	5.5	4.7	4.3	5.3	4.8	5.9	5.7
	BP <sub>S</sub>	n.a.	n.a.	n.a.	n.a.	50.8	35.8	31.1	26.2	23.0	21.4	19.5	19.6
	LBM <sub>S</sub>	n.a.	n.a.	n.a.	n.a.	50.9	36.0	31.7	26.8	24.0	22.6	20.3	20.8
	BP <sub>W</sub>	4.2	4.1	5.0	4.3	2.6	2.6	2.5	2.3	2.2	1.7	1.1	0.6
	LBM <sub>W</sub>	4.2	4.2	5.0	4.7	2.7	2.9	2.5	2.3	2.3	1.8	1.1	0.7
BP <sub>SN</sub>	4.3	5.0	4.1	4.4	4.6	5.2	4.2	4.1	3.7	3.5	3.8	4.0	
LBM <sub>SN</sub>	4.4	5.0	4.1	4.6	4.8	5.4	4.5	4.3	3.8	3.7	4.3	4.2	
N = 2,000		1	2	3	4	5	6	7	8	9	10	11	12
	BP <sub>S</sub> (1)	n.a.	n.a.	n.a.	n.a.	32.0	19.5	17.1	13.8	12.7	12.3	12.7	11.3
	BP <sub>S</sub> (2)	n.a.	n.a.	n.a.	n.a.	32.1	22.1	17.5	15.8	14.1	12.7	12.4	10.9
	LBM <sub>S</sub> (1)	n.a.	n.a.	n.a.	n.a.	32.0	19.5	17.3	14.0	12.9	12.3	12.8	11.6
	LBM <sub>S</sub> (2)	n.a.	n.a.	n.a.	n.a.	32.2	22.1	17.6	15.9	14.2	12.8	12.4	11.1
	BP <sub>W</sub> (1)	4.3	4.2	5.8	4.7	4.5	4.0	4.3	4.7	3.9	3.6	4.6	4.6
	BP <sub>W</sub> (2)	3.8	4.1	4.0	4.9	4.5	4.2	4.8	4.7	4.7	4.1	3.9	2.9
	LBM <sub>W</sub> (1)	4.3	4.2	5.8	4.7	4.6	4.0	4.3	4.8	3.9	3.6	4.6	4.6
	LBM <sub>W</sub> (2)	3.8	4.2	4.0	4.9	4.5	4.4	4.9	4.9	4.7	4.1	3.9	3.0
	BP <sub>SN</sub> (1)	4.5	4.5	4.2	3.7	4.3	4.0	5.1	5.7	4.8	5.6	5.5	5.5
	BP <sub>SN</sub> (2)	4.9	4.5	4.5	4.8	4.8	4.4	5.3	5.7	6.3	5.3	5.0	4.4
	LBM <sub>SN</sub> (1)	4.5	4.6	4.2	3.7	4.3	4.0	5.2	5.8	4.8	5.7	5.6	5.5
	LBM <sub>SN</sub> (2)	4.9	4.5	4.5	4.9	4.8	4.6	5.3	5.7	6.3	5.3	5.1	4.4
	BP <sub>S</sub>	n.a.	n.a.	n.a.	n.a.	51.0	33.8	27.4	24.6	20.4	18.8	18.0	16.3
	LBM <sub>S</sub>	n.a.	n.a.	n.a.	n.a.	51.2	33.9	27.5	24.9	20.5	18.8	18.2	16.3
	BP <sub>W</sub>	3.5	4.1	3.4	3.9	3.7	3.5	3.9	4.0	3.9	3.6	3.7	3.1
LBM <sub>W</sub>	3.5	4.1	3.4	3.9	3.8	3.5	3.9	4.1	3.9	3.7	3.7	3.1	
BP <sub>SN</sub>	4.7	4.7	3.7	4.2	4.4	5.4	5.4	5.0	4.3	4.7	5.1	5.4	
LBM <sub>SN</sub>	4.7	4.7	3.7	4.2	4.4	5.5	5.4	5.1	4.4	4.8	5.3	5.5	

Table 5.2 – Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a weak SPARMA<sub>2</sub>(1, 1)(1, 1) given by (5.37) where the innovation  $\epsilon_{nT+v} \sim (5.38)$ .

Length N	Tests	Lag $m$											
		1	2	3	4	5	6	7	8	9	10	11	12
N = 400	BP <sub>S</sub> (1)	n.a.	n.a.	n.a.	n.a.	35.8	24.1	21.8	18.0	16.9	16.0	14.7	13.7
	BP <sub>S</sub> (2)	n.a.	n.a.	n.a.	n.a.	35.1	23.6	19.5	17.5	15.9	14.7	13.2	12.7
	LBM <sub>S</sub> (1)	n.a.	n.a.	n.a.	n.a.	36.2	24.3	22.5	18.4	17.6	16.4	15.2	14.2
	LBM <sub>S</sub> (2)	n.a.	n.a.	n.a.	n.a.	35.2	24.0	19.7	17.9	16.2	15.3	13.9	13.4
	BP <sub>W</sub> (1)	4.5	4.4	5.1	4.4	4.6	3.6	3.1	2.6	2.3	2.3	2.4	2.4
	BP <sub>W</sub> (2)	5.1	5.5	5.9	5.0	4.0	3.6	3.5	2.2	1.9	1.7	1.4	1.4
	LBM <sub>W</sub> (1)	4.5	4.5	5.1	4.5	4.6	3.6	3.2	2.7	2.5	2.4	2.4	2.5
	LBM <sub>W</sub> (2)	5.1	5.5	5.9	5.1	4.2	3.7	3.7	2.4	1.9	2.0	1.5	1.4
	BP <sub>SN</sub> (1)	4.6	5.0	5.2	4.9	4.9	5.2	5.4	4.4	4.5	4.1	4.8	4.0
	BP <sub>SN</sub> (2)	4.5	3.6	3.6	4.8	4.6	4.1	4.5	5.2	5.3	4.8	4.7	4.1
	LBM <sub>SN</sub> (1)	4.6	5.0	5.4	4.9	5.1	5.3	5.5	4.4	4.6	4.2	5.0	4.3
	LBM <sub>SN</sub> (2)	4.5	3.6	3.6	4.9	4.8	4.1	4.6	5.4	5.3	5.0	4.7	4.4
	BP <sub>S</sub>	n.a.	n.a.	n.a.	n.a.	53.6	39.8	32.7	27.2	24.3	23.8	21.8	19.3
	LBM <sub>S</sub>	n.a.	n.a.	n.a.	n.a.	53.8	40.0	33.1	28.3	25	24.6	22.7	20.7
	BP <sub>W</sub>	4.1	4.4	4.2	2.6	2.8	2.4	2.2	1.9	2.0	1.8	1.8	1.3
	LBM <sub>W</sub>	4.1	4.4	4.3	2.8	2.8	2.6	2.3	2.1	2.0	1.9	1.9	1.4
	BP <sub>SN</sub>	3.7	3.5	3.9	3.1	4.1	3.6	3.4	2.8	2.8	1.8	1.7	1.7
	LBM <sub>SN</sub>	3.7	3.5	4.1	3.1	4.1	3.6	3.7	2.8	2.9	1.8	1.9	1.9
N = 2,000		1	2	3	4	5	6	7	8	9	10	11	12
	BP <sub>S</sub> (1)	n.a.	n.a.	n.a.	n.a.	39.1	26.4	22.8	19.5	17.8	17.7	17.4	16.2
	BP <sub>S</sub> (2)	n.a.	n.a.	n.a.	n.a.	33.0	23.7	19.5	17.0	15.2	14.9	12.0	11.6
	LBM <sub>S</sub> (1)	n.a.	n.a.	n.a.	n.a.	39.2	26.4	22.9	19.6	18.0	17.7	17.5	16.3
	LBM <sub>S</sub> (2)	n.a.	n.a.	n.a.	n.a.	33.1	23.8	19.6	17.1	15.4	15.0	12.1	11.6
	BP <sub>W</sub> (1)	7.3	5.7	6.1	6.3	6.9	5.8	5.9	5.8	5.5	6.0	6.1	5.4
	BP <sub>W</sub> (2)	6.2	5.6	5.2	4.3	4.4	4.6	4.0	3.5	4.0	3.4	3.0	2.9
	LBM <sub>W</sub> (1)	7.3	5.7	6.1	6.4	6.9	5.8	5.9	5.8	5.5	6.1	6.1	5.4
	LBM <sub>W</sub> (2)	6.2	5.6	5.2	4.3	4.4	4.6	4.0	3.6	4.0	3.4	3.1	2.9
	BP <sub>SN</sub> (1)	6.3	5.8	6.3	5.4	6.2	6.2	5.6	5.6	5.3	6.2	5.0	5.4
	BP <sub>SN</sub> (2)	5.9	5.7	5.2	5.1	5.3	5.2	4.9	6.0	5.6	5.5	5.8	5.9
	LBM <sub>SN</sub> (1)	6.3	5.8	6.3	5.4	6.2	6.2	5.6	5.6	5.4	6.2	5.0	5.4
	LBM <sub>SN</sub> (2)	5.9	5.7	5.2	5.2	5.3	5.2	4.9	6.0	5.6	5.5	5.8	5.9
	BP <sub>S</sub>	n.a.	n.a.	n.a.	n.a.	56.1	41.1	33.4	29.7	26.3	23.3	21.3	20.0
	LBM <sub>S</sub>	n.a.	n.a.	n.a.	n.a.	56.2	41.1	33.6	29.9	26.5	23.4	21.5	20.1
	BP <sub>W</sub>	6.2	5.9	5.5	4.8	4.8	4.4	4.3	4.0	3.9	3.9	3.6	3.9
	LBM <sub>W</sub>	6.2	6.1	5.5	4.8	4.8	4.5	4.4	4.0	4.0	3.9	3.6	3.9
	BP <sub>SN</sub>	6.7	5.2	5.3	5.1	5.0	4.6	4.8	5.3	5.6	4.7	4.7	4.6
LBM <sub>SN</sub>	6.7	5.2	5.3	5.1	5.0	4.6	4.9	5.3	5.6	4.7	4.8	4.6	

Table 5.3 – Empirical size of the standard and proposed tests: relative frequencies (in %) of rejection of a weak SPARMA<sub>2</sub>(1, 1)(1, 1) given by (5.37) where the innovation  $\epsilon_{nT+v} \sim (5.39)$ .

Length N	Tests	Lag $m$											
		1	2	3	4	5	6	7	8	9	10	11	12
N = 400	BP <sub>S</sub> (1)	n.a.	n.a.	n.a.	n.a.	40.8	30.7	26.4	22.0	20.9	19.8	18.3	19.2
	BP <sub>S</sub> (2)	n.a.	n.a.	n.a.	n.a.	39.9	29.4	24.6	22.5	20.8	19.4	17.5	17.2
	LBM <sub>S</sub> (1)	n.a.	n.a.	n.a.	n.a.	41.0	30.9	26.6	22.4	21.6	20.3	19.0	19.7
	LBM <sub>S</sub> (2)	n.a.	n.a.	n.a.	n.a.	40.4	29.9	24.9	22.9	21.6	19.9	18.3	18.0
	BP <sub>W</sub> (1)	7.1	5.7	6.3	5.7	4.8	4.7	4.2	4.5	3.6	3.5	3.5	3.1
	BP <sub>W</sub> (2)	6.9	5.7	6.9	6.4	5.1	5.1	5.1	4.8	4.0	3.6	3.4	3.3
	LBM <sub>W</sub> (1)	7.1	5.7	6.4	5.8	4.9	4.8	4.4	4.5	3.8	3.5	3.6	3.3
	LBM <sub>W</sub> (2)	6.9	5.7	6.9	6.5	5.1	5.2	5.2	5.1	4.1	3.9	3.5	3.7
	BP <sub>SN</sub> (1)	6.9	5.4	5.5	5.7	5.4	6.4	6.0	6.9	6.3	6.6	6.5	5.7
	BP <sub>SN</sub> (2)	7.0	6.3	6.0	6.2	7.0	5.8	6.6	6.5	6.3	5.9	4.7	5.4
	LBM <sub>SN</sub> (1)	6.9	5.5	5.5	5.7	5.4	6.4	6.1	7.1	6.4	6.6	6.6	5.9
	LBM <sub>SN</sub> (2)	7.0	6.3	6.0	6.3	7.0	6.0	6.8	6.7	6.3	6.1	5.0	5.5
	BP <sub>S</sub>	n.a.	n.a.	n.a.	n.a.	62.2	45.3	39.0	34.6	29.7	29.3	27.3	25.7
	LBM <sub>S</sub>	n.a.	n.a.	n.a.	n.a.	62.2	45.9	39.7	36.2	30.4	30.3	28.4	26.9
	BP <sub>W</sub>	6.3	4.6	5.7	5.0	4.4	3.7	3.2	3.3	3.0	2.7	2.8	2.6
	LBM <sub>W</sub>	6.3	4.7	5.9	5.0	4.4	3.8	3.3	3.4	3.0	2.8	2.9	2.6
	BP <sub>SN</sub>	5.9	5.5	6.1	6.3	4.9	4.2	4.8	4.7	4.8	5.0	4.1	3.7
	LBM <sub>SN</sub>	5.9	5.5	6.1	6.4	5.0	4.3	4.9	4.9	5.0	5.1	4.4	3.9
N = 2,000		1	2	3	4	5	6	7	8	9	10	11	12
	BP <sub>S</sub> (1)	n.a.	n.a.	n.a.	n.a.	53.8	44.2	41.8	39.9	39.6	39.1	38.3	38.5
	BP <sub>S</sub> (2)	n.a.	n.a.	n.a.	n.a.	52.8	41.4	37.8	35.9	35.7	34.5	33.3	32.0
	LBM <sub>S</sub> (1)	n.a.	n.a.	n.a.	n.a.	53.8	44.3	41.8	39.9	39.8	39.4	38.5	38.5
	LBM <sub>S</sub> (2)	n.a.	n.a.	n.a.	n.a.	52.8	41.4	37.8	36.0	35.8	34.5	33.5	32.1
	BP <sub>W</sub> (1)	7.6	7.0	6.9	6.1	6.4	5.5	4.8	5.2	4.1	4.4	5.1	4.4
	BP <sub>W</sub> (2)	6.7	6.1	6.1	5.8	5.2	4.9	4.3	4.1	4.1	4.0	3.6	3.3
	LBM <sub>W</sub> (1)	7.6	7.0	6.9	6.1	6.4	5.5	4.8	5.2	4.2	4.4	5.2	4.5
	LBM <sub>W</sub> (2)	6.7	6.1	6.1	5.8	5.2	4.9	4.3	4.1	4.1	4.0	3.7	3.3
	BP <sub>SN</sub> (1)	7.1	6.7	6.9	6.4	6.3	5.9	5.8	5.5	4.9	5.0	4.3	4.1
	BP <sub>SN</sub> (2)	5.6	4.8	4.8	5.1	4.7	4.6	4.7	5.0	4.9	4.2	4.3	4.1
	LBM <sub>SN</sub> (1)	7.1	6.7	7.0	6.4	6.3	6.0	5.8	5.5	5.0	5.1	4.4	4.1
	LBM <sub>SN</sub> (2)	5.6	4.8	4.8	5.2	4.7	4.6	4.7	5.0	4.9	4.2	4.3	4.1
	BP <sub>S</sub>	n.a.	n.a.	n.a.	n.a.	75.9	64.2	58.9	57.6	57.1	55.3	53.5	52.7
	LBM <sub>S</sub>	n.a.	n.a.	n.a.	n.a.	75.9	64.3	59.1	57.8	57.3	55.7	53.8	52.9
	BP <sub>W</sub>	7.6	6.2	5.6	5.0	4.5	4.5	4.0	3.8	3.3	2.7	2.8	3.4
	LBM <sub>W</sub>	7.6	6.2	5.6	5.0	4.6	4.5	4.0	3.8	3.4	2.7	2.8	3.6
	BP <sub>SN</sub>	7.6	5.8	4.7	4.5	4.0	3.5	3.1	3.3	2.5	2.5	2.0	2.1
LBM <sub>SN</sub>	7.6	5.8	4.7	4.5	4.1	3.5	3.2	3.4	2.5	2.5	2.0	2.2	

## 5.7 Appendix: Proofs

The proofs of Theorem 5.3.1 and 5.3.2 are quite technical. These are adaptations of the arguments used in [21].

In the rest of the paper  $K$  and  $\rho$  denote generic constants, such that  $K > 0$  and  $\rho \in (0, 1)$ .

**Lemma 5.7.1** For any  $\alpha \in \Delta_\delta$ , let  $(C_i(\alpha))_{i \in \mathbb{N}}$  be the sequence satisfying

$$\boldsymbol{\epsilon}_n(\alpha) = \sum_{i=0}^{\infty} C_i(\alpha) \mathbf{X}_{n-i}, \quad \frac{\partial \boldsymbol{\epsilon}_n(\alpha)}{\partial \alpha_m} = \sum_{i=0}^{\infty} C_{m,i}(\alpha) \mathbf{X}_{n-i} \quad \text{and} \quad \frac{\partial^2 \boldsymbol{\epsilon}_n(\alpha)}{\partial \alpha_m \partial \alpha_l} = \sum_{i=0}^{\infty} C_{m,l,i}(\alpha) \mathbf{X}_{n-i}. \quad (5.40)$$

Then there exists a constant  $K$  such that for all  $i \in \mathbb{N}$

$$\sup_{\alpha \in \Delta_\delta} \|C_i(\alpha)\| \leq K i^{\text{T}q^*} \left( \frac{1}{1+\delta} \right)^i, \quad \sup_{\alpha \in \Delta_\delta} \|C_{m,i}(\alpha)\| \leq K i^{\text{T}q^*} \left( \frac{1}{1+\delta} \right)^i$$

and  $\sup_{\alpha \in \Delta_\delta} \|C_{m,l,i}(\alpha)\| \leq K i^{\text{T}q^*} \left( \frac{1}{1+\delta} \right)^i$ .

**Proof 5.7.1** Consider the case where  $q^* = 0$  and  $p^* > 0$ . Let  $L = \sup_{i=0, \dots, p^*} \sup_{\alpha \in \Delta_\delta} \|C_i(\alpha)\|$  and put  $K = L(1+\delta)^{p^*}$ . It is not difficult to show that

$$\sup_{\alpha \in \Delta_\delta} \|C_i(\alpha)\| \leq K \left( \frac{1}{1+\delta} \right)^i.$$

Now consider the case  $q^* > 0$ . Define the  $q^* \text{T} \times 1$  vectors

$$\tilde{\mathbf{X}}_n = \begin{pmatrix} \mathbf{X}_n \\ \mathbf{0}_{\text{T} \times 1} \\ \vdots \\ \mathbf{0}_{\text{T} \times 1} \end{pmatrix}; \quad \underline{\boldsymbol{\epsilon}}_n(\alpha) = \begin{pmatrix} \boldsymbol{\epsilon}_n(\alpha) \\ \boldsymbol{\epsilon}_{n-1}(\alpha) \\ \vdots \\ \boldsymbol{\epsilon}_{n-q^*+1}(\alpha) \end{pmatrix}; \quad \mathbf{e}_n(\alpha) = \begin{pmatrix} \mathbf{e}_n(\alpha) \\ \mathbf{e}_{n-1}(\alpha) \\ \vdots \\ \mathbf{e}_{n-q^*+1}(\alpha) \end{pmatrix}, \quad (5.41)$$

and the  $q^* \text{T} \times q^* \text{T}$  companion matrices

$$\mathbf{H}_i = \begin{pmatrix} \mathbf{B}_0^{-1} \mathbf{A}_i & \mathbf{0}_{\text{T} \times \text{T}} & \cdots & \mathbf{0}_{\text{T} \times \text{T}} \\ \mathbf{0}_{\text{T} \times \text{T}} & \mathbf{0}_{\text{T} \times \text{T}} & \cdots & \mathbf{0}_{\text{T} \times \text{T}} \\ \vdots & & \ddots & \mathbf{0}_{\text{T} \times \text{T}} \\ \mathbf{0}_{\text{T} \times \text{T}} & \mathbf{0}_{\text{T} \times \text{T}} & \cdots & \mathbf{0}_{\text{T} \times \text{T}} \end{pmatrix}, \quad i = 1, \dots, p^*,$$

$$\mathbf{D} = \begin{pmatrix} -\mathbf{B}_0^{-1} \mathbf{B}_1 & -\mathbf{B}_0^{-1} \mathbf{B}_2 & \cdots & -\mathbf{B}_0^{-1} \mathbf{B}_{q^*-1} & -\mathbf{B}_0^{-1} \mathbf{B}_{q^*} \\ \mathbf{I}_{\text{T} \times \text{T}} & \mathbf{0}_{\text{T} \times \text{T}} & \cdots & \mathbf{0}_{\text{T} \times \text{T}} & \mathbf{0}_{\text{T} \times \text{T}} \\ \mathbf{0}_{\text{T} \times \text{T}} & \mathbf{I}_{\text{T} \times \text{T}} & \ddots & \vdots & \mathbf{0}_{\text{T} \times \text{T}} \\ \vdots & \ddots & \ddots & \mathbf{0}_{\text{T} \times \text{T}} & \mathbf{0}_{\text{T} \times \text{T}} \\ \mathbf{0}_{\text{T} \times \text{T}} & \cdot & \mathbf{0}_{\text{T} \times \text{T}} & \mathbf{I}_{\text{T} \times \text{T}} & \mathbf{0}_{\text{T} \times \text{T}} \end{pmatrix} \quad (5.42)$$

We have

$$\mathbf{H}_0 \tilde{\mathbf{X}}_n + \sum_{i=1}^{p^*} \mathbf{H}_i \tilde{\mathbf{X}}_{n-1} = \underline{\boldsymbol{\epsilon}}_n(\alpha) - \mathbf{D} \underline{\boldsymbol{\epsilon}}_{n-1}(\alpha).$$

This implies that

$$\underline{\boldsymbol{\epsilon}}_n(\alpha) = \sum_{j=0}^{\infty} \mathbf{D}^j (\mathbf{H}_0 \tilde{\mathbf{X}}_{n-j} + \mathbf{H}_1 \tilde{\mathbf{X}}_{n-j-1} + \dots + \mathbf{H}_{p^*} \tilde{\mathbf{X}}_{n-j-p^*}) = \sum_{j=0}^{\infty} \tilde{\mathbf{C}}_j(\alpha) \tilde{\mathbf{X}}_{n-j};$$

where

$$\tilde{\mathbf{C}}_i(\alpha) = \sum_{j=0}^{\min(i, p^*)} \mathbf{D}^{i-j} \mathbf{H}_j. \quad (5.43)$$

We have the characteristic polynomial of  $D$ :

$$P_D = (-1)^{q^*} Z^{q^*} B_0^{-1} \det\left(B\left(\frac{1}{Z}\right)\right)$$

for which the nonzero eigenvalues  $\lambda_i$  are equal to the inverse of the zeros of  $\det\{B(z)\}$ .

Now using a multiplicative norm it can be shown that there exist a constant  $K_1$  independent of  $\alpha$  such that

$$\|\tilde{C}_i(\alpha)\| \leq \sum_{j=0}^{p^*} \|D^{i-j}\| \|H_j\| \leq K_1 i^{Tq^*} \left(\frac{1}{1+\delta}\right)^i.$$

The conclusion follow from the fact that  $C_i(\alpha)$  is the matrix in the first row and the first column of the block matrix  $\tilde{C}_i(\alpha)$ .

Consider the Jordan decomposition  $D = P\Lambda P^{-1}$ , where the matrix  $\Lambda$  takes the form

$$\Lambda = \begin{pmatrix} \Lambda_1 & 0 & \cdots & 0 \\ 0 & \Lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & \Lambda_s \end{pmatrix} \quad \text{with} \quad \Lambda_h = \begin{pmatrix} \lambda_h & 1 & 0 & \cdots & 0 \\ 0 & \lambda_h & 1 & & 0 \\ \vdots & \ddots & \ddots & \ddots & \\ 0 & & & \lambda_h & 1 \\ 0 & 0 & \cdots & 0 & \lambda_h \end{pmatrix}.$$

Using this decomposition it can be shown that  $D^t = P\Lambda^t P^{-1}$  where  $\Lambda^t = \text{Diag}(\Lambda_1^t, \dots, \Lambda_s^t)$  with

$$\Lambda_h^t = \begin{pmatrix} \lambda_h^t & \binom{t}{1} \lambda_h^{t-1} & \cdots & \binom{t}{r_h-1} \lambda_h^{t-r_h+1} \\ 0 & \lambda_h^t & \cdots & \binom{t}{r_h-2} \lambda_h^{t-r_h+2} \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & \lambda_h^t \end{pmatrix}.$$

Therefore there exists a positive constant  $K$  independant of  $\alpha$  such that

$$\|D^t\| = \|P\Lambda^t P^{-1}\| \leq \|P\| \|P^{-1}\| \|\Lambda^t\| \leq K t^{Tq^*} \left(\frac{1}{1+\delta}\right)^t. \quad (5.44)$$

Now consider the second inequality in (5.40). We have

$$\frac{\partial \tilde{C}_i(\alpha)}{\alpha_m} = \sum_{j=0}^{\min(i, p^*)} \left\{ \frac{\partial D^{i-j}}{\partial \alpha_m} H_j + D^{i-j} \frac{\partial H_j}{\partial \alpha_m} \right\}$$

and

$$\sup_{\alpha \in \Delta_\delta} \left\| \frac{\partial \tilde{C}_i(\alpha)}{\alpha_m} \right\| \leq \sum_{j=0}^{\min(i, p^*)} \left\{ \left\| \frac{\partial D^{i-j}}{\partial \alpha_m} \right\| \|H_j\| + \|D^{i-j}\| \left\| \frac{\partial H_j}{\partial \alpha_m} \right\| \right\}. \quad (5.45)$$

From  $D^t = P\Lambda^t P^{-1}$ , we have

$$\frac{\partial D^t}{\partial \alpha_m} = \frac{\partial P}{\partial \alpha_m} \Lambda^t P^{-1} + P \frac{\partial \Lambda^t}{\partial \alpha_m} P^{-1} + P \Lambda^t \frac{\partial P^{-1}}{\partial \alpha_m},$$

with  $\frac{\partial \Lambda^t}{\partial \alpha_m} = \text{Diag}\left(\frac{\partial \Lambda_1^t}{\partial \alpha_m}, \dots, \frac{\partial \Lambda_s^t}{\partial \alpha_m}\right)$  where

$$\frac{\partial \Lambda_h^t}{\partial \alpha_m} = \frac{\partial \lambda_h}{\partial \alpha_m} \begin{pmatrix} t\lambda_h^{t-1} & \binom{t}{1} (t-1)\lambda_h^{t-2} & \cdots & \binom{t}{r_h-1} (t-r_h+1)\lambda_h^{t-r_h} \\ 0 & n\lambda_h^{t-1} & \cdots & \binom{t}{r_h-2} (t-r_h+2)\lambda_h^{t-r_h+1} \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & t\lambda_h^{t-1} \end{pmatrix}.$$



This implies that

$$\left\| \frac{\partial D^t}{\partial \alpha_m} \right\| \leq \left\| \frac{\partial P}{\partial \alpha_m} \right\| \|\Lambda^t\| \|P^{-1}\| + \|P\| \left\| \frac{\partial \Lambda^t}{\partial \alpha_m} \right\| \|P^{-1}\| + \|P\| \|\Lambda^t\| \left\| \frac{\partial P^{-1}}{\partial \alpha_m} \right\|.$$

Since  $\max_h |\lambda_h| \leq 1/(1+\delta)$  there exists a positive constant  $K_2$  independent of  $\alpha$  such that

$$\left| \frac{\partial D^t}{\partial \alpha_m} \right| \leq K_2 t^{\Gamma q^*} \left( \frac{1}{1+\delta} \right)^t. \quad (5.46)$$

Similar inequality holds for the last equality of (5.40) by writing

$$\frac{\partial^2 \Lambda^t}{\partial \alpha_l \partial \alpha_m} = \text{Diag} \left( \frac{\partial^2 \Lambda_1^t}{\partial \alpha_l \partial \alpha_m}, \dots, \frac{\partial^2 \Lambda_s^t}{\partial \alpha_l \partial \alpha_m} \right)$$

where

$$\frac{\partial^2 \Lambda_h^t}{\partial \alpha_l \partial \alpha_m} = \frac{\partial}{\partial \alpha_l} \left( \frac{\partial \lambda_h}{\partial \alpha_m} \right) \begin{pmatrix} t(t-1)\lambda_h^{t-2} & \binom{t}{1}(t-1)(t-2)\lambda_h^{t-3} & \dots & \binom{t}{r_h-1}(t-r_h+1)(t-r_h)\lambda_h^{t-r_h-1} \\ 0 & t(t-1)\lambda_h^{t-2} & \dots & \binom{t}{r_h-2}(t-r_h+2)(t-r_h+1)\lambda_h^{t-r_h} \\ \vdots & & \ddots & \\ 0 & 0 & \dots & t(t-1)\lambda_h^{t-2} \end{pmatrix}$$

The proof is complete.

**Lemma 5.7.2** For any  $\alpha \in \Delta_\delta$  and  $m, l \in \{1, \dots, k_0 T\}$ , let

$$\mathbf{e}_n(\alpha) = \sum_{i=0}^{N-2} C_i(\alpha) \mathbf{X}_{n-i}, \quad \frac{\partial \mathbf{e}_n(\alpha)}{\partial \alpha_m} = \sum_{i=0}^{N-2} C_{m,i}(\alpha) \mathbf{X}_{n-i} \quad \text{and} \quad \frac{\partial^2 \mathbf{e}_n(\alpha)}{\partial \alpha_m \partial \alpha_l} = \sum_{i=0}^{N-2} C_{m,l,i}(\alpha) \mathbf{X}_{n-i}.$$

Then there exists a constant  $K$  such for any  $i \in \mathbb{N}$

$$\sup_{\alpha \in \Delta_\delta} \|C_i(\alpha)\| \leq K i^{\Gamma q^*} \left( \frac{1}{1+\delta} \right)^i, \quad \sup_{\alpha \in \Delta_\delta} \|C_{m,i}(\alpha)\| \leq K i^{\Gamma q^*} \left( \frac{1}{1+\delta} \right)^i \quad \text{and} \quad \sup_{\alpha \in \Delta_\delta} \|C_{m,l,i}(\alpha)\| \leq K i^{\Gamma q^*} \left( \frac{1}{1+\delta} \right)^i.$$

**Proof 5.7.2** The proof is similar to that given in Lemma 5.7.1.

**Lemma 5.7.3** We have

$$\mathbb{E} \left[ \sup_{\alpha \in \Delta_\delta} \epsilon_{nT+v}^2(\alpha) \right] < \infty \quad (5.47)$$

**Proof 5.7.3** From Lemma 5.7.1 there exists a constant  $K$  such that

$$\sup_{\alpha \in \Omega_\delta} |\epsilon_{nT+v}(\alpha)| \leq K \sum_{i=0}^{\infty} i^{\Gamma q^*} \left( \frac{1}{1+\delta} \right)^i |X_{nT+v-i}|.$$

Using the Cauchy criterion, it can be shown that the series  $\sum_{i=0}^{\infty} i^{\Gamma q^*} (1/(1+\delta))^i |X_{nT+v-i}|$  converges in mean squares. The result follows.

**Lemma 5.7.4** Let  $\epsilon_n(\alpha)$  as defined by (5.6). For any  $\alpha \in \Delta$

$$\epsilon_n(\alpha) = \epsilon_n(\alpha_0) \Rightarrow \alpha = \alpha_0.$$

**Proof 5.7.4** Since the covariance matrix  $\Sigma_\epsilon = \text{Diag}(\sigma_1^2, \dots, \sigma_T^2)$  is supposed to be strictly positive definite for any sequence of  $T \times T$  real matrices  $(\Psi_i)_{i \in \mathbb{N}}$ , we have that

$$\sum_{i=0}^{\infty} \Psi_i \mathbf{X}_{n-i} = 0 \text{ a.s.} \Rightarrow \Psi_i = 0, i \geq 0.$$

Let  $\alpha \in \Delta$ , in view of Lemma 5.7.1 we have

$$\boldsymbol{\epsilon}_n(\alpha) = \sum_{i=0}^{\infty} C_i(\alpha) \mathbf{X}_{n-i} \text{ and } \boldsymbol{\epsilon}_n(\alpha_0) = \sum_{i=0}^{\infty} C_i(\alpha_0) \mathbf{X}_{n-i}.$$

Then we obtain

$$\boldsymbol{\epsilon}_n(\alpha) - \boldsymbol{\epsilon}_n(\alpha_0) = 0 \text{ a.s.} \Rightarrow \sum_{i=0}^{\infty} (C_i(\alpha) - C_i(\alpha_0)) \mathbf{X}_{n-i} = 0 \text{ a.s.} \Rightarrow C_i(\alpha) = C_i(\alpha_0), i \geq 0.$$

The conclusion hold by invoking the identifiability assumption.

**Lemma 5.7.5** For any  $\alpha \in \Delta$  and any  $\sigma^2 = (\sigma_1^2, \dots, \sigma_T^2) > 0$ , let

$$O_\infty^{\sigma^2}(\alpha) = \sum_{v=1}^T \sigma_v^{-2} \mathbb{E}[\epsilon_v^2(\alpha)].$$

Then for any  $\alpha \neq \alpha_0$  we have

$$O_\infty^{\sigma^2}(\alpha_0) = \sum_{v=1}^T \frac{\sigma_{0v}^2}{\sigma_v^2} < O_\infty^{\sigma^2}(\alpha).$$

**Proof 5.7.5** We assume  $n = 0$  in  $t = nT + v$ , from where  $\epsilon_v(\alpha) - \epsilon_v(\alpha_0)$  belongs to Hilbert's space  $\mathcal{H}_X(v-1)$ . Therefore, the linear innovation  $\epsilon_v(\alpha_0)$  is not correlated with  $\epsilon_v(\alpha) - \epsilon_v(\alpha_0)$ .

$$\begin{aligned} \mathbb{E}[\epsilon_v^2(\alpha)] &= \mathbb{E}[\epsilon_v(\alpha) - \epsilon_v(\alpha_0) + \epsilon_v^2(\alpha_0)] \\ &= \mathbb{E}[\epsilon_v^2(\alpha_0)] + \mathbb{E}[(\epsilon_v(\alpha) - \epsilon_v(\alpha_0))^2] + 2\text{Cov}(\epsilon_v(\alpha_0), \epsilon_v(\alpha) - \epsilon_v(\alpha_0)) \\ &= \sigma_{0v}^2 + \mathbb{E}[(\epsilon_v(\alpha) - \epsilon_v(\alpha_0))^2] \end{aligned}$$

In view of Lemma 5.7.1 the second term on the right is strictly positive for at least one  $v \in \{1, \dots, T\}$ , if  $\alpha \neq \alpha_0$ . Thus

$$O_\infty^{\sigma^2}(\alpha) = \sum_{v=1}^T \frac{1}{\sigma_v^2} \mathbb{E}[\epsilon_v(\alpha)^2] > \sum_{v=1}^T \frac{\sigma_{0v}^2}{\sigma_v^2} \text{ if } \alpha \neq \alpha_0.$$

The proof is complete.

**Lemma 5.7.6** For any  $\alpha^* \in \Delta, \alpha^* \neq \alpha_0$ , and any  $\sigma^2 > 0$ , there exists a neighbourhood  $V(\alpha^*)$  of  $\alpha^*$  such that  $V(\alpha^*) \subset \Delta$  and

$$\liminf_{N \rightarrow \infty} \inf_{\alpha \in V(\alpha^*)} O_N^{\sigma^2}(\alpha) > \lim_{N \rightarrow \infty} O_N^{\sigma^2}(\alpha_0) = \sum_{v=1}^T \frac{\sigma_{0v}^2}{\sigma_v^2} \text{ a.s.}$$

**Proof 5.7.6** Let  $V_m(\alpha^*)$  be the open sphere with center  $\alpha^*$  and radius  $\frac{1}{m}$ . Let

$$S_m(n) = \inf_{\alpha \in V_m(\alpha) \cap \Delta} \sum_{v=1}^T \sigma_v^{-2} \epsilon_{nT+v}^2(\alpha).$$

The variable  $S_m(n)$  is measurable because it can be written as the inf over a dense countable subset. Moreover,  $S_m(n)$  belongs to  $L^1$ . The ergodic theorem applied to the stationary process  $\{S_m(n) : n \in \mathbb{Z}\}$  shows that almost surely

$$\inf_{\alpha \in V_m(\alpha^*) \cap \Delta_{v=1}} O_N^{\sigma^2}(\alpha) = \inf_{\alpha \in V_m(\alpha^*) \cap \Delta} \left\{ \frac{1}{N} \sum_{v=1}^T \sigma_v^2 \sum_{n=0}^{N-1} \epsilon_{nT+v}^2(\alpha) \right\} \geq \frac{1}{N} \sum_{n=0}^{N-1} S_m(n) \xrightarrow{N \rightarrow \infty} \mathbb{E}[S_m(0)].$$

Since  $S_m(0)$  increases to  $\sum_{v=1}^T \sigma_v^2 \epsilon_v^2(\alpha^*)$  as  $m$  tends to infinity, by Lemme 5.7.5 and the monotone convergence theorem, we obtain that

$$\lim_{m \rightarrow \infty} \mathbb{E}[S_m(0)] = \sum_{v=1}^T \sigma_v^{-2} \mathbb{E}[\epsilon_v^2(\alpha^*)] = O_\infty^{\sigma^2}(\alpha^*) > \sum_{v=1}^T \frac{\sigma_{0v}^2}{\sigma_v^2}.$$

This implies that

$$\liminf_{m \rightarrow \infty} \liminf_{N \rightarrow \infty} \inf_{\alpha \in V_m(\alpha^*)} O_N^{\sigma^2}(\alpha) \geq O_\infty^{\sigma^2}(\alpha^*) > \sum_{v=1}^T \frac{\sigma_{0v}^2}{\sigma_v^2},$$

and the result follows.

**Lemma 5.7.7** Let  $\sigma^2 = (\sigma_1^2, \dots, \sigma_T^2)'$  be a vector of strictly positive constants, and let  $\hat{\sigma}^2$  be a sequence of random vectors such that  $\hat{\sigma}^2 \rightarrow \sigma^2$  almost surely as  $N \rightarrow \infty$ . For any  $\alpha^* \in \Delta$ ,  $\alpha^* \neq \alpha_0$ , there exists a neighbourhood  $V(\alpha^*)$  of  $\alpha^*$  such that  $V(\alpha^*) \subset \Delta$  and

$$\liminf_{N \rightarrow \infty} \inf_{\alpha \in V_m(\alpha^*)} O_N^{\sigma^2}(\alpha) > \lim_{N \rightarrow \infty} O_N^{\sigma^2}(\alpha_0) = \sum_{v=1}^T \frac{\sigma_{0v}^2}{\sigma_v^2} \text{ a.s.}$$

**Proof 5.7.7** For any  $\epsilon > 0$ , we have almost surely

$$\max_{v=1, \dots, T} \left| \frac{1}{\hat{\sigma}_v^2} - \frac{1}{\sigma_v^2} \right| < \epsilon \quad (5.48)$$

for  $N$  large enough. By the ergodic theorem and Lemma 5.7.3, we thus have

$$\begin{aligned} \limsup_{N \rightarrow \infty} \sup_{\alpha \in \Delta} |O_N^{\hat{\sigma}^2}(\alpha) - O_N^{\sigma^2}(\alpha)| &\leq \limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \sum_{v=1}^T \left| \frac{1}{\hat{\sigma}_v^2} - \frac{1}{\sigma_v^2} \right| \sup_{\alpha \in \Delta} |\epsilon_{nT+v}^2(\alpha)| \\ &\leq \epsilon \sum_{v=1}^T \mathbb{E} \sup_{\alpha \in \Delta} |\epsilon_v^2(\alpha)|. \end{aligned}$$

Since the inequality holds for any  $\epsilon > 0$ , we have

$$\lim_{N \rightarrow \infty} \sup_{\alpha \in \Delta} |O_N^{\hat{\sigma}^2}(\alpha) - O_N^{\sigma^2}(\alpha)| = 0 \text{ a.s.}$$

In view of Lemma 5.7.6, the conclusion follows.

**Lemma 5.7.8** Let  $\epsilon_{nT+v}(\alpha)$  given by 5.8 and  $e_{nT+v}(\alpha)$  given by 5.9. Almost surely, there exist  $K > 0$  and  $\rho \in (0, 1)$  such that

$$\begin{aligned} \sup_{\alpha \in \Delta_\delta} |\epsilon_{nT+v}(\alpha) - e_{nT+v}(\alpha)| &\leq K\rho^n, \\ \sup_{\alpha \in \Delta_\delta} \left| \frac{\partial \epsilon_{nT+v}(\alpha)}{\partial \alpha_m} - \frac{\partial e_{nT+v}}{\partial \alpha_m} \right| &\leq K\rho^n, \\ \sup_{\alpha \in \Delta_\delta} \left| \frac{\partial^2 \epsilon_{nT+v}(\alpha)}{\partial \alpha_m \partial \alpha_l} - \frac{\partial^2 e_{nT+v}}{\partial \alpha_m \partial \alpha_l} \right| &\leq K\rho^n \end{aligned}$$

for all  $v \in \{1, 2, \dots, T\}$  and all  $n \geq 1$ .

**Proof 5.7.8** For  $q^* = 0$ , the result is obvious. Otherwise, consider and the  $q^*T \times 1$  vectors  $\underline{\mathbf{e}}_n(\alpha)$ ,  $\mathbf{e}_n(\alpha)$  define by (5.41) and the  $q^*T \times q^*T$  companion matrix  $D$  given by (5.42). We have

$$\underline{\mathbf{e}}_n(\alpha) - \mathbf{e}_n(\alpha) = D^{n-p^*} (\underline{\mathbf{e}}_p(\alpha) - \mathbf{e}_p(\alpha)) \quad \forall n > p^*. \quad (5.49)$$

The conclusion follows from Lemma 5.7.1.

**Lemma 5.7.9** For any constant  $\underline{\sigma} > 0$ , we have almost surely

- (i)  $\sup_{\sigma^2 > \underline{\sigma}} \sup_{\alpha \in \Delta} |O_N^{\sigma^2}(\alpha) - Q_N^{\sigma^2}(\alpha)| = O(1/N)$  a.s.  $N \rightarrow \infty$
- (ii)  $\sup_{\sigma^2 > \underline{\sigma}} \sup_{\alpha \in \Delta} \left| \frac{\partial O_N^{\sigma^2}(\alpha)}{\partial \alpha} - \frac{\partial Q_N^{\sigma^2}(\alpha)}{\partial \alpha} \right| = O(1/N)$  a.s.  $N \rightarrow \infty$
- (iii)  $\sup_{\sigma^2 > \underline{\sigma}} \sup_{\alpha \in \Delta} \left| \frac{\partial^2 O_N^{\sigma^2}(\alpha)}{\partial \alpha \partial \alpha'} - \frac{\partial^2 Q_N^{\sigma^2}(\alpha)}{\partial \alpha \partial \alpha'} \right| = O(1)$  a.s.  $N \rightarrow \infty$ .

**Proof 5.7.9** Using Lemma 5.7.8, we obtain that

$$\begin{aligned} (i) \quad & \sup_{\alpha \in \Delta_\delta, \sigma^2 > \underline{\sigma}} |O_N^{\sigma^2}(\alpha) - Q_N^{\sigma^2}(\alpha)| \\ &= \sup_{\alpha \in \Delta_\delta, \sigma^2 > \underline{\sigma}} \left| \frac{1}{N} \sum_{v=1}^T \sigma_v^{-2} \sum_{n=0}^{N-1} (e_{nT+v}^2(\alpha) - \epsilon_{nT+v}^2(\alpha)) \right| \\ &\leq \sum_{v=1}^T \underline{\sigma}^{-2} \sup_{\alpha \in \Delta_\delta} \left| \frac{1}{N} \sum_{n=0}^{N-1} (e_{nT+v}^2(\alpha) - \epsilon_{nT+v}^2(\alpha)) \right| \\ &= \sum_{v=1}^T \underline{\sigma}^{-2} \frac{1}{N} \sup_{\alpha \in \Delta_\delta} \left| \sum_{n=0}^{N-1} (e_{nT+v}(\alpha) - \epsilon_{nT+v}(\alpha))^2 + 2\epsilon_{nT+v}(\alpha)(e_{nT+v}(\alpha) - \epsilon_{nT+v}(\alpha)) \right| \\ &\leq \frac{K}{N} \sum_{v=1}^T \sum_{n=0}^{N-1} \rho^n \left( 1 + \sup_{\alpha \in \Delta_\delta} |e_{nT+v}(\alpha)| \right). \end{aligned}$$

By Lemma 5.7.3 and the Markov inequality, we have

$$\sum_{n=0}^{\infty} \mathbb{P} \left( \rho^n \left( 1 + \sup_{\alpha \in \Delta_\delta} |e_{nT+v}(\alpha)| \right) > \epsilon \right) \leq \mathbb{E} \left( 1 + \sup_{\alpha \in \Delta_\delta} |e_{nT+v}(\alpha)| \right) \sum_{n=0}^{\infty} \frac{\rho^n}{\epsilon} < \infty.$$

By Borel-Cantelli theorem,  $\rho^n \left( 1 + \sup_{\alpha \in \Delta_\delta} |e_{nT+v}(\alpha)| \right) \rightarrow 0$  almost surely as  $n \rightarrow \infty$ . The conclusion follows from the Cesaro theorem.

We also have

$$\begin{aligned} (ii) \quad & \sup_{\alpha \in \Delta_\delta, \sigma^2 > \underline{\sigma}} \left| \frac{\partial O_N^{\sigma^2}(\alpha)}{\partial \alpha} - \frac{\partial Q_N^{\sigma^2}(\alpha)}{\partial \alpha} \right| \\ &= \sup_{\alpha \in \Delta_\delta, \sigma^2 > \underline{\sigma}} \left| \sum_{v=1}^T \underline{\sigma}^{-2} \frac{2}{N} \sum_{n=0}^{N-1} (\epsilon_{nT+v} \frac{\partial \epsilon_{nT+v}}{\partial \alpha_m} - e_{nT+v} \frac{\partial e_{nT+v}}{\partial \alpha_m}) \right| \\ &\leq \sup_{\alpha \in \Delta_\delta} \left| \sum_{v=1}^T \underline{\sigma}^{-2} \frac{2}{N} \sum_{n=0}^{N-1} \left\{ \epsilon_{nT+v} \frac{\partial \epsilon_{nT+v}}{\partial \alpha_m} - e_{nT+v} \frac{\partial e_{nT+v}}{\partial \alpha_m} - \frac{\partial \epsilon_{nT+v}}{\partial \alpha_m} e_{nT+v} + \frac{\partial \epsilon_{nT+v}}{\partial \alpha_m} e_{nT+v} \right\} \right| \\ &\leq \sum_{v=1}^T \underline{\sigma}^{-2} \frac{2}{N} \sum_{n=0}^{N-1} \sup_{\alpha \in \Delta_\delta} \left| (\epsilon_{nT+v} - e_{nT+v}) \frac{\partial \epsilon_{nT+v}}{\partial \alpha_m} + \left( \frac{\partial \epsilon_{nT+v}}{\partial \alpha_m} - \frac{\partial e_{nT+v}}{\partial \alpha_m} \right) e_{nT+v} \right| \\ &\leq \frac{K}{N} \sum_{v=1}^T \sum_{n=0}^{N-1} \rho^n \left( \sup_{\alpha \in \Delta_\delta} \left| \frac{\partial \epsilon_{nT+v}}{\partial \alpha_m} \right| + \sup_{\alpha \in \Delta_\delta} |e_{nT+v}| \right). \end{aligned}$$

By Lemmas 5.7.1, 5.7.3 and the Markov inequality, we have

$$\sum_{n=0}^{\infty} \mathbb{P} \left( \rho^n \sup_{\alpha \in \Delta_\delta} \left| \frac{\partial \epsilon_{nT+v}}{\partial \alpha_m} \right| > \epsilon \right) \leq \mathbb{E} \sup_{\alpha \in \Delta_\delta} |\epsilon_{nT+v}(\alpha)| \sum_{n=0}^{\infty} \frac{\rho^n}{\epsilon} < \infty.$$

By Borel-Cantelli theorem,  $\rho^n \sup_{\alpha \in \Delta_\delta} |\partial \epsilon_{nT+v} / \partial \alpha_m| \rightarrow 0$  almost surely as  $n \rightarrow \infty$ . By the same arguments  $\rho^n \sup_{\alpha \in \Delta_\delta} |e_{nT+v}(\alpha)| \rightarrow 0$  almost surely as  $n \rightarrow \infty$ .

The conclusion follows from the Cesaro theorem. The proof of the point (iii) is similar to that of point (ii).

**Lemma 5.7.10** For any  $\alpha^* \in \Delta$ ,  $\alpha^* \neq \alpha_0$ , and any  $\sigma^2 > 0$ , there exists a neighborhood  $V(\alpha^*)$  of  $\alpha^*$  such that  $V(\alpha^*) \subset \Delta$  and

$$\liminf_{N \rightarrow \infty} \inf_{\alpha \in V(\alpha^*)} Q_N^{\sigma^2}(\alpha) > \sum_{v=1}^T \frac{\sigma_{0v}^2}{\sigma_v^2} \text{ a.s.}$$

For any neighborhood  $V(\alpha_0)$  of  $\alpha_0$ ,

$$\limsup_{N \rightarrow \infty} \inf_{\alpha \in V(\alpha_0)} Q_N^{\sigma^2}(\alpha) \leq \sum_{v=1}^T \frac{\sigma_{0v}^2}{\sigma_v^2} \text{ a.s.}$$

**Proof 5.7.10** Since  $V(\alpha^*)$  can be chosen as being included in  $\Delta_\delta$  for some  $\delta > 0$ . We have

$$\inf_{\alpha \in V(\alpha^*)} Q_N^{\sigma^2}(\alpha) \geq \inf_{\alpha \in V(\alpha^*)} O_N^{\sigma^2}(\alpha) - \sup_{\alpha \in \Delta_\delta} |O_N^{\sigma^2}(\alpha) - Q_N^{\sigma^2}(\alpha)|$$

and

$$\inf_{\alpha \in V(\alpha_0)} Q_N^{\sigma^2}(\alpha) \leq O_N^{\sigma^2}(\alpha_0) + \sup_{\alpha \in \Delta_\delta} |O_N^{\sigma^2}(\alpha) - Q_N^{\sigma^2}(\alpha)|$$

The conclusion follows from Lemmas 5.7.6 and 5.7.9.

**Lemma 5.7.11** Let  $\sigma^2 = (\sigma_1^2, \dots, \sigma_T^2)'$  be a vector of strictly positive constants, and let  $\hat{\sigma}^2$  be a sequence of random vectors such that  $\hat{\sigma}^2 \rightarrow \sigma^2$  almost surely as  $N \rightarrow \infty$ . For any  $\alpha^* \in \Delta$ ,  $\alpha^* \neq \alpha_0$ , there exists a neighborhood  $V(\alpha^*)$  of  $\alpha^*$  such that  $V(\alpha^*) \subset \Delta$  and

$$\liminf_{N \rightarrow \infty} \inf_{\alpha \in V(\alpha^*)} Q_N^{\hat{\sigma}^2}(\alpha) > \sum_{v=1}^T \frac{\sigma_{0v}^2}{\sigma_v^2} \text{ a.s.}$$

For any neighborhood  $V(\alpha_0)$  of  $\alpha_0$ ,

$$\limsup_{N \rightarrow \infty} \inf_{\alpha \in V(\alpha_0)} Q_N^{\hat{\sigma}^2}(\alpha) \leq \sum_{v=1}^T \frac{\sigma_{0v}^2}{\sigma_v^2} \text{ a.s.}$$

**Proof 5.7.11** The proof is the same as that of Lemma 5.7.10, referring to lemma 5.7.7 instead of Lemma 5.7.6.

### 5.7.1 Proof of Theorem 5.3.1

Let  $V(\alpha_0)$  be a neighborhood of  $\alpha_0$ . Clearly  $\Delta_\delta$  is covered by  $V(\alpha_0)$  and the union of all the  $V(\alpha^*)$ ,  $\alpha^* \in \Delta_\delta - V(\alpha_0)$  where  $V(\alpha^*)$  is defined in Lemma 5.7.10. By the compactness of  $\Delta_\delta$ , there exists  $\alpha_1, \dots, \alpha_k$  such that  $\Delta_\delta$  is covered by  $V(\alpha_0), V(\alpha_1), \dots, V(\alpha_k)$ . Lemma 5.7.10 shows that almost surely,

$$\inf_{\alpha \in \Delta_\delta} Q_N^{\sigma^2}(\alpha) = \min_{i=0,1,\dots,k} \inf_{\alpha \in V(\alpha_i) \cap \Delta_\delta} Q_N^{\sigma^2}(\alpha) = \inf_{\alpha \in V(\alpha_i) \cap \Delta_\delta} Q_N^{\sigma^2}(\alpha)$$

for  $N$  large enough. By using the previous lemma, we have  $\hat{\alpha}_{\text{GLS}}^{\hat{\sigma}^2} \rightarrow \alpha_0$  almost surely whenever  $\hat{\sigma}^2 \rightarrow \sigma_0^2$  as  $N \rightarrow \infty$ . It follows that the GLS estimator is consistent.

□

### 5.7.2 Proof of Theorem 5.3.2

Theorem 5.3.2 can be established using the following lemmas.

**Lemma 5.7.12** For any  $\alpha^* \in \Delta$ , and any  $m \in \{1, 2, \dots, (p + q + P + Q)T\}$ , there exists absolutely summable sequences  $(C_i(\alpha))_{i \in \mathbb{N}}$  and  $(C_{m,i}(\alpha))_{i \in \mathbb{N}}$  such that

$$\mathbf{e}_n(\alpha) = \sum_{i=0}^{\infty} C_i(\alpha) \mathbf{X}_{n-i}, \quad \frac{\partial \mathbf{e}_n(\alpha)}{\partial \alpha_m} = \sum_{i=0}^{\infty} C_{m,i}(\alpha) \mathbf{X}_{n-i} \quad \text{and} \quad \frac{\partial^2 \mathbf{e}_n(\alpha)}{\partial \alpha_m \partial \alpha_l} = \sum_{i=0}^{\infty} C_{m,l,i}(\alpha) \mathbf{X}_{n-i} \quad (5.50)$$

Moreover, there exists  $\rho \in (0, 1)$  and  $K \in [0, \infty[$  such that, for all  $i \geq 0$ ,

$$\sup_{\alpha \in \Delta_\delta} \|C_i(\alpha)\| \leq K\rho^i, \quad \sup_{\alpha \in \Delta_\delta} \|C_{m,i}(\alpha)\| \leq K\rho^i \quad \text{and} \quad \sup_{\alpha \in \Delta_\delta} \|C_{m,l,i}(\alpha)\| \leq K\rho^i \quad (5.51)$$

**Proof 5.7.12** For  $q^* = 0$ , the result is obvious. Otherwise, we use arguments similar to those used to prove Lemma 5.7.1. Indeed, as shown in Lemma 5.7.1,  $\underline{\mathbf{e}}_n(\alpha) = \sum_{i=0}^{\infty} \tilde{C}_i(\alpha) \tilde{\mathbf{X}}_{n-i}$  where  $\tilde{\mathbf{X}}_{n-i}$  and  $\underline{\mathbf{e}}_n(\alpha)$  are given by (5.41) and  $\tilde{C}_i(\alpha) = \sum_{j=0}^{\min(i, p^*)} B^{i-j} A_j$  as in (5.43). Then

$$\frac{\partial \underline{\mathbf{e}}_n(\alpha)}{\partial \alpha_m} = \sum_{i=0}^{\infty} \frac{\partial \tilde{C}_i(\alpha)}{\partial \alpha_m} \tilde{\mathbf{X}}_{n-i},$$

and since

$$\mathbf{e}_n(\alpha) = \sum_{i=0}^{\infty} L_1' \tilde{C}_i(\alpha) L_1 \mathbf{X}_{n-i} = \sum_{i=0}^{\infty} C_i(\alpha) \mathbf{X}_{n-i},$$

we have

$$\frac{\partial \mathbf{e}_n(\alpha)}{\partial \alpha_m} = \sum_{i=0}^{\infty} L_1' \frac{\partial \tilde{C}_i(\alpha)}{\partial \alpha_m} L_1 \mathbf{X}_{n-i} = \sum_{i=0}^{\infty} C_{m,i}(\alpha) \mathbf{X}_{n-i},$$

where  $C_i(\alpha) = L_1' \tilde{C}_i(\alpha) L_1$  and  $C_{m,i}(\alpha) = L_1' \frac{\partial \tilde{C}_i(\alpha)}{\partial \alpha_m} L_1$  with  $L_1 = (I_{T \times T}, 0_{T \times T}, \dots, 0_{T \times T})'$ . Also, from Lemma 5.7.1, there exists a positive constant  $K_1$  independant of  $\alpha$  such that

$$\begin{aligned} \sup_{\alpha \in \Delta_\delta} \|C_i(\alpha)\| &\leq K_1 i^{Tq^*} \left( \frac{1}{1+\delta} \right)^i \\ &\leq K_1 i^{Tq^*} \left( \frac{1}{1+(\delta/2)} \right)^i \left( \frac{1+(\delta/2)}{1+\delta} \right)^i \\ &\leq K\rho^i \end{aligned}$$

where  $\rho = \frac{1+(\delta/2)}{1+\delta} \in (0, 1)$  and  $K = K_1 \max_i \left\{ i^{Tq^*} \left( \frac{1}{1+(\delta/2)} \right)^i \right\} \leq K_1 e^{-Tq^*} \left[ \frac{q^* T}{\log(1+(\delta/2))} \right]^{Tq^*}$ .

Now, consider the second inequality in (5.51), from 5.7.1, we obtain that

$$\begin{aligned} \sup_{\alpha \in \Delta_\delta} \|C_{m,i}(\alpha)\| &\leq K_3 i^{q^*} \left( \frac{1}{1+\delta} \right)^i \\ &\leq K_3 i^{Tq^*} \left( \frac{1}{1+(\delta/2)} \right)^i \left( \frac{1+(\delta/2)}{1+\delta} \right)^i \\ &\leq K\rho^i \end{aligned}$$

where  $\rho = \frac{1+(\delta/2)}{1+\delta} \in (0, 1)$  and  $K = K_3 \max_i \left\{ i^{Tq^*} \left( \frac{1}{1+(\delta/2)} \right)^i \right\} \leq K_3 e^{-Tq^*} \left[ \frac{Tq^*}{\log(1+(\delta/2))} \right]^{Tq^*}$ .

Which completes the proof of Lemma 5.7.12.

**Lemma 5.7.13** *Let the assumption of Theorem 5.3.2 be satisfied. For all  $\alpha \in \Delta_\delta$  and all  $\sigma^2 > 0$ , the matrix*

$$\mathbf{I}(\alpha, \sigma^2) = \frac{1}{4} \lim_{N \rightarrow \infty} \text{Var} \left( \sqrt{N} \frac{\partial \mathbf{O}_N^{\sigma^2}(\alpha)}{\partial \alpha} \right)$$

exists.

**Proof 5.7.13** *Let the diagonal matrix  $\Sigma_\sigma = \text{Diag}(\sigma^2)$  and*

$$\mathbf{Y}_n = \frac{\partial \boldsymbol{\epsilon}'_n(\alpha)}{\partial \alpha} \Sigma_\sigma^{-1} \boldsymbol{\epsilon}_n(\alpha) = \sum_{v=1}^T \sigma_v^{-2} \boldsymbol{\epsilon}_{nT+v}(\alpha) \frac{\partial \boldsymbol{\epsilon}_{nT+v}(\alpha)}{\partial \alpha}$$

*The process  $\{\mathbf{Y}_n\}$  is strictly stationary and ergodic. Moreover,*

$$\frac{1}{4} \text{Var} \left( \sqrt{N} \frac{\partial \mathbf{O}_N^{\sigma^2}(\alpha)}{\partial \alpha} \right) = \frac{1}{N} \sum_{n=0}^{N-1} \sum_{n'=0}^{N-1} \text{Cov}(\mathbf{Y}_n, \mathbf{Y}_{n'}).$$

Denote by

$$\mathbf{I}_N(l, m) = \frac{1}{N} \sum_{n=0}^{N-1} \sum_{n'=0}^{N-1} \text{Cov}(\mathbf{Y}_n(l), \mathbf{Y}_{n'}(m)) = \frac{1}{N} \sum_{k=1-N}^{N-1} (N - |k|) c(k)$$

where

$$c(k) = \text{Cov}(\mathbf{Y}_n(l), \mathbf{Y}_{n-k}(m))$$

with

$$\mathbf{Y}_n(l) = \frac{\partial \boldsymbol{\epsilon}'_n(\alpha)}{\partial \alpha_l} \Sigma_\sigma^{-1} \boldsymbol{\epsilon}_n(\alpha) \text{ and } \mathbf{Y}_{n-k}(m) = \frac{\partial \boldsymbol{\epsilon}'_{n-k}(\alpha)}{\partial \alpha_l} \Sigma_\sigma^{-1} \boldsymbol{\epsilon}_{n-k}(\alpha)$$

Suppose that  $k \leq 0$ . From Lemma 5.7.12.

$$\begin{aligned} |c(k)| &= \left| \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{i'=0}^{\infty} \sum_{j'=0}^{\infty} \text{Cov} \left( \mathbf{X}'_{n-j} \mathbf{C}'_{l,j}(\alpha) \Sigma_\sigma^{-1} \mathbf{C}_i(\alpha) \mathbf{X}_{n-i}, \mathbf{X}'_{n-k-j'} \mathbf{C}'_{m,j'}(\alpha) \Sigma_\sigma^{-1} \mathbf{C}_{i'}(\alpha) \mathbf{X}_{n-k-i'} \right) \right| \\ &\leq \sum_{i,j,i',j'=0}^{\infty} \left| \text{Cov} \left( \mathbf{X}'_{n-j} \mathbf{C}'_{l,j}(\alpha) \Sigma_\sigma^{-1} \mathbf{C}_i(\alpha) \mathbf{X}_{n-i}, \mathbf{X}'_{n-k-j'} \mathbf{C}'_{m,j'}(\alpha) \Sigma_\sigma^{-1} \mathbf{C}_{i'}(\alpha) \mathbf{X}_{n-k-i'} \right) \right| \\ &\leq g_1 + g_2 + g_3 + g_4 + g_5 \end{aligned}$$

where

$$\begin{aligned} g_1 &= \sum_{i>k/2} \sum_{j=0}^{\infty} \sum_{i'=0}^{\infty} \sum_{j'=0}^{\infty} \left| \text{Cov} \left( \mathbf{X}'_{n-j} \mathbf{C}'_{l,j}(\alpha) \Sigma_\sigma^{-1} \mathbf{C}_i(\alpha) \mathbf{X}_{n-i}, \mathbf{X}'_{n-k-j'} \mathbf{C}'_{m,j'}(\alpha) \Sigma_\sigma^{-1} \mathbf{C}_{i'}(\alpha) \mathbf{X}_{n-k-i'} \right) \right|, \\ g_2 &= \sum_{i'>k/2} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{j'=0}^{\infty} \left| \text{Cov} \left( \mathbf{X}'_{n-j} \mathbf{C}'_{l,j}(\alpha) \Sigma_\sigma^{-1} \mathbf{C}_i(\alpha) \mathbf{X}_{n-i}, \mathbf{X}'_{n-k-j'} \mathbf{C}'_{m,j'}(\alpha) \Sigma_\sigma^{-1} \mathbf{C}_{i'}(\alpha) \mathbf{X}_{n-k-i'} \right) \right|, \\ g_3 &= \sum_{j>k/2} \sum_{i=0}^{\infty} \sum_{i'=0}^{\infty} \sum_{j'=0}^{\infty} \left| \text{Cov} \left( \mathbf{X}'_{n-j} \mathbf{C}'_{l,j}(\alpha) \Sigma_\sigma^{-1} \mathbf{C}_i(\alpha) \mathbf{X}_{n-i}, \mathbf{X}'_{n-k-j'} \mathbf{C}'_{m,j'}(\alpha) \Sigma_\sigma^{-1} \mathbf{C}_{i'}(\alpha) \mathbf{X}_{n-k-i'} \right) \right|, \\ g_4 &= \sum_{j'>k/2} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{i'=0}^{\infty} \left| \text{Cov} \left( \mathbf{X}'_{n-j} \mathbf{C}'_{l,j}(\alpha) \Sigma_\sigma^{-1} \mathbf{C}_i(\alpha) \mathbf{X}_{n-i}, \mathbf{X}'_{n-k-j'} \mathbf{C}'_{m,j'}(\alpha) \Sigma_\sigma^{-1} \mathbf{C}_{i'}(\alpha) \mathbf{X}_{n-k-i'} \right) \right|, \\ g_5 &= \sum_{i \leq k/2} \sum_{j \leq k/2} \sum_{i' \leq k/2} \sum_{j' \leq k/2} \left| \text{Cov} \left( \mathbf{X}'_{n-j} \mathbf{C}'_{l,j}(\alpha) \Sigma_\sigma^{-1} \mathbf{C}_i(\alpha) \mathbf{X}_{n-i}, \mathbf{X}'_{n-k-j'} \mathbf{C}'_{m,j'}(\alpha) \Sigma_\sigma^{-1} \mathbf{C}_{i'}(\alpha) \mathbf{X}_{n-k-i'} \right) \right|. \end{aligned}$$



By the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}
 g_1 &\leq \sum_{i>k/2} \sum_{j=0}^{\infty} \sum_{i'=0}^{\infty} \sum_{j'=0}^{\infty} \left\{ \mathbb{E} \left[ \left( \mathbf{X}'_{n-j} C'_{l,j}(\alpha) \Sigma_{\sigma}^{-1} C_i(\alpha) \mathbf{X}_{n-i} \right)^2 \right] \mathbb{E} \left[ \left( \mathbf{X}'_{n-k-j'} C'_{m,j'}(\alpha) \Sigma_{\sigma}^{-1} C_{i'}(\alpha) \mathbf{X}_{n-k-i'} \right)^2 \right] \right\}^{1/2} \\
 &\leq \sum_{i>k/2} \sum_{j=0}^{\infty} \sum_{i'=0}^{\infty} \sum_{j'=0}^{\infty} \left\{ \mathbb{E} \left[ \|\mathbf{X}'_{n-j}\|^2 \|C'_{l,j}(\alpha)\|^2 \|\Sigma_{\sigma}^{-1}\|^2 \|C_i(\alpha)\|^2 \|\mathbf{X}_{n-i}\|^2 \right] \right. \\
 &\quad \left. \mathbb{E} \left[ \|\mathbf{X}'_{n-k-j'}\|^2 \|C'_{m,j'}(\alpha)\|^2 \|\Sigma_{\sigma}^{-1}\|^2 \|C_{i'}(\alpha)\|^2 \|\mathbf{X}_{n-k-i'}\|^2 \right] \right\}^{1/2} \\
 &\leq \mathbb{E} [\|\mathbf{X}_n\|^4] \|\Sigma_{\sigma}^{-1}\|^2 \sum_{i>k/2} \|C_i(\alpha)\| \sum_{j=0}^{\infty} \|C_{l,j}(\alpha)\| \sum_{i'=0}^{\infty} \|C_{i'}(\alpha)\| \sum_{j'=0}^{\infty} \|C_{m,j'}(\alpha)\| \\
 &\leq M_1 \sum_{i>k/2} \|C_i(\alpha)\| \leq \Delta_1 \rho^{k/2}
 \end{aligned}$$

for some positive constants  $M_1$  and  $\Delta_1$ . Using the same arguments we obtain that  $g_i$  ( $i = 2, 3, 4$ ) is bounded by  $\Delta_i \rho^{k/2}$ .

In other hand, Davydov inequality (see [19]) implies that, there exists a positive constant  $M_5$  such that

$$\begin{aligned}
 g_5 &\leq \sum_{i \leq k/2} \sum_{j \leq k/2} \sum_{i' \leq k/2} \sum_{j' \leq k/2} M_5 \|\mathbf{X}'_{n-j} C'_{l,j}(\alpha) \Sigma_{\sigma}^{-1} C_i(\alpha) \mathbf{X}_{n-i}\|_{2+\tau} \|\mathbf{X}'_{n-k-j'} C'_{m,j'}(\alpha) \Sigma_{\sigma}^{-1} C_{i'}(\alpha) \mathbf{X}_{n-k-i'}\|_{2+\tau} \\
 &\quad \left[ \underline{\alpha}_{\mathbf{X}}(\min\{k+i'-j, k+j'-j, k+i'-i, k+j'-i\}) \right]^{\frac{\tau}{2+\tau}} \\
 &\leq M [\underline{\alpha}_{\mathbf{X}}(k/2)]^{\frac{\tau}{2+\tau}}
 \end{aligned}$$

for some positive constant  $M$ . Therefore, for  $k \geq 0$ , there exist positive constants  $M$  and  $\Delta$  such that

$$|c(k)| \leq \Delta \rho^{k/2} + M [\underline{\alpha}_{\mathbf{X}}(k/2)]^{\frac{\tau}{2+\tau}}.$$

A similar inequality hold for  $k < 0$ . Therefore

$$\sum_{k=-\infty}^{\infty} |c(k)| < \infty$$

Then, the dominated convergences theorem gives

$$\mathbf{I}_N(l, m) = \frac{1}{N} \sum_{k=1-N}^{N-1} (N - |k|) c(k) \xrightarrow{N \rightarrow \infty} \sum_{k=-\infty}^{\infty} |c(k)|.$$

This implies that

$$\mathbf{I}(\alpha, \sigma^2) = \sum_{v=1}^T \sum_{v'=1}^T \sigma_v^{-2} \sigma_{v'}^{-2} \sum_{k=-\infty}^{\infty} \mathbb{E} \left[ \left( \epsilon_{nT+v}(\alpha) \left( \frac{\partial \epsilon_{nT+v}(\alpha)}{\partial \alpha} \right) \right) \left( \epsilon_{(n-k)T+v}(\alpha) \left( \frac{\partial \epsilon_{(n-k)T+v}(\alpha)}{\partial \alpha} \right) \right) \right]$$

exists. Which completes the proof of Lemma 5.7.13.

**Lemma 5.7.14** Under the assumptions of Theorem 5.3.2, the random vector

$$\sqrt{N} \left( \frac{\partial Q_N^{\sigma^2}(\alpha)}{\partial \alpha} \right)_{\alpha=\alpha_0} \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, 4\mathbf{I}(\alpha_0, \sigma^2)) \quad a.s. N \rightarrow \infty$$

**Proof 5.7.14** By Lemma 5.7.10, we have

$$\left( \frac{\partial Q_N^{\sigma^2}(\alpha)}{\partial \alpha} \right)_{\alpha=\alpha_0} - \left( \frac{\partial O_N^{\sigma^2}(\alpha)}{\partial \alpha} \right)_{\alpha=\alpha_0} = o_p(1).$$

Thus,  $\left(\frac{\partial Q_N^{\sigma^2}(\alpha)}{\partial \alpha}\right)_{\alpha=\alpha_0}$  and  $\left(\frac{\partial O_N^{\sigma^2}(\alpha)}{\partial \alpha}\right)_{\alpha=\alpha_0}$  have same asymptotic distribution. Therefore, it remains to show that  $\sqrt{N}\left(\frac{\partial O_N^{\sigma^2}(\alpha)}{\partial \alpha}\right)_{\alpha=\alpha_0} \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, 4\mathbf{I}(\alpha_0, \sigma^2))$  a.s.  $N \rightarrow \infty$ . Moreover, note that

$$\begin{aligned} \sqrt{N}\left(\frac{\partial O_N^{\sigma^2}(\alpha)}{\partial \alpha}\right) &= \frac{2}{\sqrt{N}} \sum_{n=0}^{N-1} \mathbf{Y}_n = \frac{2}{\sqrt{N}} \sum_{n=0}^{N-1} \frac{\partial \boldsymbol{\epsilon}'_n(\alpha)}{\partial \alpha} \Sigma_{\sigma}^{-1} \boldsymbol{\epsilon}_n(\alpha) \\ &= \frac{2}{\sqrt{N}} \sum_{n=0}^{N-1} \sum_{j=0}^{\infty} \left(\mathbf{I}_d \times \mathbf{X}'_{n-j}\right) \mathbf{E}_j(\alpha) \Sigma_{\sigma}^{-1} \sum_{i=0}^{\infty} C_i(\alpha) \mathbf{X}_{n-j} \end{aligned}$$

where  $d = T(p + q + P + Q)$  and  $\mathbf{E}_j(\alpha) = (C_{1,j}(\alpha), \dots, C_{l,j}(\alpha), \dots, C_{d,j}(\alpha))'$ . Since  $\left(\frac{\partial \epsilon_{nT+v}(\alpha)}{\partial \alpha}\right)_{\alpha=\alpha_0}$  belongs to the Hilbert space  $\mathcal{H}_X(nT + v - 1)$ , the random variables  $\epsilon_{nT+v}(\alpha_0)$  and  $\left(\frac{\partial \epsilon_{nT+v}(\alpha)}{\partial \alpha}\right)_{\alpha=\alpha_0}$  are orthogonal and  $\mathbb{E}\left[\sqrt{N}\left(\frac{\partial O_N^{\sigma^2}(\alpha)}{\partial \alpha}\right)_{\alpha=\alpha_0}\right] = \mathbf{0}$ .

Now, for any positive integer  $r$ , we have

$$\sqrt{N}\left(\frac{\partial O_N^{\sigma^2}(\alpha)}{\partial \alpha}\right)_{\alpha=\alpha_0} = \frac{2}{\sqrt{N}} \sum_{n=0}^{N-1} (\mathbf{Y}_{n,r} - \mathbb{E}[\mathbf{Y}_{n,r}]) + \frac{2}{\sqrt{N}} \sum_{n=0}^{N-1} (\mathbf{Z}_{n,r} - \mathbb{E}[\mathbf{Z}_{n,r}])$$

where

$$\begin{aligned} \mathbf{Z}_{n,r} &= \mathbf{U}_{n,r} + \mathbf{V}_{n,r} \\ \mathbf{U}_{n,r} &= \sum_{i=r+1}^{\infty} \sum_{j=0}^{\infty} \left(\mathbf{I}_d \times \mathbf{X}'_{n-j}\right) \mathbf{E}_j(\alpha_0) \Sigma_{\sigma}^{-1} C_i(\alpha_0) \mathbf{X}_{n-i} \\ \mathbf{V}_{n,r} &= \sum_{i=0}^r \sum_{j=r+1}^{\infty} \left(\mathbf{I}_d \times \mathbf{X}'_{n-j}\right) \mathbf{E}_j(\alpha_0) \Sigma_{\sigma}^{-1} C_i(\alpha_0) \mathbf{X}_{n-i} \\ \text{and} \\ \mathbf{Y}_{n,r} &= \sum_{i=0}^r \sum_{j=0}^r \left(\mathbf{I}_d \times \mathbf{X}'_{n-j}\right) \mathbf{E}_j(\alpha_0) \Sigma_{\sigma}^{-1} C_i(\alpha_0) \mathbf{X}_{n-i} \end{aligned}$$

Note that  $\mathbf{Y}_{n,r}$  is function of a finite number of values of the process  $\{\mathbf{X}_n\}$ . Therefore, the process  $(\mathbf{Y}_{n,r})_{n \in \mathbb{Z}}$  satisfies the strong mixing condition in Assumption (A3). The central limit theorem for strong mixing processes (see [26]) implies that

$$\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} (\mathbf{Y}_{n,r} - \mathbb{E}[\mathbf{Y}_{n,r}]) \xrightarrow{N \rightarrow \infty} \mathcal{N}(\mathbf{0}, \tilde{\mathbf{I}}_r).$$

Moreover,  $\tilde{\mathbf{I}}_r = \lim_{N \rightarrow \infty} \text{Var}\left(\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} (\mathbf{Y}_{n,r} - \mathbb{E}[\mathbf{Y}_{n,r}])\right) \xrightarrow{r \rightarrow \infty} \mathbf{I}(\alpha_0, \sigma^2)$ .

$$\mathbb{E}\left[\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} (\mathbf{Z}_{n,r} - \mathbb{E}[\mathbf{Z}_{n,r}]) \left(\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} (\mathbf{Z}_{n,r} - \mathbb{E}[\mathbf{Z}_{n,r}])\right)'\right] \xrightarrow{r \rightarrow \infty} \mathbf{0}, \quad \forall N. \quad (5.52)$$

For  $m \in \{1, \dots, T(p + q + P + Q)\}$ , we have

$$\text{Var}\left(\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \mathbf{U}_{n,r}(m)\right) = \frac{1}{N} \sum_{n=0}^{N-1} \sum_{n'=0}^{N-1} \text{Cov}(\mathbf{U}_{n,r}(m), \mathbf{U}_{n',r}(m)) = \frac{1}{N} \sum_{h=1-N}^{N-1} (N - |h|) c_r(h) \leq \sum_{h=-\infty}^{\infty} |c_r(h)|$$

where

$$\begin{aligned} c_r(h) &= \text{Cov}(\mathbf{U}_{n,r}(m), \mathbf{U}_{n+h,r}(m)) \\ &= \sum_{i=r+1}^{\infty} \sum_{j=0}^{\infty} \sum_{i'=r+1}^{\infty} \sum_{j'=0}^{\infty} \text{Cov}\left(\mathbf{X}'_{n-j} C_{m,j}(\alpha_0) \Sigma_{\sigma}^{-1} C_i(\alpha_0) \mathbf{X}_{n-i}, \mathbf{X}'_{n+h-j'} C_{m,j'}(\alpha_0) \Sigma_{\sigma}^{-1} C_{i'}(\alpha_0) \mathbf{X}_{n+h-i'}\right) \end{aligned}$$

Consider first the case  $h \geq 0$ . For  $\lfloor h/2 \rfloor \geq r$ , using Cauchy-Shwarz inequality, we obtain that,

$$|c_r(h)| \leq M_1 \sum_{i=r+1}^{\infty} \|C_i(\alpha_0)\| \sum_{j=0}^{\infty} \|C'_{m,j}(\alpha_0)\| \sum_{i'=r+1}^{\infty} \|C_{i'}(\alpha_0)\| \sum_{j'=0}^{\infty} \|C'_{m,j'}(\alpha_0)\| \leq \Delta_1 \rho^r$$

for some positive constants  $M_1$  and  $\Delta_1$ . For  $r < \lfloor h/2 \rfloor$ , using the Cauchy-Shwarz inequality and Davydov inequality, we get that,

for some positive constants  $M_2$  and  $\Delta_2$ . The same inequality holds for  $h < 0$ . Therefore, there exists a constant  $\Delta$  such that

$$\sum_{h=-\infty}^{\infty} |c_r(h)| = \sum_{|h| \leq 2r+1} |c_r(h)| + \sum_{|h| \geq 2(r+1)} |c_r(h)| \leq \Delta r \rho^r + \Delta \rho^r + \Delta r \rho^r \sum_k [\underline{\alpha} \mathbf{X}(k)]^{\frac{1}{2+\tau}} \xrightarrow{r \rightarrow \infty} 0.$$

This implies that

$$\sup_N \text{Var} \left( \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \mathbf{U}_{n,r}(m) \right) \xrightarrow{r \rightarrow \infty} 0. \quad (5.53)$$

In a similar way, it can be shown that

$$\sup_N \text{Var} \left( \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \mathbf{V}_{n,r}(m) \right) \xrightarrow{r \rightarrow \infty} 0. \quad (5.54)$$

Finally, (5.52) follows from (5.53) and (5.54). Which completes the proof of Lemma 5.7.14.

#### Lemma 5.7.15

$$\sqrt{N} \left( \frac{\partial O_N^{\hat{\sigma}^2}(\alpha)}{\partial \alpha} - \frac{\partial O_N^{\sigma_0^2}(\alpha)}{\partial \alpha} \right)_{\alpha=\alpha_0} = \sum_{v=1}^T \left( \frac{1}{\hat{\sigma}_v^2} - \frac{1}{\sigma_{0v}^2} \right) \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \epsilon_{nT+v}(\alpha_0) \left( \frac{\partial \epsilon_{nT+v}}{\partial \alpha} \right)_{\alpha=\alpha_0} = o_p(1).$$

**Proof 5.7.15** For any  $\epsilon > 0$ , we have almost surely

$$\max_{v=1, \dots, T} \left| \frac{1}{\hat{\sigma}_v^2} - \frac{1}{\sigma_{0v}^2} \right| < \epsilon$$

for  $N$  large enough. By lemma 5.7.14, we have

$$\begin{aligned} \sqrt{N} \left( \frac{\partial O_N^{\hat{\sigma}^2}(\alpha)}{\partial \alpha} - \frac{\partial O_N^{\sigma_0^2}(\alpha)}{\partial \alpha} \right)_{\alpha=\alpha_0} &= \sum_{v=1}^T \left( \frac{1}{\hat{\sigma}_v^2} - \frac{1}{\sigma_{0v}^2} \right) \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \epsilon_{nT+v}(\alpha_0) \left( \frac{\partial \epsilon_{nT+v}}{\partial \alpha} \right)_{\alpha=\alpha_0} \\ &= \sum_{v=1}^T \epsilon \sqrt{N} \frac{\partial O_N^{\sigma_0^2}(\alpha)}{\partial \alpha} \xrightarrow{N \rightarrow \infty} 0 \end{aligned}$$

**Lemma 5.7.16** For all  $\sigma^2 > 0$ , almost surely the matrix

$$\mathbf{J}(\alpha_0, \sigma^2) = \frac{1}{2} \lim_{N \rightarrow \infty} \left[ \left( \frac{\partial^2 O_N^{\sigma^2}(\alpha)}{\partial \alpha \partial \alpha'} \right)_{\alpha=\alpha_0} \right]$$

exists and is strictly positive definite.

**Proof 5.7.16** By lemma 5.7.9,  $\left\| \frac{\partial^2 \mathbf{e}_n(\alpha)}{\partial \alpha_l \partial \alpha_m} - \frac{\partial^2 \mathbf{e}_n(\alpha)}{\partial \alpha_l \partial \alpha_m} \right\|_{N \rightarrow \infty} \rightarrow 0$ . Therefore,  $\left( \frac{\partial^2 \mathbf{Q}_N^{\omega^2}(\alpha)}{\partial \alpha_l \partial \alpha_m} \right)_{\alpha=\alpha_0}$  and  $\left( \frac{\partial^2 \mathbf{Q}_N^{\sigma^2}(\alpha)}{\partial \alpha_l \partial \alpha_m} \right)_{\alpha=\alpha_0}$  have almost surely the same limit.

By lemma 5.7.12 implies that  $\frac{\partial^2 \epsilon_n(\alpha)}{\partial \alpha_l \partial \alpha_m}$  belongs to  $L^2$ . We have

$$\begin{aligned} \left[ \left( \frac{\partial^2 \mathbf{Q}_N^{\sigma^2}(\alpha)}{\partial \alpha_l \partial \alpha_m} \right)_{\alpha=\alpha_0} \right] &= 2 \sum_{v=1}^T \sigma_v^{-2} \frac{1}{N} \sum_{n=0}^{N-1} \epsilon_{nT+v}(\alpha_0) \left( \frac{\partial^2 \epsilon_{nT+v}(\alpha)}{\partial \alpha \partial \alpha'} \right)_{\alpha=\alpha_0} \\ &\quad + 2 \sum_{v=1}^T \sigma_v^{-2} \frac{1}{N} \sum_{n=0}^{N-1} \left( \frac{\partial^2 \epsilon_{nT+v}(\alpha)}{\partial \alpha \partial \alpha'} \right)_{\alpha=\alpha_0} \left( \frac{\partial^2 \epsilon_{nT+v}(\alpha)}{\partial \alpha \partial \alpha'} \right)'_{\alpha=\alpha_0} \\ &\xrightarrow{N \rightarrow \infty} 2 \sum_{v=1}^T \sigma_v^{-2} \mathbb{E} \left[ \epsilon_{nT+v}(\alpha_0) \left( \frac{\partial^2 \epsilon_{nT+v}(\alpha)}{\partial \alpha \partial \alpha'} \right)_{\alpha=\alpha_0} \right] \\ &\quad + 2 \sum_{v=1}^T \sigma_v^{-2} \mathbb{E} \left[ \left( \frac{\partial^2 \epsilon_{nT+v}(\alpha)}{\partial \alpha \partial \alpha'} \right)_{\alpha=\alpha_0} \left( \frac{\partial^2 \epsilon_{nT+v}(\alpha)}{\partial \alpha \partial \alpha'} \right)'_{\alpha=\alpha_0} \right] \end{aligned}$$

Since  $\left( \frac{\partial^2 \epsilon_{nT+v}(\alpha)}{\partial \alpha \partial \alpha'} \right)_{\alpha=\alpha_0}$  belongs to the Hilbert space  $\mathcal{H}_X(nT+v-1)$ , the random variables  $\epsilon_{nT+v}(\alpha_0)$  and  $\left( \frac{\partial^2 \epsilon_{nT+v}(\alpha)}{\partial \alpha \partial \alpha'} \right)_{\alpha=\alpha_0}$  are orthogonal and the first term in the limit is zero. Therefore,

$$\mathbf{J}(\alpha_0, \sigma^2) = \sum_{v=1}^T \sigma_v^{-2} \mathbb{E} \left[ \left( \frac{\partial^2 \epsilon_{nT+v}(\alpha)}{\partial \alpha \partial \alpha'} \right)_{\alpha=\alpha_0} \left( \frac{\partial^2 \epsilon_{nT+v}(\alpha)}{\partial \alpha \partial \alpha'} \right)'_{\alpha=\alpha_0} \right].$$

Again, invoking the fact that  $\sigma_1^2, \dots, \sigma_T^2$  are all strictly positive and the identifiability Assumption (A1), we can conclude that  $\mathbf{J}(\alpha_0, \sigma^2)$  is a strictly positive definite matrix.

**Proof of Theorem 5.3.2.** Consider the first-order Taylor expansion around  $\alpha_0$ , we obtain

$$0 = \sqrt{N} \left( \frac{\partial \mathbf{Q}_N^{\sigma^2}(\alpha)}{\partial \alpha} \right)_{\alpha=\alpha_0} = \sqrt{N} \left( \frac{\partial \mathbf{Q}_N^{\sigma^2}(\alpha)}{\partial \alpha} \right)_{\alpha=\alpha_0} + \left[ \left( \frac{\partial^2 \mathbf{Q}_N^{\sigma^2}(\alpha)}{\partial \alpha_l \partial \alpha_m} \right)_{\alpha=\alpha_{N,l,m}^*} \right] \sqrt{N} (\hat{\alpha}_{\text{GLS}} - \alpha_0)$$

where  $\alpha_{N,l,m}^*$  is between  $\hat{\alpha}_{\text{GLS}}$  and  $\alpha_0$ . Using again a Taylor expansion around  $\alpha_0$ , we have

$$\begin{aligned} \left| \left( \frac{\partial^2 \mathbf{Q}_N^{\sigma^2}(\alpha)}{\partial \alpha_l \partial \alpha_m} \right)_{\alpha=\alpha_{N,l,m}^*} - \left( \frac{\partial^2 \mathbf{Q}_N^{\sigma^2}(\alpha)}{\partial \alpha_l \partial \alpha_m} \right)_{\alpha=\alpha_0} \right| &\leq \sup_{\alpha \in \sigma_\delta} \left\| \frac{\partial}{\partial \alpha} \left( \frac{\partial^2 \mathbf{Q}_N^{\sigma^2}(\alpha)}{\partial \alpha_l \partial \alpha_m} \right) \right\| \|\alpha_{N,l,m}^* - \alpha_0\| \\ &\xrightarrow{N \rightarrow \infty} 0 \text{ a.s.} \end{aligned}$$

This implies that, as  $N \rightarrow \infty$

$$\sqrt{N} (\hat{\alpha}_{\text{GLS}} - \alpha_0) = - \left[ \left( \frac{\partial^2 \mathbf{Q}_N^{\sigma^2}(\alpha)}{\partial \alpha_l \partial \alpha_m} \right)_{\alpha=\alpha_0} \right]^{-1} \sqrt{N} \left( \frac{\partial \mathbf{Q}_N^{\sigma^2}(\alpha)}{\partial \alpha} \right)_{\alpha=\alpha_0} + o_p(1).$$

Using Lemmas 5.7.14 and 5.7.16, we conclude the proof of Theorem 5.3.2.

□

### 5.7.3 Proof of Theorem 5.4.1

The proof of theorem 5.4.1 is based on a series lemmas. We use the multiplicative matrix norm defined by:  $\|A\| = \sup_{\|x\| \leq 1} \|Ax\| = \rho^{\frac{1}{2}}(A'A)$ , where  $A$  is a  $d_1 \times d_2$  matrix,  $\|x\|$  is the Euclidean norm of the vector  $x \in \mathbb{R}^{d_2}$ , and  $\rho(\cdot)$  denotes the spectral radius. This norm satisfies

$$\|A\|^2 = \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} a_{i,j}^2 \quad (5.55)$$

with  $a_{i,j}$  the entries of  $A \in \mathbb{R}^{k_1 \times k_2}$ . The choice of the norm is crucial for the following results to hold( with e.g the Euclidean, this result is not valid).

We denote

$$\Sigma_{Y, \underline{Y}_r} = \mathbb{E} Y_n \underline{Y}'_{r,n} \quad \Sigma_Y = \mathbb{E} Y_n Y'_n \quad \Sigma_{\underline{Y}_r} = \mathbb{E} \underline{Y}_{r,n} \underline{Y}'_{r,n}$$

where  $\underline{Y}_{r,n} = (Y_{n-1}, \dots, Y_{n-r})$ . For any  $N \geq 0$  we have

$$\begin{aligned} \hat{I}_N^{\text{SP}} &= \hat{\Phi}_r^{-1}(1) \hat{\Sigma}_{\hat{u}_r} \hat{\Phi}_r^{-1}(1) = (\hat{\Phi}_r^{-1}(1) - \Phi^{-1}(1)) \hat{\Sigma}_{\hat{u}_r} \hat{\Phi}_r^{-1}(1) + \Phi^{-1}(1) (\hat{\Sigma}_{\hat{u}_r} - \Sigma_u) \hat{\Phi}_r^{-1}(1)' \\ &\quad + \Phi^{-1}(1) \Sigma_u (\hat{\Phi}_r^{-1}(1) - \Phi^{-1}(1)') + \Phi^{-1}(1) \Sigma_u \Phi^{-1}(1)'. \end{aligned}$$

We the obtain

$$\begin{aligned} \|\hat{I}_N^{\text{SP}} - I(\alpha_0)\| &\leq \|\hat{\Phi}_r^{-1}(1) - \Phi^{-1}(1)\| \|\hat{\Sigma}_{\hat{u}_r}\| \|\hat{\Phi}_r^{-1}(1)'\| + \|\Phi^{-1}(1)\| \|\hat{\Sigma}_{\hat{u}_r} - \Sigma_u\| \|\hat{\Phi}_r^{-1}(1)'\| \\ &\quad + \|\Phi^{-1}(1)\| \|\Sigma_u\| \|\hat{\Phi}_r^{-1}(1)' - \Phi^{-1}(1)'\| \\ &\leq \|\hat{\Phi}_r^{-1}(1) - \Phi^{-1}(1)\| (\|\hat{\Sigma}_{\hat{u}_r}\| \|\hat{\Phi}_r^{-1}(1)'\| + \|\Phi^{-1}(1)\| \|\Sigma_u\|) \\ &\quad + \|\hat{\Sigma}_{\hat{u}_r} - \Sigma_u\| \|\hat{\Phi}_r^{-1}(1)'\| \|\Phi^{-1}(1)\| \\ &\leq \|\hat{\Phi}_r^{-1}(1)'\| \|\Phi - \hat{\Phi}_r(1)\| \|\Phi^{-1}(1)\| (\|\hat{\Sigma}_{\hat{u}_r}\| \|\hat{\Phi}_r^{-1}(1)'\| + \|\Phi^{-1}(1)\| \|\Sigma_u\|) \\ &\quad + \|\hat{\Sigma}_{\hat{u}_r} - \Sigma_u\| \|\hat{\Phi}_r^{-1}(1)'\| \|\Phi^{-1}(1)\|. \end{aligned} \quad (5.56)$$

It suffices to show that  $\hat{\Phi}_r(1)' \rightarrow \Phi$  and  $\hat{\Sigma}_{\hat{u}_r} \rightarrow \Sigma_u$  in probability, to have the proof of Theorem 5.4.1. Let the  $r \times 1$  vector  $\mathbf{1}_r = (1, \dots, 1)'$  and the  $rT(p+q+P+Q) \times T(p+q+P+Q)$  matrix  $\mathbb{E}_r = I_{T(p+q+P+Q)} \otimes \mathbf{1}_r$ , where  $\otimes$  denotes the matrix kronecker product and  $I_m$  the  $m \otimes m$  identity matrix. Write  $\underline{\Phi}^* = (\phi_1, \dots, \phi_r)$  where the  $\phi_i$ 's are defined by (5.16). We have

$$\begin{aligned} \|\hat{\Phi}_r(1) - \Phi(1)\| &= \left\| \sum_{k=1}^r \hat{\Phi}_{r,k} - \sum_{k=1}^r \Phi_{r,k} + \sum_{k=1}^r \Phi_{r,k} - \sum_{k=1}^{\infty} \Phi_{r,k} \right\| \\ &\leq \left\| \sum_{k=1}^r (\hat{\Phi}_{r,k} - \Phi_{r,k}) \right\| + \left\| \sum_{k=1}^r (\Phi_{r,k} - \Phi_k) \right\| + \left\| \sum_{k=r+1}^{\infty} \Phi_k \right\| \\ &\leq \left\| (\hat{\underline{\Phi}}_r - \underline{\Phi}_r) \mathbb{E}_r \right\| + \left\| (\underline{\Phi}_r^* - \underline{\Phi}_r) \mathbb{E}_r \right\| + \left\| \sum_{k=r+1}^{\infty} \Phi_k \right\| \\ &\leq \sqrt{r} (\|\hat{\underline{\Phi}}_r - \underline{\Phi}_r\| + \|\underline{\Phi}_r^* - \underline{\Phi}_r\|) + \left\| \sum_{k=r+1}^{\infty} \Phi_k \right\|. \end{aligned} \quad (5.57)$$

In view of Assumptions of Theorem 5.4.1 we have

$$\left\| \sum_{k=r+1}^{\infty} \Phi_k \right\| \leq \sum_{k=r+1}^{\infty} \|\Phi_k\| \xrightarrow{N \rightarrow \infty} 0$$

To obtain the convergence in probability of  $\hat{\Phi}_r(1)$  towards  $\Phi(1)$ , one must show the convergence in probability towards 0 of  $\sqrt{r} \|\hat{\underline{\Phi}}_r - \underline{\Phi}_r\|$  and  $\sqrt{r} \|\underline{\Phi}_r^* - \underline{\Phi}_r\|$ .

By (5.18) we obtain

$$Y_n(\alpha_0) = \underline{\Phi}_r \underline{Y}_{r,n}(\alpha_0) + u_{r,n}, \quad (5.58)$$

and

$$\Sigma_{u_r} = \text{Var}(u_{r,n}) = \mathbb{E} \left[ u_{r,n} (Y_n(\alpha_0) - \underline{\Phi}_r \underline{Y}_{r,n}(\alpha_0))' \right].$$

The vector  $u_{r,n}$  is orthogonal to  $\underline{Y}_{r,n}(\alpha_0)$ . Therefore

$$\begin{aligned} \text{Var}(u_{r,n}) &= \mathbb{E} \left[ (Y_n(\alpha_0) - \underline{\Phi}_r \underline{Y}_{r,n}(\alpha_0)) Y'_n(\alpha_0) \right] \\ &= \Sigma_Y - \underline{\Phi}_r \Sigma'_{Y, Y_r}. \end{aligned}$$

Consequently the least squares estimator of  $\Sigma_{u_r}$  can be rewritten in the form:

$$\hat{\Sigma}_{\hat{u}_r} = \hat{\Sigma}_{\hat{Y}} - \hat{\Phi}_r \hat{\Sigma}'_{\hat{Y}, \underline{Y}_r}, \quad (5.59)$$

where

$$\hat{\Sigma}_{\hat{Y}} = \frac{1}{N} \sum_{n=0}^{N-1} \hat{Y}_n \hat{Y}'_n. \quad (5.60)$$

Similar arguments combined with (5.16) yied

$$\begin{aligned} \Sigma_u &= \mathbb{E}[u_n u'_n] = \mathbb{E}[u_n Y_n(\alpha_0)'] \\ &= \mathbb{E}[Y_n(\alpha_0) Y_n(\alpha_0)'] - \sum_{k=1}^r \Phi_k \mathbb{E}[Y_{n-k}(\alpha_0) Y_n(\alpha_0)'] - \sum_{k=r+1}^{\infty} \Phi_k \mathbb{E}[Y_{n-k}(\alpha_0) Y_n(\alpha_0)'] \\ &= \Sigma_Y - \Phi_r^* \Sigma'_{Y, \underline{Y}_r} - \sum_{k=r+1}^{\infty} \Phi_k \mathbb{E}[Y_{n-k}(\alpha_0) Y_n(\alpha_0)']. \end{aligned}$$

By (5.59) we obtain

$$\begin{aligned} \|\hat{\Sigma}_{\hat{u}_r} - \Sigma_u\| &= \left\| \hat{\Sigma}_{\hat{w}} - \hat{\Phi}_r \hat{\Sigma}'_{\hat{Y}, \underline{Y}_r} - \Sigma_Y + \Phi_r^* \Sigma'_{Y, \underline{Y}_r} + \sum_{k=r+1}^{\infty} \Phi_k \mathbb{E}[Y_{n-k}(\alpha_0) Y_n(\alpha_0)'] \right\| \\ &= \left\| \hat{\Sigma}_{\hat{Y}} - \Sigma_Y - (\hat{\Phi}_r - \Phi_r^*) \hat{\Sigma}'_{\hat{Y}, \underline{Y}_r} - \Phi_r^* (\hat{\Sigma}'_{\hat{Y}, \underline{Y}_r} - \Sigma'_{Y, \underline{Y}_r}) + \sum_{k=r+1}^{\infty} \Phi_k \mathbb{E}[Y_{n-k}(\alpha_0) Y_n(\alpha_0)'] \right\| \\ &\leq \|\hat{\Sigma}_{\hat{Y}} - \Sigma_Y\| + \left\| (\hat{\Phi}_r - \Phi_r^*) (\hat{\Sigma}'_{\hat{Y}, \underline{Y}_r} - \Sigma'_{Y, \underline{Y}_r}) \right\| + \left\| (\hat{\Phi}_r - \Phi_r^*) \hat{\Sigma}'_{\hat{Y}, \underline{Y}_r} \right\| \\ &\quad + \left\| \Phi_r^* (\hat{\Sigma}'_{\hat{Y}, \underline{Y}_r} - \Sigma'_{Y, \underline{Y}_r}) \right\| + \left\| \sum_{k=r+1}^{\infty} \Phi_k \mathbb{E}[Y_{n-k}(\alpha_0) Y_n(\alpha_0)'] \right\|. \end{aligned} \quad (5.61)$$

By Lemma 5.7.3, under Assumption of Theorem 5.4.1 and (5.55) we obtain

$$\begin{aligned} \left\| \sum_{k=r+1}^{\infty} \Phi_k \mathbb{E}[Y_{n-k}(\alpha_0) Y_n(\alpha_0)'] \right\| &\leq \|\Phi_k\| \left\| \sum_{k=r+1}^{\infty} \mathbb{E}[Y_{n-k}(\alpha_0) Y_n(\alpha_0)'] \right\| \\ &\leq K \sum_{k=r+1}^{\infty} \frac{1}{k^2} \xrightarrow{N \rightarrow \infty} 0 \end{aligned}$$

and we also have

$$\|\Phi_r^*\|^2 \leq \sum_{k \geq r} \text{Tr}(\Phi_k \Phi_k') < \infty.$$

**Lemma 5.7.17** *Under the assumptions of the theorem 5.4.1 we have*

$$\sup_{r \geq 1} \{ \|\Sigma_{Y, \underline{Y}_r}\|, \|\Sigma_{Y, \underline{Y}_r}\|, \|\Sigma_{\underline{Y}_r}\| \} < \infty.$$

**Proof 5.7.17** *See Lemma 1 in the supplementary material of [13].*

**Lemma 5.7.18** *Under the assumptions of the theorem 5.4.1 there exists a finite positive constant  $K$  such that, for  $1 \leq r_1, r_2 \leq r$  and  $1 \leq m_1, m_2 \leq T(p + q + P + Q)$  we have*

$$\sup_{n \in \mathbb{Z}} \sum_{h=-\infty}^{\infty} |\text{cov}(Y_{n-r_1, m_1}(\alpha_0) Y_{n-r-2, m_2}(\alpha_0), Y_{n-r_1-h, m_1}(\alpha_0) w_{n-r_2-h, m_2}(\alpha_0))| < K.$$

**Proof 5.7.18** We can take the supremum over the integers  $n > 0$ , and write the proof in the case  $m_1 = m_2 = m$ . We have

$$\begin{aligned} & \sum_{h=-\infty}^{\infty} |\text{cov}(Y_{n-r_1, m_1}(\alpha_0) Y_{n-r_2, m_2}(\alpha_0), Y_{n-r_1-h, m_1}(\alpha_0) Y_{n-r_2-h, m_2}(\alpha_0))| \\ & \leq K^8 \sum_{h_1, \dots, h_8=0}^{\infty} \text{cov}(Y_{1, h_1, h_2} Y_{1+s, h_3, h_4}, Y_{1+h, h_5, h_6} Y_{1+s+h, h_7, h_8}) \end{aligned}$$

where

$$Y_{t, h_1, h_2} = \epsilon_{t-h_1} \epsilon_{t-h_2-m} - \mathbb{E} \epsilon_{t-h_1} \epsilon_{t-h_2-m}$$

A slight extension of the proof of Lemma 5.7.14 concludes.

**Lemma 5.7.19** Under the assumptions of the theorem 5.4.1,  $\sqrt{r} \|\hat{\Sigma}_{\underline{Y}_r} - \Sigma_{\underline{Y}_r}\|$ ,  $\sqrt{r} \|\hat{\Sigma}_{Y, \underline{Y}_r} - \Sigma_{Y, \underline{Y}_r}\|$  and  $\sqrt{r} \|\hat{\Sigma}_Y - \Sigma_Y\|$  tend to 0 in probability as  $N \rightarrow \infty$  when  $r = o(N^{\frac{1}{3}})$ .

**Proof 5.7.19** For  $1 \leq m_1, m_2 \leq T(p+q+P+Q)$  and  $1 \leq r_1, r_2 \leq r$ , the  $(\{(r_1-1)(T(p+q+P+Q)+m_1)\}, \{(r_2-1)(T(p+q+P+Q)+m_2)\})$ -th element of  $\hat{\Sigma}_{\underline{Y}_r}$  is given by:

$$\frac{1}{N} \sum_{n=0}^{N-1} Y_{n-r_1, m_1}(\alpha_0) Y_{n-r_2, m_2}(\alpha_0).$$

For all  $\beta > 0$  we use (5.55) and we obtain

$$\begin{aligned} & \mathbb{P}(\sqrt{r} \|\hat{\Sigma}_{\underline{Y}_r} - \Sigma_{\underline{Y}_r}\| \geq \beta) \leq \frac{r}{\beta^2} \mathbb{E} \|\hat{\Sigma}_{\underline{Y}_r} - \Sigma_{\underline{Y}_r}\|^2 \\ & \leq \frac{r}{\beta^2} \mathbb{E} \left\| \frac{1}{N} \sum_{n=0}^{N-1} \underline{Y}_{r,n} \underline{Y}'_{r,n} - \mathbb{E} [\underline{Y}_{r,n} \underline{Y}'_{r,n}] \right\|^2 \\ & \leq \frac{r}{\beta^2} \sum_{r_1=0}^r \sum_{r_2=0}^r \sum_{m_1=1}^{T(p+q+P+Q)} \sum_{m_2=1}^{T(p+q+P+Q)} \mathbb{E} \left( \frac{1}{N} \sum_{n=0}^{N-1} Y_{n-r_1, m_1}(\alpha_0) Y_{n-r_2, m_2}(\alpha_0) \right. \\ & \quad \left. - \mathbb{E} [Y_{n-r_1, m_1}(\alpha_0) Y_{n-r_2, m_2}(\alpha_0)] \right)^2. \end{aligned}$$

The stationarity of the process  $(Y_{n-r_1, m_1}(\alpha_0) Y_{n-r_2, m_2}(\alpha_0))_n$  and Lemma 5.7.18 imply

$$\begin{aligned} & \mathbb{P}(\sqrt{r} \|\hat{\Sigma}_{\underline{Y}_r} - \Sigma_{\underline{Y}_r}\| \geq \beta) \leq \frac{r}{\beta^2} \sum_{r_1=0}^r \sum_{r_2=0}^r \sum_{m_1=1}^{T(p+q+P+Q)} \sum_{m_2=1}^{T(p+q+P+Q)} \text{Var} \left( \frac{1}{N} \sum_{n=0}^{N-1} Y_{n-r_1, m_1}(\alpha_0) Y_{n-r_2, m_2}(\alpha_0) \right) \\ & \leq \frac{r}{N\beta^2} \sum_{r_1=0}^r \sum_{r_2=0}^r \sum_{m_1=1}^{T(p+q+P+Q)} \sum_{m_2=1}^{T(p+q+P+Q)} \sum_{n=0}^{N-1} \sum_{n'=0}^{N-1} \text{Cov}(Y_{n-r_1, m_1}(\alpha_0) Y_{n-r_2, m_2}(\alpha_0), Y_{n'-r_1, m_1}(\alpha_0) Y_{n'-r_2, m_2}(\alpha_0)) \\ & \leq \frac{r}{N\beta^2} \sum_{r_1=0}^r \sum_{r_2=0}^r \sum_{m_1=1}^{T(p+q+P+Q)} \sum_{m_2=1}^{T(p+q+P+Q)} \sum_{h=2-N}^{N-2} (N-|h|) \text{Cov}(Y_{n-r_1, m_1}(\alpha_0) Y_{n-r_2, m_2}(\alpha_0), Y_{n'-r_1, m_1}(\alpha_0) Y_{n'-r_2, m_2}(\alpha_0)) \\ & \leq \frac{r}{N\beta^2} \sum_{r_1=0}^r \sum_{r_2=0}^r \sum_{m_1=1}^{T(p+q+P+Q)} \sum_{m_2=1}^{T(p+q+P+Q)} \sup_{n \in \mathbb{Z}} \sum_{h=-\infty}^{\infty} |\text{cov}(Y_{n-r_1, m_1}(\alpha_0) Y_{n-r_2, m_2}(\alpha_0), Y_{n-r_1-h, m_1}(\alpha_0) Y_{n-r_2-h, m_2}(\alpha_0))| \\ & \leq \frac{K(T(p+q+P+Q))^2 r^3}{N\beta^2}. \end{aligned}$$

Finally we have

$$\begin{aligned} \mathbb{E} \left[ r \|\hat{\Sigma}_Y - \Sigma_Y\|^2 \right] & \leq \mathbb{E} \left[ r \|\hat{\Sigma}_{Y, \underline{Y}_r} - \Sigma_{Y, \underline{Y}_r}\|^2 \right] \leq \mathbb{E} \left[ r \|\hat{\Sigma}_{\underline{Y}_r} - \Sigma_{\underline{Y}_r}\|^2 \right] \\ & \leq \frac{K(T(p+q+P+Q))^2 r^3}{N} \xrightarrow{N \rightarrow \infty} 0, \end{aligned}$$

when  $r = o(N^{\frac{1}{3}})$ . The result follows.



We show in the following lemma that the previous lemma remains valid when we replace  $Y_n(\alpha_0)$  by  $\hat{Y}_n$ .

**Lemma 5.7.20** *Under the assumptions of the theorem 5.4.1,  $\sqrt{r} \left\| \hat{\Sigma}_{\hat{Y}_r} - \Sigma_{\underline{Y}_r} \right\|$ ,  $\sqrt{r} \left\| \hat{\Sigma}_{\hat{Y}_r, \underline{Y}_r} - \Sigma_{\underline{Y}_r, \underline{Y}_r} \right\|$  and  $\sqrt{r} \left\| \hat{\Sigma}_{\hat{Y}} - \Sigma_{\underline{Y}} \right\|$  tend to 0 in probability as  $N \rightarrow \infty$  when  $r = o(N^{\frac{1}{3}})$ .*

**Proof 5.7.20** We denote  $\hat{\Sigma}_{\underline{Y}_{r,N}}$  the matrix obtained by replacing  $e_n(\hat{\alpha}_N)$  by  $\epsilon_n(\hat{\alpha}_N)$  in  $\hat{\Sigma}_{\hat{Y}_r}$ . We have

$$\sqrt{r} \left\| \hat{\Sigma}_{\hat{Y}_r} - \Sigma_{\underline{Y}_r} \right\| \leq \sqrt{r} \left\| \hat{\Sigma}_{\hat{Y}_r} - \hat{\Sigma}_{\underline{Y}_{r,N}} \right\| + \sqrt{r} \left\| \hat{\Sigma}_{\underline{Y}_{r,N}} - \hat{\Sigma}_{\underline{Y}_r} \right\| + \sqrt{r} \left\| \hat{\Sigma}_{\underline{Y}_r} - \Sigma_{\underline{Y}_r} \right\|$$

By Lemma 5.7.19 the term  $\sqrt{r} \left\| \hat{\Sigma}_{\hat{Y}_r} - \hat{\Sigma}_{\underline{Y}_{r,N}} \right\|$  converges in probability. The lemma will be proved as soon as we show that

$$\sqrt{r} \left\| \hat{\Sigma}_{\hat{Y}_r} - \hat{\Sigma}_{\underline{Y}_{r,N}} \right\| = o_p(1) \quad (5.62)$$

$$\text{and } \sqrt{r} \left\| \hat{\Sigma}_{\underline{Y}_{r,N}} - \hat{\Sigma}_{\underline{Y}_r} \right\| = o_p(1) \quad (5.63)$$

when  $r = o(N^{\frac{1}{3}})$ . This is done in two separate steps.

### Step 1: proof of 5.62.

For all  $\beta > 0$  we have

$$\begin{aligned} \mathbb{P} \left( \sqrt{r} \left\| \hat{\Sigma}_{\hat{Y}_r} - \hat{\Sigma}_{\underline{Y}_{r,N}} \right\| \geq \beta \right) &\leq \frac{\sqrt{r}}{\beta} \mathbb{E} \left\| \hat{\Sigma}_{\hat{Y}_r} - \hat{\Sigma}_{\underline{Y}_{r,N}} \right\| \\ &\leq \frac{\sqrt{r}}{\beta} \mathbb{E} \left\| \frac{1}{N} \sum_{n=0}^{N-1} \hat{Y}_{r,n} \hat{Y}'_{r,n} - \frac{1}{N} \sum_{n=0}^{N-1} \underline{Y}_{r,n} \underline{Y}'_{r,n} \right\| \\ &\leq \frac{K\sqrt{r}}{\beta} \sum_{r_1=0}^r \sum_{r_2=0}^r \sum_{m_1=1}^{T(p+q+P+Q)} \sum_{m_2=1}^{T(p+q+P+Q)} \mathbb{E} \left| \frac{1}{N} \sum_{n=0}^{N-1} \hat{Y}_{n-r_1, m_1} \hat{Y}_{n-r_2, m_2} \right. \\ &\quad \left. - \frac{1}{N} \sum_{n=0}^{N-1} \underline{Y}_{n-r_1, m_1} \underline{Y}_{n-r_2, m_2} \right|, \end{aligned}$$

where

$$\underline{Y}_{n,m}^{(N)} = \sum_{v=1}^T \sigma_v^{-2} \epsilon_{nT+v}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_m} \epsilon_{nT+v}(\hat{\alpha}_N) \quad \text{and} \quad \underline{Y}_{r,n}^{(N)} = \left( Y_{n-1}^{(N)}, \dots, Y_{n-r}^{(N)} \right)'$$

It is follow that

$$\begin{aligned} \mathbb{P} \left( \sqrt{r} \left\| \hat{\Sigma}_{\hat{Y}_r} - \hat{\Sigma}_{\underline{Y}_{r,N}} \right\| \geq \beta \right) &\leq \frac{K\sqrt{r}}{N\beta} \sum_{v=1}^T \left( \frac{1}{\hat{\sigma}_v^2} - \frac{1}{\sigma_v^2} \right) \sum_{r_1=0}^r \sum_{r_2=0}^r \sum_{m_1=1}^{T(p+q+P+Q)} \sum_{m_2=1}^{T(p+q+P+Q)} \\ &\quad \mathbb{E} \left| \sum_{n=0}^{N-1} \hat{A}_{n,r_1} \hat{A}_{n,r_2} - A_{n,r_1} A_{n,r_2} \right| \quad (5.64) \end{aligned}$$

where

$$A_{n,r_i} = \epsilon_{nT+v-r_i}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_{m_i}} \epsilon_{nT+v-r_i}(\hat{\alpha}_N), \text{ text for } i = \{1, 2\}.$$

Now notice that

$$\begin{aligned}
 & e_{nT+v-r_1}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_{m_1}} e_{nT+v-r_1}(\hat{\alpha}_N) e_{nT+v-r_2}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_{m_2}} e_{nT+v-r_2}(\hat{\alpha}_N) - \\
 & \epsilon_{nT+v-r_1}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\hat{\alpha}_N) \epsilon_{nT+v-r_2}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_{m_2}} \epsilon_{nT+v-r_2}(\hat{\alpha}_N) \\
 & = (e_{nT+v-r_1}(\hat{\alpha}_N) - \epsilon_{nT+v-r_1}(\hat{\alpha}_N)) \frac{\partial}{\partial \alpha_{m_1}} e_{nT+v-r_1}(\hat{\alpha}_N) e_{nT+v-r_2}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_{m_2}} e_{nT+v-r_2}(\hat{\alpha}_N) \\
 & + \epsilon_{nT+v-r_1}(\hat{\alpha}_N) \left( \frac{\partial}{\partial \alpha_{m_1}} e_{nT+v-r_1}(\hat{\alpha}_N) - \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\hat{\alpha}_N) \right) e_{nT+v-r_2}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_{m_2}} e_{nT+v-r_2}(\hat{\alpha}_N) \\
 & + \epsilon_{nT+v-r_1}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\hat{\alpha}_N) (e_{nT+v-r_2}(\hat{\alpha}_N) - \epsilon_{nT+v-r_2}(\hat{\alpha}_N)) \frac{\partial}{\partial \alpha_{m_2}} e_{nT+v-r_2}(\hat{\alpha}_N) \\
 & + \epsilon_{nT+v-r_1}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\hat{\alpha}_N) \epsilon_{nT+v-r_2} \left( \frac{\partial}{\partial \alpha_{m_2}} e_{nT+v-r_2}(\hat{\alpha}_N) - \frac{\partial}{\partial \alpha_{m_2}} \epsilon_{nT+v-r_2}(\hat{\alpha}_N) \right).
 \end{aligned}$$

We replace the above identity in (5.64) and we obtain by Hölder's inequality that

$$\mathbb{P} \left( \sqrt{r} \left\| \hat{\Sigma}_{\underline{Y}_r} - \hat{\Sigma}_{\underline{Y}_r, N} \right\| \geq \beta \right) \leq \frac{K\sqrt{r}}{N\beta} \sum_{v=1}^T \left( \frac{1}{\hat{\sigma}_v} - \frac{1}{\sigma_v} \right) \sum_{r_1=1}^r \sum_{r_2=1}^r \sum_{m_1=1}^{T(p+q+P+Q)} \sum_{m_2=1}^{T(p+q+P+Q)} \mathbb{E} (D_{N,1} + D_{N,2} + D_{N,3} + D_{N,4}) \quad (5.65)$$

where

$$\begin{aligned}
 D_{N,1} &= \sum_{n=0}^{N-1} (e_{nT+v-r_1}(\hat{\alpha}_N) - \epsilon_{nT+v-r_1}(\hat{\alpha}_N)) \frac{\partial}{\partial \alpha_{m_1}} e_{nT+v-r_1}(\hat{\alpha}_N) e_{nT+v-r_2}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_{m_2}} e_{nT+v-r_2}(\hat{\alpha}_N) \\
 D_{N,2} &= \sum_{n=0}^{N-1} \epsilon_{nT+v-r_1}(\hat{\alpha}_N) \left( \frac{\partial}{\partial \alpha_{m_1}} e_{nT+v-r_1}(\hat{\alpha}_N) - \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\hat{\alpha}_N) \right) e_{nT+v-r_2}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_{m_2}} e_{nT+v-r_2}(\hat{\alpha}_N) \\
 D_{N,3} &= \sum_{n=0}^{N-1} \epsilon_{nT+v-r_1}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\hat{\alpha}_N) (e_{nT+v-r_2}(\hat{\alpha}_N) - \epsilon_{nT+v-r_2}(\hat{\alpha}_N)) \frac{\partial}{\partial \alpha_{m_2}} e_{nT+v-r_2}(\hat{\alpha}_N) \\
 D_{N,4} &= \sum_{n=0}^{N-1} \epsilon_{nT+v-r_1}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\hat{\alpha}_N) \epsilon_{nT+v-r_2} \left( \frac{\partial}{\partial \alpha_{m_2}} e_{nT+v-r_2}(\hat{\alpha}_N) - \frac{\partial}{\partial \alpha_{m_2}} \epsilon_{nT+v-r_2}(\hat{\alpha}_N) \right).
 \end{aligned}$$

We have

$$\begin{aligned}
 |D_{N,1}| &\leq K \sum_{n=0}^{N-1} |e_{nT+v-r_1}(\hat{\alpha}_N) - \epsilon_{nT+v-r_1}(\hat{\alpha}_N)| \\
 &\leq K \left( r + \sum_{n=r_1+1}^{N-1} \rho^n \right).
 \end{aligned}$$

Then we obtain

$$D_{N,1} \leq Kr. \quad (5.66)$$

The same calculations hold for the terms  $D_{N,2}$ ,  $D_{N,3}$  and  $D_{N,4}$ . Thus

$$|D_{N,1}| + |D_{N,2}| + |D_{N,3}| + |D_{N,4}| \leq Kr \quad (5.67)$$

and reporting this estimation in (5.65) and using (5.48), we have

$$\mathbb{P} \left( \sqrt{r} \left\| \hat{\Sigma}_{\underline{Y}_r} - \hat{\Sigma}_{\underline{Y}_r, N} \right\| \geq \beta \right) \leq \epsilon \frac{K\sqrt{r}}{N\beta} r T^2 (p+q+P+Q)^2 \leq \frac{Kr^{\frac{5}{2}}}{N}.$$

Then the sequence  $\sqrt{r} \left\| \hat{\Sigma}_{\underline{Y}_r} - \hat{\Sigma}_{\underline{Y}_r, N} \right\|$  converge in probability to 0 as  $N \rightarrow \infty$  when  $r = r(N) = o(N^{\frac{5}{2}})$

**Step 2: proof of (5.63).**

First we follow the same approach than in the previous step. We have

$$\begin{aligned}
\left\| \hat{\Sigma}_{\underline{Y}_{r,N}} - \hat{\Sigma}_{\underline{Y}_r} \right\|^2 &= \left\| \frac{1}{N} \sum_{n=0}^{N-1} \underline{Y}_{r,n} \underline{Y}'_{r,n} - \underline{Y}_{r,n} \underline{Y}'_{r,n} \right\|^2 \\
&\leq \frac{K\sqrt{r}}{\beta} \sum_{r_1=0}^r \sum_{r_2=0}^r \sum_{m_1=1}^{T(p+q+P+Q)} \sum_{m_2=1}^{T(p+q+P+Q)} \left( \frac{1}{N} \sum_{n=0}^{N-1} \Upsilon_{n-r_1, m_1}^{(n)} \Upsilon_{n-r_2, m_2}^{(n)} \right. \\
&\quad \left. - \Upsilon_{n-r_1, r_1} \Upsilon_{n-r_2, m_2} \right)^2 \\
&\leq \frac{K\sqrt{r}}{\beta} \sum_{r_1=0}^r \sum_{r_2=0}^r \sum_{m_1=1}^{T(p+q+P+Q)} \sum_{m_2=1}^{T(p+q+P+Q)} \sum_{v=1}^T \left( \frac{1}{\hat{\sigma}_v^2} - \frac{1}{\sigma_v^2} \right)^2 \\
&\quad \times \frac{1}{N} \sum_{n=0}^{N-1} \left( A_{n, r_1, m_1}(\hat{\alpha}_N) A_{n, r_1} A_{n, r_2}(\hat{\alpha}_N) - A_{n, r_1}(\alpha_0) A_{n, r_2}(\alpha_0) \right)^2.
\end{aligned}$$

Since

$$\begin{aligned}
&\epsilon_{nT+v-r_1}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\hat{\alpha}_N) \epsilon_{nT+v-r_2}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_{m_2}} \epsilon_{nT+v-r_2}(\hat{\alpha}_N) - \\
&\quad \epsilon_{nT+v-r_1}(\alpha_0) \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\alpha_0) \epsilon_{nT+v-r_2}(\alpha_0) \frac{\partial}{\partial \alpha_{m_2}} \epsilon_{nT+v-r_2}(\alpha_0) \\
&= \left( \epsilon_{nT+v-r_1}(\hat{\alpha}_N) - \epsilon_{nT+v-r_1}(\alpha_0) \right) \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\hat{\alpha}_N) \epsilon_{nT+v-r_2}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_{m_2}} \epsilon_{nT+v-r_2}(\hat{\alpha}_N) \\
&+ \epsilon_{nT+v-r_1}(\alpha_0) \left( \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\hat{\alpha}_N) - \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\alpha_0) \right) \epsilon_{nT+v-r_2}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_{m_2}} \epsilon_{nT+v-r_2}(\hat{\alpha}_N) \\
&\quad + \epsilon_{nT+v-r_1}(\alpha_0) \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\alpha_0) \left( \epsilon_{nT+v-r_2}(\hat{\alpha}_N) - \epsilon_{nT+v-r_2}(\alpha_0) \right) \frac{\partial}{\partial \alpha_{m_2}} \epsilon_{nT+v-r_2}(\hat{\alpha}_N) \\
&+ \epsilon_{nT+v-r_1}(\alpha_0) \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\alpha_0) \epsilon_{nT+v-r_2}(\alpha_0) \left( \frac{\partial}{\partial \alpha_{m_2}} \epsilon_{nT+v-r_2}(\hat{\alpha}_N) - \frac{\partial}{\partial \alpha_{m_2}} \epsilon_{nT+v-r_2}(\alpha_0) \right)
\end{aligned}$$

one has

$$\left\| \hat{\Sigma}_{\underline{Y}_{r,N}} - \hat{\Sigma}_{\underline{Y}_r} \right\| \leq \epsilon \frac{K\sqrt{r}}{\beta} \sum_{r_1=0}^r \sum_{r_2=0}^r \sum_{m_1=1}^{T(p+q+P+Q)} (E_{N,1} + E_{N,2} + E_{N,3} + E_{N,4}) \quad (5.68)$$

where

$$\begin{aligned}
E_{N,1} &= \frac{1}{N} \sum_{n=0}^{N-1} \left| \epsilon_{nT+v-r_1}(\hat{\alpha}_N) - \epsilon_{nT+v-r_1}(\alpha_0) \right| \left| \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\hat{\alpha}_N) \right| \left| \epsilon_{nT+v-r_2}(\hat{\alpha}_N) \right| \left| \frac{\partial}{\partial \alpha_{m_2}} \epsilon_{nT+v-r_2}(\hat{\alpha}_N) \right| \\
E_{N,2} &= \frac{1}{N} \sum_{n=0}^{N-1} \left| \epsilon_{nT+v-r_1}(\alpha_0) \right| \left| \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\hat{\alpha}_N) - \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\alpha_0) \right| \left| \epsilon_{nT+v-r_2}(\hat{\alpha}_N) \right| \left| \frac{\partial}{\partial \alpha_{m_2}} \epsilon_{nT+v-r_2}(\hat{\alpha}_N) \right| \\
E_{N,3} &= \frac{1}{N} \sum_{n=0}^{N-1} \left| \epsilon_{nT+v-r_1}(\alpha_0) \right| \left| \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\alpha_0) \right| \left| \epsilon_{nT+v-r_2}(\hat{\alpha}_N) - \epsilon_{nT+v-r_2}(\alpha_0) \right| \left| \frac{\partial}{\partial \alpha_{m_2}} \epsilon_{nT+v-r_2}(\hat{\alpha}_N) \right| \\
E_{N,4} &= \frac{1}{N} \sum_{n=0}^{N-1} \left| \epsilon_{nT+v-r_1}(\alpha_0) \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\alpha_0) \right| \left| \epsilon_{nT+v-r_2}(\alpha_0) \right| \left| \frac{\partial}{\partial \alpha_{m_2}} \epsilon_{nT+v-r_2}(\hat{\alpha}_N) - \frac{\partial}{\partial \alpha_{m_2}} \epsilon_{nT+v-r_2}(\alpha_0) \right|.
\end{aligned}$$

Taylor expansions around  $\alpha_0$  yield that there exists  $\underline{\alpha}$  and  $\bar{\alpha}$  between  $\hat{\alpha}_N$  and  $\alpha_0$  such that

$$\left| \epsilon_{nT+v-r_1}(\hat{\alpha}_N) - \epsilon_{nT+v-r_1}(\alpha_0) \right| \leq r_{nT+v} \|\hat{\alpha}_N - \alpha_0\|$$

and

$$\left| \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\hat{\alpha}_N) - \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\alpha_0) \right| \leq s_{nT+v} \|\hat{\alpha}_N - \alpha_0\|$$

with  $r_{nT+v} = \|\partial \epsilon_{nT+v}(\underline{\alpha}) / \partial \underline{\alpha}\|$  and  $s_{nT+v} = \|\partial^2 \epsilon_{nT+v}(\bar{\alpha}) / \partial \alpha' \partial \alpha_m\|$ . Using the fact that

$$\mathbb{E} \left| r_{nT+v-r_1} \frac{\partial}{\partial \alpha_{m_1}} \epsilon_{nT+v-r_1}(\hat{\alpha}_N) \epsilon_{nT+v-r_2}(\hat{\alpha}_N) \frac{\partial}{\partial \alpha_{m_2}} \epsilon_{nT+v-r_2}(\hat{\alpha}_N) \right| < \infty$$

and that  $(\sqrt{N}(\hat{\alpha}_N - \alpha_0))_N$  is a tight sequence (which implies that  $\|\hat{\alpha}_N - \alpha_0\| = O_{\mathbb{P}}(1/\sqrt{N})$ ), we deduce that

$$E_{N,1} = O_{\mathbb{P}} \left( \frac{1}{\sqrt{N}} \right).$$

The same arguments are valid for  $E_{N,2}$ ,  $E_{N,3}$  and  $E_{N,4}$ . Consequently  $E_{N,1} + E_{N,2} + E_{N,3} + E_{N,4} = O_{\mathbb{P}}(1/\sqrt{N})$  and (5.68) yields

$$\left\| \hat{\Sigma}_{\underline{Y}_r, N} - \hat{\Sigma}_{\underline{Y}_r} \right\|^2 = O_{\mathbb{P}} \left( \frac{r^2}{N} \right)$$

when  $r = o(N^{\frac{1}{3}})$  we finally obtain  $\sqrt{r} \left\| \hat{\Sigma}_{\underline{Y}_r, N} - \hat{\Sigma}_{\underline{Y}_r} \right\| = O_{\mathbb{P}}(1)$ .

**Lemma 5.7.21** Under the assumptions of the theorem 5.4.1, we have

$$\sqrt{r} \left\| \underline{\Phi}_r^* - \underline{\Phi}_r \right\| = O_{\mathbb{P}}(1) \text{ as } r \rightarrow \infty.$$

**Proof 5.7.21** Recall that (5.16) and (5.58) we have

$$\underline{Y}_n(\alpha_0) = \underline{\Phi}_r \underline{Y}_{r,n} + u_{r,n} = \underline{\Phi}_r^* \underline{Y}_{r,n} + \sum_{k=r+1}^{\infty} \Phi_r Y_{n-k}(\alpha_0) + u_n := \underline{\Phi}_r^* \underline{Y}_{r,n} + u_{r,n}^*.$$

By the orthogonality conditions in (5.16) and (5.58), one has

$$\Sigma_{u_r^*, \underline{Y}_r} := \mathbb{E} \left[ u_{r,n}^* \underline{Y}'_{r,n} \right] = \mathbb{E} \left[ \left( \underline{Y}_n(\alpha_0) - \underline{\Phi}_r^* \underline{Y}_{r,n} \right) \underline{Y}'_{r,n} \right] = \mathbb{E} \left[ \left( \underline{\Phi}_r \underline{Y}_{r,n} + u_{r,n} - \underline{\Phi}_r^* \underline{Y}_{r,n} \right) \underline{Y}'_{r,n} \right] = (\underline{\Phi}_r - \underline{\Phi}_r^*) \Sigma_{\underline{Y}_r}$$

and consequently

$$\underline{\Phi}_r^* - \underline{\Phi}_r = \Sigma_{u_r^*, \underline{Y}_r} \Sigma_{\underline{Y}_r}^{-1} \tag{5.69}$$

Using Lemmas 5.7.17, 5.7.18 and (5.69) implies that

$$\begin{aligned} \mathbb{P}(\sqrt{r} \|\underline{\Phi}_r^* - \underline{\Phi}_r\| \geq \beta) &\leq \frac{\sqrt{r}}{\beta} \left\| \Sigma_{u_r^*, \underline{Y}_r} \right\| \left\| \Sigma_{\underline{Y}_r}^{-1} \right\| \\ &\leq \kappa \frac{\sqrt{r}}{\beta} \left\| \mathbb{E} \left[ \left( \sum_{k=r+1}^{\infty} \Phi_r Y_{n-k}(\alpha_0) + u_n \right) \underline{Y}'_{r,n} \right] \right\| \\ &\leq \kappa \frac{\sqrt{r}}{\beta} \sum_{k \geq r+1} \|\Phi_k\| \left\| \mathbb{E} \left[ Y_{n-k}(\alpha_0) \underline{Y}'_{r,n} \right] \right\| \\ &\leq \kappa \frac{\sqrt{r}}{\beta} \sum_{l \geq 1} \|\Phi_{l+r}\| \left\| \mathbb{E} \left[ Y_{n-k}(\alpha_0) (Y'_{n-k}(\alpha_0), \dots, Y'_{n-r}) \right] \right\| \\ &\leq \kappa \frac{\sqrt{r}}{\beta} \sum_{l \geq 1} \|\Phi_{l+r}\| \left( \sum_{j=1}^{T(p+q+P+Q)} \sum_{k=1}^{T(p+q+P+Q)} \sum_{r_1}^r \left| \mathbb{E} \left[ Y_{n-r-l,j}(\alpha_0) Y_{n-r-l,k}(\alpha_0) \right] \right|^2 \right)^{\frac{1}{2}} \\ &\leq \kappa \frac{\sqrt{r}}{\beta} \sum_{l \geq 1} \|\Phi_{l+r}\| \left( \sum_{j=1}^{T(p+q+P+Q)} \sum_{k=1}^{T(p+q+P+Q)} \sum_{r_1}^r \left| \mathbb{E} \left[ Y_{n-r-l,j}^2(\alpha_0) \right] \left[ Y_{n-r-l,k}^2(\alpha_0) \right] \right|^2 \right)^{\frac{1}{2}} \\ &\leq \frac{\kappa T(p+q+P+Q)r}{\beta} \sum_{l \geq 1} \|\Phi_{l+r}\|. \end{aligned}$$

By the assumptions of Theorem 5.4.1,  $\sum_{l \geq 1} \|\Phi_{l+r}\| = o(1)$  as  $r \rightarrow \infty$ . The proof of the lemma follows.

**Lemma 5.7.22** *Under the assumptions of the theorem (5.4.1) we have*

$$\sqrt{r} \left\| \hat{\Sigma}_{\underline{Y}_r}^{-1} - \Sigma_{\underline{Y}_r}^{-1} \right\| = o_{\mathbb{P}}(1)$$

as  $N \rightarrow \infty$  when  $r = o(N^{\frac{1}{3}})$  and  $r \rightarrow \infty$

**Proof 5.7.22** *We have*

$$\left\| \hat{\Sigma}_{\underline{Y}_r}^{-1} - \Sigma_{\underline{Y}_r}^{-1} \right\| \leq \left( \left\| \hat{\Sigma}_{\underline{Y}_r}^{-1} - \Sigma_{\underline{Y}_r}^{-1} \right\| + \left\| \Sigma_{\underline{Y}_r}^{-1} \right\| \right) \left\| \Sigma_{\underline{Y}_r} - \hat{\Sigma}_{\underline{Y}_r} \right\| \left\| \Sigma_{\underline{Y}_r}^{-1} \right\|.$$

and by induction we obtain

$$\left\| \hat{\Sigma}_{\underline{Y}_r}^{-1} - \Sigma_{\underline{Y}_r}^{-1} \right\| \leq \left\| \Sigma_{\underline{Y}_r}^{-1} \right\| \sum_{k=1}^{\infty} \left\| \Sigma_{\underline{Y}_r} - \hat{\Sigma}_{\underline{Y}_r} \right\|^k \left\| \Sigma_{\underline{Y}_r}^{-1} \right\|^k.$$

We have

$$\begin{aligned} \mathbb{P} \left( \sqrt{r} \left\| \hat{\Sigma}_{\underline{Y}_r}^{-1} - \Sigma_{\underline{Y}_r}^{-1} \right\| > \beta \right) &\leq \mathbb{P} \left( \left\| \Sigma_{\underline{Y}_r}^{-1} \right\| \sum_{k=1}^{\infty} \left\| \Sigma_{\underline{Y}_r} - \hat{\Sigma}_{\underline{Y}_r} \right\|^k \left\| \Sigma_{\underline{Y}_r}^{-1} \right\|^k > \beta \right) \\ &\leq \mathbb{P} \left( \left\| \Sigma_{\underline{Y}_r}^{-1} \right\| \sum_{k=1}^{\infty} \left\| \Sigma_{\underline{Y}_r} - \hat{\Sigma}_{\underline{Y}_r} \right\|^k \left\| \Sigma_{\underline{Y}_r}^{-1} \right\|^k > \beta \text{ and } \left\| \Sigma_{\underline{Y}_r} - \hat{\Sigma}_{\underline{Y}_r} \right\| \left\| \Sigma_{\underline{Y}_r}^{-1} \right\| < 1 \right) \\ &\quad + \mathbb{P} \left( \left\| \Sigma_{\underline{Y}_r}^{-1} \right\| \sum_{k=1}^{\infty} \left\| \Sigma_{\underline{Y}_r} - \hat{\Sigma}_{\underline{Y}_r} \right\|^k \left\| \Sigma_{\underline{Y}_r}^{-1} \right\|^k > \beta \text{ and } \left\| \Sigma_{\underline{Y}_r} - \hat{\Sigma}_{\underline{Y}_r} \right\| \left\| \Sigma_{\underline{Y}_r}^{-1} \right\| \geq 1 \right) \\ &\leq \mathbb{P} \left( \frac{\sqrt{r} \left\| \Sigma_{\underline{Y}_r}^{-1} \right\|^2 \left\| \Sigma_{\underline{Y}_r} - \hat{\Sigma}_{\underline{Y}_r} \right\|}{1 - \left\| \Sigma_{\underline{Y}_r} - \hat{\Sigma}_{\underline{Y}_r} \right\| \left\| \Sigma_{\underline{Y}_r}^{-1} \right\|} \right) + \mathbb{P} \left( \sqrt{r} \left\| \Sigma_{\underline{Y}_r} - \hat{\Sigma}_{\underline{Y}_r} \right\| \left\| \Sigma_{\underline{Y}_r}^{-1} \right\| \geq 1 \right) \\ &\leq \mathbb{P} \left( \sqrt{r} \left\| \Sigma_{\underline{Y}_r} - \hat{\Sigma}_{\underline{Y}_r} \right\| > \frac{\beta}{\left\| \Sigma_{\underline{Y}_r}^{-1} \right\|^2 + \beta r^{-\frac{1}{2}} \left\| \Sigma_{\underline{Y}_r}^{-1} \right\|} \right) \\ &\quad + \mathbb{P} \left( \sqrt{r} \left\| \Sigma_{\underline{Y}_r} - \hat{\Sigma}_{\underline{Y}_r} \right\| \geq \left\| \Sigma_{\underline{Y}_r}^{-1} \right\| \right) \end{aligned}$$

Lemmas 5.7.17 and 5.7.20 imply the result.

**Lemma 5.7.23** *Under the assumptions of the theorem 5.4.1, we have*

$$\sqrt{r} \left\| \hat{\Phi}_r - \Phi_r \right\| = O_{\mathbb{P}}(1) \text{ as } r \rightarrow \infty \text{ and } r = o(N^{\frac{1}{3}}).$$

**Proof 5.7.23** Lemmas 5.7.17 and 5.7.22 yield

$$\left\| \hat{\Sigma}_{\underline{Y}_r}^{-1} \right\| \leq \left\| \hat{\Sigma}_{\underline{Y}_r}^{-1} - \Sigma_{\underline{Y}_r}^{-1} \right\| + \left\| \Sigma_{\underline{Y}_r}^{-1} \right\| = O_{\mathbb{P}}(1). \quad (5.70)$$

By (5.58), we have

$$0 = \mathbb{E} \left[ u_{r,n} \underline{Y}'_{r,n} \right] = \mathbb{E} \left[ \left( \underline{Y}_n(\alpha_0) - \Phi_r \underline{Y}'_{r,n} \right) \underline{Y}'_{r,n} \right] = \Sigma_{\underline{Y}, \underline{Y}_r} - \Phi_r \Sigma_{\underline{Y}_r}$$

and so we have  $\Phi_r = \Sigma_{\underline{Y}, \underline{w}_r} \Sigma_{\underline{Y}_r}^{-1}$ . Lemmas 5.7.17, 5.7.20, 5.7.22 and (5.70) imply

$$\begin{aligned} \sqrt{r} \left\| \hat{\Phi}_r - \Phi_r \right\| &= \sqrt{r} \left\| \hat{\Sigma}_{\hat{\underline{Y}}, \hat{\underline{Y}}_r} \hat{\Sigma}_{\hat{\underline{Y}}_r}^{-1} - \Sigma_{\underline{Y}, \underline{Y}_r} \Sigma_{\underline{Y}_r}^{-1} \right\| \\ &= \sqrt{r} \left\| \left( \hat{\Sigma}_{\hat{\underline{Y}}, \hat{\underline{Y}}_r} - \Sigma_{\underline{Y}, \underline{Y}_r} \right) \hat{\Sigma}_{\hat{\underline{Y}}_r}^{-1} + \Sigma_{\underline{Y}, \underline{Y}_r} \left( \hat{\Sigma}_{\hat{\underline{Y}}_r}^{-1} - \Sigma_{\underline{Y}_r}^{-1} \right) \right\| \\ &= o_{\mathbb{P}}(1). \end{aligned}$$

The lemma is proved

**Proof of Theorem 5.4.1**

**Proof 5.7.24** Since by Lemma 5.7.20 we have  $\left\| \hat{\Sigma}_{\hat{\underline{Y}}} - \Sigma_{\underline{Y}} \right\| = o_{\mathbb{P}}(r^{-\frac{1}{2}}) = o_{\mathbb{P}}(1)$  and  $\left\| \hat{\Sigma}_{\hat{\underline{Y}}, \hat{\underline{Y}}_r} - \Sigma_{\underline{Y}, \underline{Y}_r} \right\| = o_{\mathbb{P}}(r^{-\frac{1}{2}}) = o_{\mathbb{P}}(1)$ , and by Lemma 5.7.21:  $\left\| \hat{\Phi}_r - \Phi_r \right\| = o_{\mathbb{P}}(r^{-\frac{1}{2}}) = o_{\mathbb{P}}(1)$ . The theorem 5.4.1 is then proved.

## Bibliography

- [1] Aknouche, A. and Bibi, A. (2009). Quasi-maximum likelihood estimation of periodic garch and periodic arma-garch processes. *Journal of Time Series Analysis*, 30(1):19–46. [151](#)
- [2] Akutowicz, E. J. (1958). On an explicit formula in linear least squares prediction. *Mathematica Scandinavica*, pages 261–266. [156](#)
- [3] Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3):817–858. [156](#), [161](#)
- [4] Basawa, I. V. and Lund, R. (2001). Large sample properties of parameter estimates for periodic ARMA models. *J. Time Ser. Anal.*, 22(6):651–663. [150](#)
- [5] Basawa, I. V., Lund, R., and Shao, Q. (2004). First-order seasonal autoregressive processes with periodically varying parameters. *Statist. Probab. Lett.*, 67(4):299–306. [150](#), [151](#), [154](#), [155](#)
- [6] Battaglia, F., Cucina, D., and Rizzo, M. (2018). A generalization of periodic autoregressive models for seasonal time series. Technical report, Tech. Rep. 2, Department of Statistical Sciences, University La Sapienza . . . . [150](#)
- [7] Battaglia, F., Cucina, D., and Rizzo, M. (2020). Parsimonious periodic autoregressive models for time series with evolving trend and seasonality. *Statistics and Computing*, 30(1):77–91. [150](#)
- [8] Berk, K. N. (1974). Consistent autoregressive spectral estimates. *Ann. Statist.*, 2:489–502. Collection of articles dedicated to Jerzy Neyman on his 80th birthday. [156](#), [161](#)
- [9] Bollerslev, T. and Ghysels, E. (1996). Periodic autoregressive conditional heteroscedasticity. *Journal of Business & Economic Statistics*, 14(2):139–151. [151](#)
- [10] Boubacar Mainassara, Y. (2011). Multivariate portmanteau test for structural VARMA models with uncorrelated but non-independent error terms. *J. Statist. Plann. Inference*, 141(8):2961–2975. [152](#)
- [11] Boubacar Maïnassara, Y. (2012). Selection of weak VARMA models by modified Akaike’s information criteria. *J. Time Series Anal.*, 33(1):121–130. [152](#)
- [12] Boubacar Mainassara, Y., Carbon, M., and Francq, C. (2012a). Computing and estimating information matrices of weak ARMA models. *Comput. Statist. Data Anal.*, 56(2):345–361. [156](#)
- [13] Boubacar Mainassara, Y., Carbon, M., and Francq, C. (2012b). Computing and estimating information matrices of weak ARMA models. *Comput. Statist. Data Anal.*, 56(2):345–361. [183](#)
- [14] Boubacar Mainassara, Y. and Francq, C. (2011). Estimating structural VARMA models with uncorrelated but non-independent error terms. *J. Multivariate Anal.*, 102(3):496–505. [152](#)
- [15] Boubacar Maïnassara, Y. and Kokonendji, C. C. (2016). Modified Schwarz and Hannan-Quinn information criteria for weak VARMA models. *Stat. Inference Stoch. Process.*, 19(2):199–217. [152](#)
- [16] Boubacar Maïnassara, Y. and Saussereau, B. (2018). Diagnostic checking in multivariate arma models with dependent errors using normalized residual autocorrelations. *J. Amer. Statist. Assoc.*, 113(524):1813–1827. [152](#), [161](#)

- [17] Box, G. E. P. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J. Amer. Statist. Assoc.*, 65:1509–1526. [158](#)
- [18] Brockwell, P. J. and Davis, R. A. (1991). *Time series: theory and methods*. Springer Series in Statistics. Springer-Verlag, New York, second edition. [152](#), [156](#)
- [19] Davydov, J. A. (1968). The convergence of distributions which are generated by stationary random processes. *Teor. Veroyatnost. i Primenen.*, 13:730–737. [178](#)
- [20] den Haan, W. J. and Levin, A. T. (1997). A practitioner's guide to robust covariance matrix estimation. In *Robust inference*, volume 15 of *Handbook of Statist.*, pages 299–342. North-Holland, Amsterdam. [156](#), [161](#)
- [21] Francq, C., Roy, R., and Saidi, A. (2011a). Asymptotic properties of weighted least squares estimation in weak PARMA models. *J. Time Series Anal.*, 32(6):699–723. [151](#), [154](#), [168](#)
- [22] Francq, C., Roy, R., and Saidi, A. (2011b). Asymptotic properties of weighted least squares estimation in weak PARMA models. *J. Time Series Anal.*, 32(6):699–723. [152](#)
- [23] Francq, C. and Zakoïan, J.-M. (2007). HAC estimation and strong linearity testing in weak ARMA models. *J. Multivariate Anal.*, 98(1):114–144. [156](#)
- [24] Giovanis, E. et al. (2014). The Turn-of-the-Month-Effect: Evidence from Periodic Generalized Autoregressive Conditional Heteroskedasticity (PGARCH) Model. *International Journal of Economic Sciences and Applied Research*, 7(3):43–61. [151](#)
- [25] Hipel, K. and McLeod, A. I. (1994). *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier, Amsterdam. [149](#), [150](#), [158](#)
- [26] Ibragimov, I. A. (1962). Some limit theorems for stationary processes. *Teor. Veroyatnost. i Primenen.*, 7:361–392. [179](#)
- [27] Imhof, J. P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika*, 48:419–426. [161](#)
- [28] Jones, R. H. and Brelsford, W. M. (1967). Time series with periodic structure. *Biometrika*, 54(3-4):403–408. [150](#)
- [29] Katayama, N. (2012). Chi-squared portmanteau tests for structural VARMA models with uncorrelated errors. *J. Time Series Anal.*, 33(6):863–872. [152](#)
- [30] Kuan, C.-M. and Lee, W.-M. (2006). Robust M tests without consistent estimation of the asymptotic covariance matrix. *J. Amer. Statist. Assoc.*, 101(475):1264–1275. [161](#)
- [31] Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2):pp. 297–303. [158](#)
- [32] Lobato, I. N. (2001). Testing that a dependent process is uncorrelated. *J. Amer. Statist. Assoc.*, 96(455):1066–1076. [161](#), [162](#), [163](#)
- [33] Lund, R. and Basawa, I. V. (2000). Recursive prediction and likelihood evaluation for periodic ARMA models. *J. Time Ser. Anal.*, 21(1):75–93. [150](#), [151](#)
- [34] Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer-Verlag, Berlin. [152](#)

- [35] McLeod, A. I. (1993). Parsimony, model adequacy and periodic correlation in time series forecasting. *International Statistical Review/Revue Internationale de Statistique*, pages 387–393. [150](#)
- [36] McLeod, A. I. (1994). Diagnostic checking of periodic autoregression models with application. *J. Time Ser. Anal.*, 15(2):221–233. [158](#)
- [37] McLeod, A. I. (1995). Addendum: Diagnostic checking of periodic autoregression models with application [J. Time Ser. Anal. **15** (1994), no. 2, 221–233]. *J. Time Ser. Anal.*, 16(6):647–648. [158](#)
- [38] Morgan, J. and Tatar, J. (1972). Calculation of the residual sum of squares for all possible regressions. *Technometrics*, 14(2):317–325. [150](#)
- [39] Newey, W. K. and West, K. D. (1987). A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708. [156](#), [161](#)
- [40] Noakes, D. J., McLeod, A. I., and Hipel, K. W. (1985). Forecasting monthly riverflow time series. *International Journal of Forecasting*, 1(2):179–190. [150](#)
- [41] Pagano, M. (1978). On periodic and multiple autoregressions. *Ann. Statist.*, 6(6):1310–1317. [150](#)
- [42] Reinsel, G. C. (1997). *Elements of multivariate time series analysis*. Springer Series in Statistics. Springer-Verlag, New York, second edition. [152](#)
- [43] Salas, J. D. (1980). *Applied modeling of hydrologic time series*. Water Resources Publication. [150](#)
- [44] Salas, J. D., Boes, D. C., and Smith, R. A. (1982). Estimation of arma models with seasonal parameters. *Water Resources Research*, 18(4):1006–1010. [150](#)
- [45] Shao, Q. and Lund, R. (2004). Computation and characterization of autocorrelations and partial autocorrelations in periodic ARMA models. *J. Time Ser. Anal.*, 25(3):359–372. [158](#)
- [46] Shao, X. (2010a). A self-normalized approach to confidence interval construction in time series. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 72(3):343–366. [161](#)
- [47] Shao, X. (2010b). Corrigendum: A self-normalized approach to confidence interval construction in time series. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 72(5):695–696. [161](#)
- [48] Shao, X. (2012). Parametric inference in stationary time series models with dependent errors. *Scand. J. Stat.*, 39(4):772–783. [161](#)
- [49] Shao, X. (2015). Self-normalization for time series: a review of recent developments. *J. Amer. Statist. Assoc.*, 110(512):1797–1817. [161](#)
- [50] Thompstone, R. M., Hipel, K. W., and McLeod, A. I. (1985). Grouping of periodic autoregressive models. *Time series analysis: Theory and practice*, 6:35–49. [150](#)
- [51] Ursu, E. and Duchesne, P. (2009). Estimation and model adequacy checking for multivariate seasonal autoregressive time series models with periodically varying parameters. *Stat. Neerl.*, 63(2):183–212. [150](#)



- [52] Ursu, E. and Perea, J.-C. (2016). Application of periodic autoregressive process to the modeling of the garonne river flows. *Stochastic Environmental Research and Risk Assessment*, 30(7):1785–1795. [150](#)
- [53] Vecchia, A. (1985a). Maximum likelihood estimation for periodic autoregressive moving average models. *Technometrics*, 27(4):375–384. [150](#)
- [54] Vecchia, A. (1985b). Periodic autoregressive-moving average (parma) modeling with applications to water resources 1. *JAWRA Journal of the American Water Resources Association*, 21(5):721–730. [150](#)
- [55] Vecchia, A. V. (1985c). Periodic autoregressive-moving average modeling with applications to water resources. *Journal of the American Water Resources Association*, 21(5):721–730. [152](#)

# Conclusion générale et perspectives

Dans cette thèse nous avons étudié les modèles ARMA saisonniers et/ou périodiques faibles dont les termes d'erreur sont non corrélés mais qui peuvent contenir des dépendances non linéaires, plus particulièrement la validation de ces modèles. Nous avons montré que les hypothèses habituellement imposé dans le cadre des modèles ARMA saisonniers et/ou périodiques forts sont moins restrictive dans le cas faible. Ceci permet aux modèles SARMA, PARMA et SPARMA faibles d'élargir considérablement leurs champs d'applications. Nous avons établi la consistance et la normalité asymptotique de l'estimateur du quasi-maximum de vraisemblance et de l'estimateur des moindres carrés quasi-généralisé des modèles SARMA faibles et SPARMA faibles. Nous avons estimé la matrice de variance asymptotique de l'estimateur des moindres carrés quasi-généralisés des modèles SPARMA faibles.

Nous avons traité le problème de la validation des modèles ARMA saisonniers et/ou périodiques faibles avec deux techniques différentes qui permettent de tester l'adéquation de ces modèles. Nous avons proposé des tests portmanteau modifiés qui prennent en compte les dépendances des termes d'erreurs. Nous avons montré que les autocorrélations résiduelles de ces modèles ont une distribution asymptotique normal de matrice de covariance différente de celle obtenue dans le cas standard (c'est-à-dire sous les hypothèses iid sur le bruit). Cela nous a permis d'avoir le comportement asymptotique des statistiques de tests portmanteau et de proposer ainsi des versions modifiées de ceux ci. La distribution asymptotique des tests portmanteau est approximée par un chi-deux lorsque le terme d'erreur est supposé iid. Dans le cas général, nous avons montré que cette distribution asymptotique est celle d'une somme pondérée de chi-deux, elle peut être très différente de l'approximation chi-deux usuelle du cas fort (iid). Nous avons proposé également une approche qui nous permet d'éviter l'estimation de la matrice de covariance asymptotique des autocorrélations empiriques dite d'auto-normalisation. Ainsi, nous avons proposé une nouvelle statistique de tests dont la distribution dépend uniquement des ordres autorégressifs et moyennes mobiles.

Nombreuses extensions sont possibles pour le travail que nous avons développé. Parmi ces développements, nous citons :

## **Identification des modèles ARMA saisonniers et/ou périodiques faibles**

Dans cette thèse nous avons développé l'estimation et la validation des modèles, nous n'avons pas abordé l'étape d'identification qui est l'une de trois étapes importantes de la modélisation. Cette étape est généralement la plus compliquée. Un des problèmes le plus délicat est celui de la sélection d'un petit nombre de valeurs pour les ordres des modèles. Les méthodes d'identifications les plus utilisées sont celles basées sur l'optimisation d'un critère d'information, tel que le critère d'information de Akaike (noté AIC pour Akaike's information criterion)([2]) ou

le critère d'information bayésienne (noté BIC pour Bayesian information criterion) ([85]. [18] a étudié le problème de sélection des ordres  $p_0$  et  $q_0$  de modèles VARMA faibles en considérant le critère AIC. [100] et al ont proposé une nouvelle procédure automatique au problème de sélection du modèle PAR qui utilise un algorithme génétique et le critère BIC pour identifier l'ordre du modèle PAR. Une de voie de recherche serait donc de développer cette étape pour les modèles ARMA saisonniers et/ou périodiques faibles.

### Estimations et validations des modèles PVARMA et SPVARMA faibles

[16] a déjà réalisé des travaux dans le cadre de modèles VARMA faibles (voir aussi [20], [17], [21]). D'un autre coté, nous avons [98], [96], [95] qui ont étudié l'estimation des paramètres (avec ou sans contraintes linéaires sur les paramètres) et la validations des modèles SVAR, PVAR et SPVAR avec des termes d'erreurs iid. Une autre voix de recherche serait de généraliser ces travaux en ajoutant la partie moyenne mobile (MA) et d'étudier dans quelle mesure, on peut étendre aux modèles PVARMA et SPVARMA avec des termes d'erreurs non corrélés mais non nécessairement indépendants.

## Bibliography

[BMI]

- [1] Ahn, S. K. (1988). Distribution for residual autocovariances in multivariate autoregressive models with structured parameterization. *Biometrika*, 75(3):590–593.
- [2] Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723. [x](#)
- [3] Aknouche, A. and Bibi, A. (2009). Quasi-maximum likelihood estimation of periodic garch and periodic arma-garch processes. *Journal of Time Series Analysis*, 30(1):19–46.
- [4] Akutowicz, E. J. (1957). On an explicit formula in linear least squares prediction. *Math. Scand.*, 5:261–266.
- [5] Andrews, B., Davis, R. A., and Breidt, F. J. (2006). Maximum likelihood estimation for all-pass time series models. *Journal of Multivariate Analysis*, 97(7):1638–1659.
- [6] Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3):817–858.
- [7] Basawa, I. V. and Lund, R. (2001). Large sample properties of parameter estimates for periodic ARMA models. *J. Time Ser. Anal.*, 22(6):651–663.
- [8] Basawa, I. V., Lund, R., and Shao, Q. (2004). First-order seasonal autoregressive processes with periodically varying parameters. *Statist. Probab. Lett.*, 67(4):299–306.
- [9] Battaglia, F., Cucina, D., and Rizzo, M. (2018). A generalization of periodic autoregressive models for seasonal time series. Technical report, Tech. Rep. 2, Department of Statistical Sciences, University La Sapienza . . . .
- [10] Battaglia, F., Cucina, D., and Rizzo, M. (2020). Parsimonious periodic autoregressive models for time series with evolving trend and seasonality. *Statistics and Computing*, 30(1):77–91.
- [11] Bauwens, L., Laurent, S., and Rombouts, J. V. K. (2006). Multivariate GARCH models: a survey. *J. Appl. Econometrics*, 21(1):79–109.

- [12] Berk, K. N. (1974). Consistent autoregressive spectral estimates. *Ann. Statist.*, 2:489–502. Collection of articles dedicated to Jerzy Neyman on his 80th birthday.
- [13] Bloomfield, P., Hurd, H. L., and Lund, R. B. (1994). Periodic correlation in stratospheric ozone data. *Journal of Time Series Analysis*, 15(2):127–150.
- [14] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.
- [15] Bollerslev, T. and Ghysels, E. (1996). Periodic autoregressive conditional heteroscedasticity. *Journal of Business & Economic Statistics*, 14(2):139–151.
- [16] Boubacar Mainassara, Y. (2009). *Estimation, validation et identification des modèles ARMA faibles multivariés*. PhD thesis. [xi](#)
- [17] Boubacar Mainassara, Y. (2011). Multivariate portmanteau test for structural VARMA models with uncorrelated but non-independent error terms. *J. Statist. Plann. Inference*, 141(8):2961–2975. [xi](#)
- [18] Boubacar Mainassara, Y. (2012). Selection of weak varma models by modified akaike's information criteria. *Journal of Time Series Analysis*, 33(1):121–130. [xi](#)
- [19] Boubacar Maïnassara, Y. (2014). Estimation of the variance of the quasi-maximum likelihood estimator of weak VARMA models. *Electron. J. Stat.*, 8(2):2701–2740.
- [20] Boubacar Mainassara, Y., Carbon, M., and Francq, C. (2012). Computing and estimating information matrices of weak ARMA models. *Comput. Statist. Data Anal.*, 56(2):345–361. [xi](#)
- [21] Boubacar Mainassara, Y. and Francq, C. (2011). Estimating structural VARMA models with uncorrelated but non-independent error terms. *J. Multivariate Anal.*, 102(3):496–505. [xi](#)
- [22] Boubacar Maïnassara, Y. and Ilmi Amir, A. (2019). Distribution of residual autocorrelations for multiplicative seasonal arma models with uncorrelated but nonindependent error terms. *Statistica Neerlandica*, 73(4):454–474.
- [23] Boubacar Maïnassara, Y. and Ilmi Amir, A. (2020). Multivariate portmanteau tests for weak multiplicative seasonal VARMA models. *Statist. Papers*, 61(6):2529–2560.
- [24] Boubacar Maïnassara, Y. and Kokonendji, C. C. (2016). Modified schwarz and hannan-quinn information criteria for weak varma models. *Statistical Inference for Stochastic Processes*, 19(2):199–217.
- [25] Boubacar Maïnassara, Y. and Sausseureau, B. (2017). Diagnostic checking in multivariate ARMA models with dependent errors using normalized residual autocorrelations. *J. Amer. Statist. Assoc.*, to appear.
- [26] Boubacar Maïnassara, Y. and Sausseureau, B. (2018). Diagnostic checking in multivariate arma models with dependent errors using normalized residual autocorrelations. *J. Amer. Statist. Assoc.*, 113(524):1813–1827.
- [27] Box, G. E. P. and Jenkins, G. M. (1970). *Times series analysis. Forecasting and control*. Holden-Day, San Francisco, Calif.-London-Amsterdam.
- [28] Box, G. E. P. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J. Amer. Statist. Assoc.*, 65:1509–1526.

- [29] Brockwell, P. J. and Davis, R. A. (1991a). *Time series: theory and methods*. Springer Series in Statistics. Springer-Verlag, New York, second edition.
- [30] Brockwell, P. J. and Davis, R. A. (1991b). *Time series: theory and methods*. Springer Series in Statistics. Springer-Verlag, New York, second edition.
- [31] Chitturi, R. V. (1974). Distribution of residual autocorrelations in multiple autoregressive schemes. *J. Amer. Statist. Assoc.*, 69:928–934.
- [32] Davidson, J. (1994). *Stochastic limit theory*. Advanced Texts in Econometrics. The Clarendon Press, Oxford University Press, New York. An introduction for econometricians.
- [33] Davydov, J. A. (1968). The convergence of distributions which are generated by stationary random processes. *Teor. Veroyatnost. i Primenen.*, 13:730–737.
- [34] den Haan, W. J. and Levin, A. T. (1997). A practitioner's guide to robust covariance matrix estimation. In *Robust inference*, volume 15 of *Handbook of Statist.*, pages 299–342. North-Holland, Amsterdam.
- [35] Duchesne, P. (2007). On consistent testing for serial correlation in seasonal time series models. *Canad. J. Statist.*, 35(2):193–213.
- [36] Dufour, J.-M. and Pelletier, D. (2008). Practical methods for modelling weak varma processes: Identification, estimation and specification with a macroeconomic application. *Manuscript, McGill University*.
- [37] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007.
- [38] Francq, C., Roy, R., and Saidi, A. (2011). Asymptotic properties of weighted least squares estimation in weak PARMA models. *J. Time Series Anal.*, 32(6):699–723.
- [39] Francq, C., Roy, R., and Zakoïan, J.-M. (2005). Diagnostic checking in ARMA models with uncorrelated errors. *J. Amer. Statist. Assoc.*, 100(470):532–544.
- [40] Francq, C. and Zakoïan, J.-M. (1998). Estimating linear representations of nonlinear processes. *J. Statist. Plann. Inference*, 68(1):145–165.
- [41] Francq, C. and Zakoïan, J.-M. (2000). Covariance matrix estimation for estimators of mixing weak ARMA models. *J. Statist. Plann. Inference*, 83(2):369–394.
- [42] Francq, C. and Zakoïan, J.-M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10(4):605–637.
- [43] Francq, C. and Zakoïan, J.-M. (2005). Recent results for linear time series models with non independent innovations. In *Statistical modeling and analysis for complex data problems*, volume 1 of *GERAD 25th Anniv. Ser.*, pages 241–265. Springer, New York.
- [44] Francq, C. and Zakoïan, J.-M. (2010). *GARCH Models: Structure, Statistical Inference and Financial Applications*. Wiley.
- [45] Franses, P. H. and Paap, R. (2004). *Periodic time series models*. OUP Oxford.
- [46] Gallagher, C. M. and Fisher, T. J. (2015). On weighted portmanteau tests for time-series goodness-of-fit. *Journal of Time Series Analysis*, 36(1):67–83.

- [47] Giovanis, E. et al. (2014). The turn-of-the-month-effect: Evidence from periodic generalized autoregressive conditional heteroskedasticity (pgarch) model. *International Journal of Economic Sciences and Applied Research*, 7(3):43–61.
- [48] Hannan, E. (1975). The estimation of arma models. *The Annals of Statistics*, pages 975–981.
- [49] Hannan, E. J. (1976). The identification and parametrization of ARMAX and state space forms. *Econometrica*, 44(4):713–723.
- [50] Hannan, E. J. and Rissanen, J. (1982). Recursive estimation of mixed autoregressive-moving average order. *Biometrika*, 69(1):81–94.
- [51] Herrndorf, N. (1984). A functional central limit theorem for weakly dependent sequences of random variables. *Ann. Probab.*, 12(1):141–153.
- [52] Hipel, K. and McLeod, A. I. (1994). *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier, Amsterdam.
- [53] Hipel, K. W., McLeod, A. I., and Lennox, W. C. (1977). Advances in box-jenkins modeling: 1. model construction. *Water Resources Research*, 13(3):567–575.
- [54] Holan, S. H., Lund, R., Davis, G., et al. (2010). The arma alphabet soup: A tour of arma model variants. *Statistics Surveys*, 4:232–274.
- [55] Hosking, J. (1981). Equivalent forms of the multivariate portmanteau statistic. *Journal of the Royal Statistical Society: Series B (Methodological)*, 43(2):261–262.
- [56] Hosking, J. (1989). Corrigendum: Equivalent forms of the multivariate portmanteau statistic. *Journal of the Royal Statistical Society. Series B (Methodological)*, 51(2):303–303.
- [57] Hosking, J. R. M. (1980). The multivariate portmanteau statistic. *J. Amer. Statist. Assoc.*, 75(371):602–608.
- [58] Ibragimov, I. A. (1962). Some limit theorems for stationary processes. *Teor. Veroyatnost. i Primenen.*, 7:361–392.
- [59] Imhof, J. P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika*, 48:419–426.
- [60] Jeantheau, T. (1998). Strong consistency of estimators for multivariate ARCH models. *Econometric Theory*, 14(1):70–86.
- [61] Jones, R. H. and Brelsford, W. M. (1967). Time series with periodic structure. *Biometrika*, 54(3-4):403–408.
- [62] Kuan, C.-M. and Lee, W.-M. (2006). Robust M tests without consistent estimation of the asymptotic covariance matrix. *J. Amer. Statist. Assoc.*, 101(475):1264–1275.
- [63] Li, W. and McLeod, A. (1981). Distribution of the residual autocorrelations in multivariate arma time series models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 43(2):231–239.
- [64] Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2):pp. 297–303.
- [65] Lobato, I. N. (2001). Testing that a dependent process is uncorrelated. *J. Amer. Statist. Assoc.*, 96(455):1066–1076.

- [66] Lund, R. and Basawa, I. V. (2000). Recursive prediction and likelihood evaluation for periodic ARMA models. *J. Time Ser. Anal.*, 21(1):75–93.
- [67] Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer-Verlag, Berlin.
- [68] Mahdi, E. (2016). Portmanteau test statistics for seasonal serial correlation in time series models. *SpringerPlus*, 5(1):1485.
- [69] McLeod, A. I. (1978). On the distribution of residual autocorrelations in Box-Jenkins models. *J. Roy. Statist. Soc. Ser. B*, 40(3):296–302.
- [70] McLeod, A. I. (1993). Parsimony, model adequacy and periodic correlation in time series forecasting. *International Statistical Review/Revue Internationale de Statistique*, pages 387–393.
- [71] McLeod, A. I. (1994). Diagnostic checking of periodic autoregression models with application. *J. Time Ser. Anal.*, 15(2):221–233.
- [72] McLeod, A. I. (1995). Addendum: Diagnostic checking of periodic autoregression models with application [J. Time Ser. Anal. 15 (1994), no. 2, 221–233]. *J. Time Ser. Anal.*, 16(6):647–648.
- [73] McLeod, A. I. and Li, W. K. (1983). Diagnostic checking arma time series models using squared-residual autocorrelations. *Journal of time series analysis*, 4(4):269–273.
- [74] Morgan, J. and Tatar, J. (1972). Calculation of the residual sum of squares for all possible regressions. *Technometrics*, 14(2):317–325.
- [75] Newey, W. K. and West, K. D. (1987). A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.
- [76] Noakes, D. J., McLeod, A. I., and Hipel, K. W. (1985). Forecasting monthly riverflow time series. *International Journal of Forecasting*, 1(2):179–190.
- [77] Osborn, D. R. and Smith, J. P. (1989). The performance of periodic autoregressive models in forecasting seasonal uk consumption. *Journal of Business & Economic Statistics*, 7(1):117–127.
- [78] Pagano, M. (1978). On periodic and multiple autoregressions. *Ann. Statist.*, 6(6):1310–1317.
- [79] Parzen, E. and Pagano, M. (1979). An approach to modeling seasonally stationary time series. *Journal of Econometrics*, 9(1-2):137–153.
- [80] Reinsel, G. C. (1997). *Elements of multivariate time series analysis*. Springer Series in Statistics. Springer-Verlag, New York, second edition.
- [81] Romano, J. P. and Thombs, L. A. (1996). Inference for autocorrelations under weak assumptions. *J. Amer. Statist. Assoc.*, 91(434):590–600.
- [82] Salas, J. and Obeysekera, J. (1982). Arma model identification of hydrologic time series. *Water Resources Research*, 18(4):1011–1021.
- [83] Salas, J. D. (1980). *Applied modeling of hydrologic time series*. Water Resources Publication.

- [84] Salas, J. D., Boes, D. C., and Smith, R. A. (1982). Estimation of arma models with seasonal parameters. *Water Resources Research*, 18(4):1006–1010.
- [85] Schwarz, G. et al. (1978). Estimating the dimension of a model. *Annals of statistics*, 6(2):461–464. [xi](#)
- [86] Shao, Q. and Lund, R. (2004). Computation and characterization of autocorrelations and partial autocorrelations in periodic ARMA models. *J. Time Ser. Anal.*, 25(3):359–372.
- [87] Shao, X. (2010a). A self-normalized approach to confidence interval construction in time series. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 72(3):343–366.
- [88] Shao, X. (2010b). Corrigendum: A self-normalized approach to confidence interval construction in time series. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 72(5):695–696.
- [89] Shao, X. (2012). Parametric inference in stationary time series models with dependent errors. *Scand. J. Stat.*, 39(4):772–783.
- [90] Shao, X. (2015). Self-normalization for time series: a review of recent developments. *J. Amer. Statist. Assoc.*, 110(512):1797–1817.
- [91] Slutsky, E. (1937). The summation of random causes as the source of cyclic processes. *Econometrica: Journal of the Econometric Society*, pages 105–146.
- [92] Thompstone, R. M., Hipel, K. W., and McLeod, A. I. (1985). Grouping of periodic autoregressive models. *Time series analysis: Theory and practice*, 6:35–49.
- [93] Tiao, G. and Grupe, M. (1980). Hidden periodic autoregressive-moving average models in time series data. *Biometrika*, 67(2):365–373.
- [94] Tiao, G. C. and Tsay, R. S. (1983). Consistency properties of least squares estimates of autoregressive parameters in arma models. *The Annals of Statistics*, pages 856–871.
- [95] Ursu, E. and Duchesne, P. (2009a). Estimation and model adequacy checking for multivariate seasonal autoregressive time series models with periodically varying parameters. *Stat. Neerl.*, 63(2):183–212. [xi](#)
- [96] Ursu, E. and Duchesne, P. (2009b). On modelling and diagnostic checking of vector periodic autoregressive time series models. *J. Time Series Anal.*, 30(1):70–96. [xi](#)
- [97] Ursu, E. and Duchesne, P. (2009c). On multiplicative seasonal modelling for vector time series. *Statist. Probab. Lett.*, 79(19):2045–2052.
- [98] Ursu, E. and Duchesne, P. (2009d). On multiplicative seasonal modelling for vector time series. *Statistics & probability letters*, 79(19):2045–2052. [xi](#)
- [99] Ursu, E. and Pereau, J.-C. (2016). Application of periodic autoregressive process to the modeling of the garonne river flows. *Stochastic Environmental Research and Risk Assessment*, 30(7):1785–1795.
- [100] Ursu, E. and Turkman, K. F. (2012). Periodic autoregressive model identification using genetic algorithms. *Journal of Time Series Analysis*, 33(3):398–405. [xi](#)
- [101] Vecchia, A. (1985a). Maximum likelihood estimation for periodic autoregressive moving average models. *Technometrics*, 27(4):375–384.



- [102] Vecchia, A. (1985b). Periodic autoregressive-moving average (parma) modeling with applications to water resources 1. *JAWRA Journal of the American Water Resources Association*, 21(5):721–730.
- [103] Wold, H. (1938). *A study in the analysis of stationary time series*. PhD thesis, Almqvist & Wiksell.
- [104] Yule, G. U. (1927). Vii. on a method of investigating periodicities disturbed series, with special reference to wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 226(636-646):267–298.
- [105] Zhu, K. (2013). A mixed portmanteau test for ARMA-GARCH models by the quasi-maximum exponential likelihood estimation approach. *J. Time Series Anal.*, 34(2):230–237.