UNIVERSITÉ MENTOURI

CONSTANTINE

DÉPARTEMENT DE MATHÉMATIQUES

FACULTÉ DES SCIENCES EXACTES

Inférence Statistique dans les Processus GARCH à Coefficients Dépendant du Temps

Ines LESCHEB

Docteur en Mathématiques

Avril 2011

Supervisor: Prof. Abdelouahab BIBI

Assessor:

Contents

Introduction

L'analyse des modèles de séries chronologiques exhibant des changements structurales remonte aux années cinquante par Rubin [43]. Excellente introduction (et une bibliographie abondante) sur le sujet a été donnée récemment par Hallin [27]. Dans cette classe de modèles, on peut distinguer deux catégories de modèles à coefficients variables, selon que cette évolution est de nature déterministe ou non. Ce sont les modèles à coefficients stochastiques (cf. Nicolls et Quinn [40]) et les modèles à coefficients dépendant du temps introduits par Cramér [19]. Les modèles de la première catégorie visent cependant tout comme les modèles à coefficients constants, à décrire principalement des processus de nature stationnaire. Par contre, les modèles de la seconde catégorie, sont introduits afin de modéliser des séries non stationnaires lorsque les méthodes de filtrages et de différentiations ne permettent cependant pas. Ce sont les modèles à coefficients (presque-) périodiques qui, assez curieusement ont reçu jusqu'à présent le plus d'attention. D'excellents et récents travaux de synthèse sont disponibles sur ce sujet (citons, notamment Bezandry and Diagana [7], Hurd et Miamee [28] et Franses et Paap [25]) auxquelles nous renvoyons le lecteure intéressé. Les applications de ces modèles sont multiples et nous les retrouvons, en science économique (Cleveland et Tio [18], Franses [24], Franses et Paap [25], Parzen et Pagano [42]), en climatologie (Bloomfield et al. [14]), à l'ingénierie électrique (Meyer et Burrus [37], Bittanti et De Nicolao [13], Adams et Goodwin [1], Gardner et Franks [26]) et à l'hydrologie (Vecchia [48]).

En économétrie et en finance emirique, une classe de modèles désormais assez populaires, ce sont les modèles $GARCH$ (Autorégressifs Conditionnellement Hétéroskédastiques Généralisés) périodiques ($PGARCH$) introduits pour la première fois par Bollerslev et Ghysels [15], puis popularisée à travers les travaux de Aknouche et Bibi [2], Bibi et Aknouche [8]. Ces modèles sont généralement non stationnaires mais ils sont stationnaires entre chaque période. Ils sont devenus un outil puissant et fondamental pour modéliser des séries financière à volatilité saisonnière. La structure des modèles $PGARCH$ est semblable à celle des modèles linéaires périodiques, ils partagent donc beaucoup de similarités avec les modèles périodiques linéaires mais ont aussi, à cause des non linéarités, des caractéristiques spécifiques que nous les étudions à travers les différents chapitres de cette thèse.

Historiquement, les motivations majeures qui se trouvent à la base d'introduction des modèles $PGARCH$ sont d'origines empiriques. En effet, l'observation d'une structure saisonnière non-constante des autocorrélations des rendements boursiers nécessite le recours à une classe de modèles plus riches que les modèles linéaires standards des séries temporelles qui supposent une constance de la structure d'autocorrélation. Ce dernier point a amené certains auteurs tels Bessembinder et Hertzel [6] à utiliser des modèles de séries temporelles périodiques qui admettent explicitement une structure d'autocorrélation qui peut varier au travers de la semaine. Cette classe de modèles périodiques a été largement étudiée tant d'un point de vue théorique qu'empirique comme en témoigne les récents livres de Franses [24] et de Hurd et Miamee [28]. Elle couvre une multitude de modèles univariés ou multivariés qui s'avèrent fort utiles pour modéliser des séries économiques saisonnières. Leur utilisation en finance empirique reste néanmoins relativement peu courante en comparaison aux simple modèles linéaires de régression. Citons les travaux de Bessembinder et Hertzel [6] qui utilisent des modèles autorégressifs périodiques (PAR) pour l'analyse de la structure d'autocorrélation des rendements aux alentours des jours ouvrables alors que dans une contribution assez importante Bollerslev et Ghysels [15] appliquent un raisonnement similaire à la modélisation de la dynamique de la volatilité des séries financières. Pour ce faire, ils proposent un modèle $PGARCH$ qu'ils s'appliquent avec succès à des séries de taux de change ainsi qu'à certains indices boursiers. L'avantage évident de cette approche est qu'elle permet d'une représentation assez flexible des effets saisonniers, et des périodicités diverses sur la volatilité des séries financières. Franses et Paap [25] quant à eux, unifient ces deux types d'études en proposant une modélisation économétrique des rendements financières intégrant à la fois périodicité observée en moyenne avec celle observée en volatilité: Le modèle PAR-$PGARCH$. Leurs résultats, ainsi que ceux obtenus par Bessembinder et Hertzel [6] mettent assez clairement en évidence non seulement une structure périodique dans l'autocorrélation des rendements, mais aussi des effets saisonniers dans la persistance de la volatilité. Malgré le nombre important des paramètres qui apparaissent dans un modèle $PGARCH$, et par conséquence leurs estimations en l'absence de la stationnarité et de l'ergodicité, les modèles $PGARCH$ ont gagnés un intérêt considérable et continu à attirés l'attention des chercheurs, cependant une grande littérature a été observée témoignant l'intérêt particulier de cette classe de modèles (cf. [44]).

Notons ici que dans la classe des modèles $GARCH$ stationnaires, nous trouvons ainsi une littérature abondante. Cette abondance est due aux conditions sous lesquelles le modèle devient ergodique. Cependant de nombreux travaux de recherche ont développés les propriétés probabilistes et statistiques notamment: l'identification, les tests et l'estimation des paramètres (Pour une bibliographie récente, riche et exhaustive, voir Francq et Zakoïan [21]). En revanche, dans la classe des modèles $GARCH$ à coefficients dépendants du temps, les méthodes classiques d'estimation ne s'appliquent pas directement. Car, par exemple, les conditions de régularité sous lesquelles l'estimateur du quasi-maximum de vraisemblance est convergent et efficient ont été dérivées pour les modèles $(I)GARCH$. A notre modeste connaissance, aucun résultat théorique n'existe sur l'estimation pour des modèles $PGARCH$ autre que cel de Aknouche et Bibi [2] (Ce papier est cité jusqu'a présent plus de 10 fois).

Certes, l'étude des modèles $GARCH$ périodiques est loin d'être achevée. Cependant de nombreux problèmes de nature statistiques restent ouverts. Néanmoins, on peut se demander s'il est possible de résoudre, par exemple, le problème de l'identification des modèles $PGARCH$ au sens de réduire le nombre des paramètres incorporés dans le modèle comme ce fut pour les modèles $GARCH$ stationnaires dans la mesure où la classe de modèles considérés est très riche et assez complexe. La théorie des tests qui est jusqu'à présent a été peu étudiée (dans le cas stationnaire) doit permettre d'aboutir assez rapidement à quelques résultats: Outre les tests de stationnarité (cf. Francq et Zakoïan [22]) pour lesquels quelques procédures ont été proposées, on a besoin de tests portant sur le choix de l'évolution des coefficients (périodique ou presque périodique), autrement dit le choix de modèle. Ainsi, le but de notre travail est de contribuer à l'étude des modèles $PGARCH$ à travers l'estimation et quelques tests de périodicité. Cette thèse que nous présentons permet de faire le point sur l'état actuel des recherches concernant les modèles $PGARCH$ ainsi que sur quelques points non encore traités et indispensable pour mieux comprendre ces modèles. Outre les résultats de l'auteur, dont les articles correspondants se trouvent vers la fin de la thèse, on trouve aussi d'autres résultats présentés sans preuves. Ceux-ci pourront être consultés à travers les références citées dans la bibliographie générale.

0.1 Apport et présentation de la thèse

Notre thèse intitulée "Inférence Statistique dans les Processus $GARCH$ à Coefficients Dépendant du Temps " se compose en cinq chapitres principaux:

Chapitre 1 : On the structures of $PGARCH$ models

Ce chapitre présente la structure de $\mathbb{L}_2$ et les propriétés propabilistes. En basant sur une représentation vectorielle appropriée, nous donnons des conditions nécessaires et suffisantes assurant l'existence et l'unicité de solutions stationnaires (au sens périodique) et l'existence de moments d'ordre supérieurs.

Chapitre 2 : The LSE approach for $PGARCH$ models

Ce chapitre traite les propriétés asymptotiques de l'estimateur (LSE) (non standard) pour les $PGARCH$ et les $PARMA-PGARCH$ modèles. Premièrement, nous donnons des conditions nécessaires et suffisantes qui assurent l'existence de solutions stationnaires (au sens périodique) et pour l'existence de moments d'ordre supérieurs. Deuxièmement, une approche basée sur des moindres carrés (non standard) pour estimer les modèles $PGARCH$ et les modèles $PARMA$-$PGARCH$ modèles est présentée. La consistance forte et la normalité asymptotique des estimateurs sont établées.

Chapitre 3 : The CLS approach for $PGARCH$ models

Ce chapitre étudie la consistance forte et la normalité asymptotique de l'estimateur des moindres carrés conditionnels (CLS) dans les modèles $GARCH$ périodiques dont le carré centré des innovations est une différence de martingale. Cette approche est étendue aux modèles $PARMA-PGARCH$. Les résultats sont obtenus sans aucune contrainte sur les moments des innovations. Nos preuves ont étés adaptées à celles de Francq et Zakoîan [20] pour des innovations $i.i.d.$

Chapitre 4 : Yule-Walker equations for $GARCH(1,1)$ models

Ce chapitre étudie l'inférence asymptotique des modèles $PGARCH(1,1)$. Tout d'abord, nous établissons des conditions nécessaires et suffisantes pour l'existence et l'unicité de solutions stationnaires (au sens périodique) et pour l'existence de moments de tout ordre. Deuxièmement, en utilisant la représentation $PARMA(1,1)$ basée sur le carré de $PGARCH(1,1)$, nous considérons alors des estimateurs des paramètres de type Yule-Walker, nous dérivons ensuite leurs propriétés asymptotiques. Comme une application, on construit la statistique de Wald pour tester une hypothèse nulle contre une alternative. Nous utilisons un bootstrap basé sur les résidus afin de construire des estimateurs bootstrapés pour les estimations de Yule-Walker et de prouver la robustesse de cette méthode. Un ensemble d'expériences numériques illustre l'importance pratique de nos résultats théoriques.

Chapitre 5 : The LAN properties for $PARCH$ processes

Dans ce chapitre, nous considérons l'estimateur des moindres carrés conditionnels CLS pour les modèles $ARCH$ périodiques ($PARCH$). L'estimateur CLS appliqué sur le carré d'un $PARCH$ a une forme explicite indépendante de la distribution des innovations. Comme l'estimateur CLS n'est pas asymptotiquement efficace en général, nous donnons des conditions nécessaires et suffisantes assurant son efficacité asymptotique basées sur l'approche LAN.

Nous terminons notre thèse par un chapitre additif comportant une conclusion générale, des remarques, quelques perspectives et nos occupations futures.

Part I

Étude Probabiliste

Chapter 1

On the structure of $PGARCH$ models

Abstract: This chapter analyzes the $\mathbb{L}_2$ structures and the asymptotic properties of parameter least squares estimates (LSE) for periodic $GARCH$ ($PGARCH$) models. In this class of models, the parameters are allowed to switch between different regimes. Firstly, we give necessary and sufficient conditions ensuring the existence of stationary solutions (in periodic sense) and for the existence of moments of any order. Secondary, a least squares estimation approach for estimating $PGARCH$ model is developed. The strong consistency and the asymptotic normality of the estimator are studied given mild regularity conditions, requiring strict stationarity and the finiteness of moments of some order for the errors term.

1.1 $PGARCH$ models and its probabilistic properties

A discrete-time stochastic process $(\epsilon_n)_{n\in\mathbb{Z}}$ defined on some probability space $(\Omega, \mathcal{A}, P)$ with finite second order moments is said to have a periodic generalized autoregressive conditional heteroscedastic representation with period $s > 0$ and orders p and q [denoted by $PGARCH(p, q)$] if it satisfies the non-linear equations

$$\forall n \in \mathbb{Z}\text{: } \epsilon_n = e_n\sqrt{h_n} \text{ and } h_n = a_0(s_n) + \sum_{i=1}^{q} a_i(s_n)\epsilon_{n-i}^2 + \sum_{j=1}^{p} b_j(s_n)h_{n-j} \qquad (1.1)$$

where $(e_n)_{n\in\mathbb{Z}}$ is a sequence of independent identically distributed ($i.i.d.$) random variables defined on the same probability space $(\Omega, \mathcal{A}, P)$ with $E\{e_n\} = 0$ and $E\{e_n^2\} = 1$ where $s_n := \sum_{k=1}^{s} k\mathbb{I}_{\Delta(k)}(n)$ is the stage of the period cycle at time n with $\Delta(k) := \{sn + k, n \in \mathbb{Z}\}$ so, by setting $n = st + v$, Model (1.1) may be equivalently written as

$$\epsilon_{st+v} = e_{st+v}\sqrt{h_{st+v}} \text{ and } h_{st+v} = a_0(v) + \sum_{i=1}^{q} a_i(v)\epsilon_{st+v-i}^2 + \sum_{j=1}^{p} b_j(v)h_{st+v-j} \tag{1.2}$$

which we will make heavy use of (1.2). In the difference Equations (1.2), ϵ_{st+v} (respectively h_{st+v}, e_{st+v}) refers to ϵ_t (respectively h_t, e_t) during the $v - th$ "season" $1 \leq v \leq s$ of cycle t, $(a_i(v), 0 \leq i \leq q)$ and $(b_i(v), 1 \leq i \leq p)$ are the model coefficients at season $v \in \{1, ..., s\}$ such that $a_0(v) > 0$, $a_i(v) \geq 0, b_j(v) \geq 0$ for all $v \in \{1, ..., s\}$, $i \in \{1, ..., q\}$ and $j \in \{1, ..., p\}$. In what follows, we assume that e_k is independent of ϵ_t for $k > t$ and we shall continue to use the non periodic notations (ϵ_t), (e_t) and (h_t) in preference to (ϵ_{st+v}), (e_{st+v}) and (h_{st+v}) whenever the seasonality is not paramount.

Since the seminal paper by Pagano [41], with periodic coefficients, it is possible to embed regimes into a multivariate process. More precisely $\underline{\epsilon}_t = (\epsilon_{st+1}, ..., \epsilon_{st+s})'$ is a weak $s-$variate $GARCH$ model in the sense that

$$\underline{\epsilon}_t = \{diag\underline{h}_t\}^{\frac{1}{2}} \underline{e}_t \text{ and } \underline{h}_t = \underline{a}_0 + \sum_{i=0}^{q^*} A_i\underline{\epsilon}_{t-i}^2 + \sum_{j=0}^{p^*} B_j\underline{h}_{t-j} \tag{1.3}$$

where $\underline{\epsilon}_t^2 = \left(\epsilon_{st+1}^2, ..., \epsilon_{st+s}^2\right)'$, $\underline{h}_t = (h_{st+1}, ..., h_{st+s})'$ and where $\underline{e}_t = (e_{st+1}, ..., e_{st+s})'$. The model orders in (1.3) are $p^* = \left[\frac{p}{s}\right]$ and $q^* = \left[\frac{q}{s}\right]$ where $[x]$ denotes the smallest integer greater than or equal to x. The $s \times s$ matrices $(A_i)_{0\leq i\leq q^*}$ and $(B_i)_{0\leq i\leq p^*}$ are computed as follows (see Basawa and Lund [4]). A_0 and B_0 have $(i, j)\,th$ entries

$$(B_0)_{i,j} = \begin{cases} b_{i-j}(i) & \text{if } i > j \\ 0 & \text{otherwise} \end{cases} \qquad (A_0)_{i,j} = \begin{cases} a_{i-j}(i) & \text{if } i > j \\ 0 & \text{otherwise} \end{cases}$$

$(B_m)_{i,j} = b_{ms+i-j}(i)$ for $1 \leq m \leq p^*$ and $(A_m)_{i,j} = a_{ms+i-j}(i)$ for $1 \leq m \leq q^*$ and the intercept vector $\underline{a}_0 = (a_0(1), ..., a_0(s))'$. In view of (1.3), it is obvious that the $PGARCH$ process is SPS if the process $(\underline{h}_t)_{t\in\mathbb{Z}}$ is strictly stationary. So if we want to study the probabilistic properties and the higher order moments of a $PGARCH$ process it is enough to do so for the process $(\underline{h}_t)_{t\in\mathbb{Z}}$. For this purpose, we have to introduce further notations to obtain similar results for the standard $GARCH$ processes. Let

$$\begin{aligned}
\underline{a}_0(t) &= \left(I_{(s)} - A_0 diag\left\{\underline{e}_t^2\right\} - B_0\right)^{-1} \underline{a}_0, \\
A_i(t) &= \left(I_{(s)} - A_0 diag\left\{\underline{e}_t^2\right\} - B_0\right)^{-1} A_i, i = 2, ..., q^*, \\
B_1(t) &= \left(I_{(s)} - A_0 diag\left\{\underline{e}_t^2\right\} - B_0\right)^{-1} \left(A_1 diag\left\{\underline{e}_{t-1}^2\right\} + B_1\right), \\
B_i(t) &= \left(I_{(s)} - A_0 diag\left\{\underline{e}_t^2\right\} - B_0\right)^{-1} B_i,\ i = 2, ..., p^*
\end{aligned}$$

Clearly the matrix $I_{(s)} - A_0 diag\left\{\underline{e}_t^2\right\} - B_0$ is invertible and

$$\left(I_{(s)} - A_0 diag\left\{\underline{e}_t^2\right\} - B_0\right)^{-1} \geq O_{(s)}.$$

With this notation, Equation (1.3) is equivalent to $\underline{\epsilon}_t = \left\{diag \underline{h}_t\right\}^{\frac{1}{2}} \underline{e}_t$ and $\underline{h}_t = \underline{a}_0(t) + \sum_{i=2}^{q^*} A_i(t) \underline{\epsilon}_{t-i}^2 + \sum_{i=1}^{p^*} B_i(t) \underline{h}_{t-i}$. Now, set $r^* = p^* + q^* - 1$ and define the $r^* \times 1$ bloc vectors $\underline{Y}_t = \left(\underline{h}_t', ..., \underline{h}_{t-p^*+1}', \underline{\epsilon}_{t-1}^{2\prime}, ..., \underline{\epsilon}_{t-q^*+1}^{2\prime}\right)'$,
$\underline{\omega}_t = (\underline{a}_0'(t), \underline{O}_{(s)}', \ldots, \underline{O}_{(s)}', \underline{O}_{(s)}', \ldots, \underline{O}_{(s)}')'$ and $r^* \times r^*$ bloc matrix

$$M_t := M_0(\underline{e}_t) M_1(\underline{e}_{t-1})$$

where

$$\left(M_0(\underline{e}_t)\right)_{i,j} = \begin{cases} \left(I_{(s)} - A_0 diag\left\{\underline{e}_t^2\right\} - B_0\right)^{-1} & \text{if } i = j = 1 \\ I_{(s)} & \text{if } 1 < i = j \leq r^* \\ O_{(s)} & \text{otherwise} \end{cases}$$

and where

$$M_1(\underline{e}_t) = \\
\begin{pmatrix}
A_1 diag\left\{\underline{e}_t^2\right\} + B_1 & B_2 & \dots & \dots & B_{p^*} & A_2 & \dots & \dots & \dots & A_{q^*} \\
I_{(s)} & O_{(s)} & \dots & \dots & O_{(s)} & O_{(s)} & \dots & \dots & \dots & O_{(s)} \\
O_{(s)} & \ddots & \ddots & \dots & \vdots & \vdots & \dots & \dots & \dots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \dots & \dots & \dots & \vdots \\
O_{(s)} & \dots & O_{(s)} & I_{(s)} & O_{(s)} & O_{(s)} & \dots & \dots & \dots & O_{(s)} \\
diag\left\{\underline{e}_t^2\right\} & O_{(s)} & \dots & \dots & O_{(s)} & O_{(s)} & \dots & \dots & \dots & O_{(s)} \\
O_{(s)} & \dots & \dots & \dots & O_{(s)} & I_{(s)} & O_{(s)} & \dots & \dots & O_{(s)} \\
\vdots & \vdots & \vdots & \vdots & \vdots & O_{(s)} & \ddots & \ddots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & O_{(s)} & O_{(s)} \\
O_{(s)} & \dots & \dots & \dots & O_{(s)} & O_{(s)} & \dots & O_{(s)} & I_{(s)} & O_{(s)}
\end{pmatrix}.$$

Then Equation (1.3) has a stationary and ergodic solution, if and only if

$$\underline{Y}_t = M_t \underline{Y}_{t-1} + \underline{\omega}_t \tag{1.4}$$

has one. Indeed, any stationary solution of (1.3) leads via $\underline{Y}_t$ to one of (1.4) and vice versa, that the first $s-$components of a stationary solution of (1.4) are one for (1.3). Moreover, an ergodic solution of (1.4) gives also an ergodic solution of (1.3) and vice versa. In the next subsections we shall examine conditions based on (1.4) ensuring the existence of SPS solutions for Equation (1.2).

1.1.1 Strict periodic stationarity

Let $\|.\|$ denote any operator norm on the sets of $sr^* \times sr^*$ and $sr^* \times 1$ matrices and let $\log^+ x = \max\{\log x, 0\}$ for $x > 0$. Since $(\underline{e}_t)_{t\in\mathbb{Z}}$ is an $i.i.d$ process, $(M_t, \underline{\omega}_t)_{t\in\mathbb{Z}}$ is a strictly stationary ergodic sequence, so the Equation (1.4) is the same as defining the equation for a RCA model, accept that the random matrix M_t is not independent of $\underline{\omega}_t$ as is required in this model. Moreover we have $E\left\{\log^+ \|\underline{\omega}_1\|\right\} \leq E\left\{\|\underline{\omega}_1\|\right\} < +\infty$ and $E\left\{\log^+ \|M_1\|\right\} \leq E\left\{\|M_1\|\right\} < +\infty$. Therefore, from Bougerol and Picard [16], Equation (1.3) have an unique strictly stationary solution if and only if the Lyapunov exponent

$$\gamma_L(M) := \inf_{t>0} \frac{1}{t} E\left\{\log \left\| \prod_{i=0}^{t-1} M_{t-i} \right\| \right\}$$

associated with the random sequence $M := (M_t)_{t\in\mathbb{Z}}$ is strictly negative. Moreover the unique solution process $(\underline{Y}_t)_{t\in\mathbb{Z}}$ of (1.4) is ergodic, causal and given by

$$\underline{Y}_t = \sum_{k=1}^{\infty} \left\{ \prod_{i=0}^{k-1} M_{t-i} \right\} \underline{\omega}_{t-k} + \underline{\omega}_t \tag{1.5}$$

where the Series (1.5) converges $a.s.$

Example 1 *For the $PGARCH\,(1,1)$ model, after some tedious algebra we find that the necessary and sufficient condition ensuring the existence of SPS solution is that $\sum\limits_{v=1}^{s} E\left\{\log\left(a_1(v)e_0^2 + b_1(v)\right)\right\} < 0$. It is worth noting that the existence of regimes which satisfy $E\left\{\log\left(a_1(v)e_0^2 + b_1(v)\right)\right\} > 0$ does not preclude strict periodic stationarity.*

Remark 2 *Similarly to the classical results on the GARCH processes theory (see for instance Berkes et al.[5]), if $\gamma_L(M) < 0$ then there exists $\delta > 0$ such that $E\left\{h_t^{\delta}\right\} < +\infty$ and $E\left\{\epsilon_t^{2\delta}\right\} < +\infty$.*

Remark 3 *Due to the positivity of the entries of M_t, it is no difficult to show that if $\gamma_L(M) < 0$, then $\det(I_{(s)} - \sum_{j=0}^{p^*} B_j z^j) \neq 0$ for all $z \in \mathbb{C} : |z| \leq 1$. This implies that we can relate $(h_t)_{t\in\mathbb{Z}}$ and $(\epsilon_t)_{t\in\mathbb{Z}}$ through the infinite series $h_{st+v} = \alpha_0(v) + \sum_{j=1}^{\infty} \alpha_j(v)\epsilon^2_{st+v-j} = \left(\mathcal{B}^{(v)}(L)\right)^{-1} a_0(v) + \left(\mathcal{B}^{(v)}(L)\right)^{-1} \mathcal{A}^{(v)}(L)\epsilon^2_{st+v}$ for all $v \in \{1, ..., s\}$ where $\mathcal{A}^{(v)}(L) = \sum_{j=1}^{q} a_j(v)L^j$, $\mathcal{B}^{(v)}(L) = 1 - \sum_{j=1}^{p} b_j(v)L^j$, L is the back-shift operator and where the "seasonal weights" $\alpha_j(v)$ satisfy $\max_{1\leq v\leq s} \sum_{j=0}^{\infty} \alpha_j(v) < +\infty$.*

The Lyapunov exponent $\gamma_L(M)$ criterion seems difficult to obtain explicitly when $r = p + q > 1$, however a potential method to verify whether or not $\gamma_L(M) < 0$ is via Monte-Carlo simulations using Equation (1.4). This fact heavily limits the interest of the criterion in statistical applications. Indeed, the solution need to have some moments to make an estimation theory possible and Lyapunov exponent criterion does not guarantee the existence of such moments. Therefore, we have to search for conditions ensuring the existence of moments for the stationary solution for which, the top-Lyapunov exponent $\gamma_L(M)$ will be automatically negative.

1.1.2 Second order periodic stationarity

In the previous subsection, necessary and sufficient conditions ensuring the existence of a SPS solution for Equation (1.2) have been established. In this subsection we give conditions ensuring the existence of a first order stationary process $(\underline{\epsilon}_t^2, \underline{h}_t)_{t\in\mathbb{Z}}$ satisfying (1.3). Therefore, the corresponding solution process $(\epsilon_t)_{t\in\mathbb{Z}}$ has a periodic covariance structure in the sense that $Cov(\epsilon_{l+s}, \epsilon_{k+s}) = Cov(\epsilon_l, \epsilon_k)$ for all integers l and k. Such series are also called periodically correlated (PC) processes.

Theorem 4 *The $s-$variate weak $GARCH$ process (1.3) is a PC process if and only if*

$$\det\left(I_{(s)} - \sum_{j=0}^{r^*} (A_j + B_j) z^j\right) \neq 0 \text{ for all complex } z \text{ such that } |z| \leq 1. \quad (1.6)$$

Moreover, the solution process is unique, strictly stationary, ergodic, causal and given by the first $s-$block component of $(\underline{Y}_t)_{t\in\mathbb{Z}}$ defined by (1.5).

Proof. The condition is obviously necessary using (1.3). To show that (4.7) is also sufficient, we define the following $\mathbb{R}^{sr^*}$-valued processes $(\underline{S}_n(t), \underline{\Delta}_n(t))_{(t,n)\in\mathbb{Z}\times\mathbb{Z}}$

$$\underline{S}_n(t) := \begin{cases} \underline{O}_{(sr^*)} & \text{if } n < 0 \\ \underline{\omega}_t + M_t \underline{S}_{n-1}(t-1), & \text{if } n \geq 0 \end{cases}$$

and $\underline{\Delta}_n(t) := \underline{S}_n(t) - \underline{S}_{n-1}(t)$. It is easily seen that for all $n \geq 0$, $\underline{S}_n(t)$ and $\underline{\Delta}_n(t)$ are measurable functions of $\underline{e}_t, \underline{e}_{t-1}, ..., \underline{e}_{t-n}$. Hence, for any fixed $n \geq 0$ the processes $(\underline{S}_n(t))_{t\in\mathbb{Z}}$ and $(\underline{\Delta}_n(t))_{t\in\mathbb{Z}}$ are strictly stationary and ergodic. From the definition of $\underline{S}_n(t)$ and $\underline{\Delta}_n(t)$ we have

$$\underline{\Delta}_n(t) := \begin{cases} \underline{O}_{(sr^*)} & \text{if } n < 0 \\ \underline{\omega}_t & \text{if } n = 0 \\ M_t \underline{\Delta}_{n-1}(t-1), & \text{if } n > 0 \end{cases}$$

and thus for any $n \geq 1$, we have $\underline{\Delta}_n(t) = \left\{ \prod\limits_{i=0}^{n-1} M_{t-i} \right\} \underline{\omega}_{t-n}$ and $\underline{S}_n(t) = \sum\limits_{k=0}^{n} \underline{\Delta}_k(t)$. Throughout, we consider the matrix norm defined by $\|A\| = \sum\limits_{i,j} A_{i,j}$ where $A_{i,j}$ denotes the generic element of A. Since $\left\{ \prod\limits_{i=0}^{n-1} M_{t-i} \right\} \underline{\omega}_{t-n}$ has positive elements, we have for $n > 0$

$$\begin{aligned} E\left\|\underline{\Delta}_n(t)\right\| &= \left\| E\left\{M_0(\underline{e}_t)\right\} E\left\{ \prod_{i=1}^{n-1} M_1(\underline{e}_{t-i}) M_0(\underline{e}_{t-i}) \right\} E\left\{M_1(\underline{e}_{t-n})\underline{\omega}_{t-n}\right\} \right\| \\ &\leq K\left\|M^{n-1}\right\| \end{aligned}$$

where $M := E\{M_1(\underline{e}_0)M_0(\underline{e}_0)\} = E\{M_1(\underline{e}_0)\}E\{M_0(\underline{e}_0)\}$ and K is some positive constant. Thus

$$\begin{aligned} \limsup_{n\to\infty} E\left\|\underline{\Delta}_n(t)\right\| &\leq K \limsup_{n\to\infty} \left(\left(\rho\left(E\left\{M_0(\underline{e}_0)M_1(\underline{e}_0)\right\}\right)\right)^n \right) \\ &= K \limsup_{n\to\infty} \left(\left(\rho\left(M\right)\right)^n \right). \end{aligned}$$

Since $\rho(M) < 1$ if and only if the Condition (1.6) holds, then $\limsup\limits_{n\to\infty} E\left\|\underline{\Delta}_n(t)\right\| = 0$, and hence $\underline{S}_n(t)$ converges in $\mathbb{L}_1$ to some limit say $\lim\limits_{n\to\infty} \underline{S}_n(t)$ which satisfies the Equations (1.4) and (1.5). The rest of the assertions are immediate. ■

Corollary 5 *Under The Condition* (1.6)*, Equation* (1.2) *has an unique PC solution such that* $E\{\epsilon_{st+v}\} = 0$ *and* $Cov(\epsilon_{st+v}, \epsilon_{st+v'}) = (diag\{\underline{\Gamma}\})_{v,v'} \delta_{\{v=v'\}}$ *where* $\underline{\Gamma} := E\{\underline{h}_t\} = \left(I_{(s)} - \sum\limits_{i=0}^{r^*} (A_i + B_i) \right)^{-1} \underline{a}_0$ *and hence the process* $(\epsilon_t)_{t\in\mathbb{Z}}$ *may be viewed as a weak white noise.*

Remark 6 *From the theorem of Kesten and Spitzer [30], it follows that the Condition* (1.6) *implies that* $\gamma_L(M) < 0$ *and thus the result given in Remark 2 holds.*

1.1.3 The existence of higher-order moments

In this subsection, we derive necessary and sufficient conditions for the finiteness of $E\{\epsilon_t^{2m}\}$, for any integer $m > 1$. By the $\mathbb{L}_m-$theory, $m \geq 1$ the problem of existence of $E\{\epsilon_t^{2m}\}$ now reduces to the convergence of $(\underline{S}_n(t))_{n\geq 0}$ in $\mathbb{L}_m$ for all $t \in \mathbb{Z}$. As it is shown in Theorem (4) $(\underline{S}_n(t))_{n\geq 0}$ converges to $\underline{Y}_t$ in $\mathbb{L}_1$. The key quantity of interest in determining $\mathbb{L}_m$ convergence is $\underline{V}_n := E\left\{\underline{\Delta}_n^{\otimes m}(t)\right\}$. Let $M^{(m)} := E\left\{M_1^{\otimes m}(\underline{e}_0)M_0^{\otimes m}(\underline{e}_0)\right\} = E\left\{M_1^{\otimes m}(\underline{e}_0)\right\}E\left\{M_0^{\otimes m}(\underline{e}_0)\right\}$. From (1.4) there exists a constant $K > 0$ such that

$$\begin{aligned}\|\underline{Y}_t\|_m &= E\{\|\underline{Y}_t\|^m\}^{\frac{1}{m}} \leq \sum_{n\geq 0}\|\underline{\Delta}_n(t)\|_m \leq K\sum_{n\geq 0}\left\|\left\{\prod_{i=1}^{n-1}M_1(\underline{e}_{t-i})M_0(\underline{e}_{t-i})\right\}\right\|_m \\ &\leq K\left\{\sum_{n\geq 0}\left\|\left(M^{(m)}\right)^n\right\|^{\frac{1}{m}}\right\}.\end{aligned}$$

Hence, if $\rho\left(M^{(m)}\right) < 1$, then $\left\|\left(M^{(m)}\right)^n\right\|$ converges to 0 with exponential rate as $n \to \infty$. Since $\left\|\epsilon_{st+v}^2\right\|_m \leq \|\underline{\epsilon}_t^2\|_m \leq \|\underline{Y}_t\|_m$ for all $v \in \{1, ..., s\}$ thus a sufficient condition for the finiteness of $E\left\{\epsilon_{st+v}^{2m}\right\}$ is that $\rho\left(E\left\{M_t^{(m)}\right\}\right) < 1$. Moreover, when $\rho\left(M^{(m)}\right) < 1$, the process $\sum_{k=0}^{K}\underline{\Delta}_k(t)$ is strictly stationary and converges in $\mathbb{L}_m$ and *a.s.* and that its limit is strictly stationary and satisfies the Equation (1.5). Now assume that $E\{e_t^{2m}\} < +\infty$ and suppose that $\underline{Y}_t \in \mathbb{L}_m$, then

$$\begin{aligned}E\left\{\underline{Y}_t^{\otimes m}\right\} &= E\left\{\sum_{k=0}^{n}\left\{\prod_{j=0}^{k-1}M_{t-j}\right\}\underline{\omega}_{t-k} + \left\{\prod_{j=0}^{n}M_{t-j}\right\}\underline{Y}_{t-n-1}\right\}^{\otimes m} \\ &\geq \sum_{k=1}^{\infty}E\left\{\left\{\prod_{j=0}^{k-1}M_{t-j}^{\otimes m}\right\}\underline{\omega}_{t-k}^{\otimes m}\right\} \\ &= M_0^{(m)}\sum_{k=1}^{\infty}\left(M^{(m)}\right)^k E\left\{M_1^{\otimes m}\left(\underline{e}_{t-k}\right)\underline{\omega}_{t-k}^{\otimes m}\right\}.\end{aligned}$$

The above discussion leads to the following theorem.

Theorem 7 *Assume that* $E\{e_t^{2m}\} < +\infty$ *and* $\rho\left(E\left\{M_t^{\otimes m}\right\}\right) < 1$ *for any* $m \geq 1$*. Then, the* $PGARCH(p, q)$ *Model* (1.2) *has a SPS solution* $(\epsilon_t, h_t)_{t\in\mathbb{Z}}$ *such*

that $E\left\{\epsilon_t^{2m}\right\} < +\infty$. *The solution process is unique, causal and periodically ergodic.*
Conversely, if $\rho\left(E\left\{M_t^{\otimes m}\right\}\right) \geq 1$, *then there is no* SPS *solution to Model* (1.2) *such that* $E\left\{\epsilon_t^{2m}\right\} < +\infty$.

Example 8 *The* $PGARCH\,(1,1)$ *process has an unique,* SPS *and causal solution in* $\mathbb{L}_1$ *given by*

$$\epsilon_{st+v} = \sqrt{h_{st+v}}e_{st+v} \text{ with } h_{st+v} = \sum_{k=0}^{\infty}\left\{\prod_{i=0}^{k-1}\left(a_1(v-i)e_0^2 + b_1(v-i)\right)\right\}a_0(v-k)$$

if and only if $\prod_{v=1}^{s}\left(a_1(v)+b_1(v)\right) < 1$ *If* $E\left\{e_t^{2m}\right\} < +\infty$, *then* $E\left\{\epsilon_t^{2m}\right\} < +\infty$ *if and only if* $\prod_{v=1}^{s} E\left\{\left(a_1(v)e_0^2 + b_1(v)\right)^m\right\} < 1$.

Part II

Étude statistique

Chapter 2

The LSE approach for $PGARCH$ models

Abstract: This chapter deals with the asymptotic properties of parameters least squares estimates (LSE) for periodic $GARCH$ ($PGARCH$) and for $PARMA-PGARCH$ models. In this class of models, the parameters are allowed to switch between different regimes. Firstly, we give necessary and sufficient conditions ensuring the existence of stationary solutions (in periodic sense) and for the existence of moments of any order. Secondary, a least squares estimation approach for estimating $PGARCH$ and $PARMA-PGARCH$ models are discussed. The strong consistency and the asymptotic normality of the estimators are studied given mild regularity conditions, requiring strict stationarity and the finiteness of moments of some order for the errors term.

2.1 $PGARCH$ models and its probabilistic properties

A second order process $(\epsilon_n)_{n\in\mathbb{Z}}$ defined on some probability space $(\Omega, \mathcal{A}, P)$ is said to have a periodic generalized autoregressive conditional heteroscedastic representation with period $s > 0$ and orders p and q $(PGARCH(p,q))$ if it satisfies the non-linear equations

$$\forall n \in \mathbb{Z}\colon\ \epsilon_n = e_n\sqrt{h_n} \text{ and } h_n = a_0(n) + \sum_{i=1}^{q} a_i(n)\epsilon_{n-i}^2 + \sum_{j=1}^{p} b_j(n)h_{n-j} \tag{2.1}$$

where $(e_n)_{n\in\mathbb{Z}}$ is a sequence of independent identically distributed ($i.i.d.$) random variables defined on the same probability space $(\Omega, \mathcal{A}, P)$ with $E\{e_n\} = 0$ and $E\{e_n^2\} = 1$ and e_k in independent of ϵ_t for $k > t$. The parameters $(a_i(n))_{0\leq i\leq q}$ and $(b_i(n))_{1\leq i\leq p}$ are periodic in n with period s, i.e., for any $(n,k) \in \mathbb{Z}^2$: $a_i(n) = a_i(n+sk)$, $i = 0,...,q$ and $b_j(n) = b_j(n+sk)$, $j = 1,...,p$. So by setting $n = st+v$, $v = 1,...,s$, Equation (2.1) may be equivalently written in periodic notations as

$$\forall t \in \mathbb{Z}\text{: } \epsilon_{st+v} = e_{st+v}\sqrt{h_{st+v}} \text{ and } h_{st+v} = a_0(v) + \sum_{i=1}^{q} a_i(v)\epsilon_{st+v-i}^2 + \sum_{j=1}^{p} b_j(v)h_{st+v-j}, \tag{2.2}$$

which we will make heavy use of (2.2). In (2.2), ϵ_{st+v} (resp. h_{st+v}, e_{st+v}) refers to ϵ_t (resp. h_t, e_t) during the $v-th$ regime of cycle t, $(a_i(v))_{0\leq i\leq q}$ and $(b_i(v))_{1\leq i\leq p}$ are the model coefficients at season $v = 1,...,s$ such that $a_0(v) > 0$, $a_i(v) \geq 0$, $b_i(v) \geq 0$ for all v and $i \in \{1,...,p \vee q\}$. In what follows, we shall continue to use the non periodic notations (ϵ_t) (e_t) and (h_t) in preference to (ϵ_{st+v}) (e_{st+v}) and (h_{st+v}) whenever the periodicity is not paramount.

Noting that Equation (2.1) is intractable when we want to examine the probabilistic structure of this representation. Instead, we will work with the corresponding Markovian representation. Let $r = p+q$ and define
$\underline{\epsilon}_t = \left(\epsilon_t^2, \epsilon_{t-1}^2, ..., \epsilon_{t-q+1}^2, h_t, h_{t-1}, ..., h_{t-p+1}\right)'_{r\times 1}, \underline{e}_t = \left(a_0(t)e_t, \underline{O}'_{(q-1)}, a_0(t), \underline{O}'_{(p-1)}\right)'_{r\times 1}$
and let $A_t := \begin{pmatrix} A_t^1 & B_t^0 \\ A_t^0 & B_t^1 \end{pmatrix}_{r\times r}$ where

$$A_t^1 = \begin{pmatrix} a_1(t)e_t^2 ... a_q(t)e_t^2 \\ I_{(q-1)} \quad \underline{O}_{(q-1)} \end{pmatrix}, A_t^0 = \begin{pmatrix} a_1(t)...a_q(t) \\ O_{(q-1,q)} \end{pmatrix},$$
$$B_t^1 = \begin{pmatrix} b_1(t)...b_p(t) \\ I_{(p-1)} \quad \underline{O}_{(p-1)} \end{pmatrix}, B_t^0 = \begin{pmatrix} b_1(t)e_t^2 ... b_p(t)e_t^2 \\ O_{(p-1,p)} \end{pmatrix}.$$

Using the notation above, Equation (2.1) can be written as

$$\underline{\epsilon}_t = A_t\underline{\epsilon}_{t-1} + \underline{e}_t \tag{2.3}$$

and $\epsilon_t^2 = H'\underline{\epsilon}_t$ where $H = \left(1, \underline{O}'_{(r-1)}\right)'$. It is worth noting that $(A_t, \underline{e}_t)$ is an independent and periodically identically distributed pair of random matrix and vector, in the sense that $(A_{st}, \underline{e}_{st})_{t\in\mathbb{Z}}$ is an $i.i.d.$ process. In the next subsection, we are interested for the existence of causal solutions i.e., solutions which ϵ_t is measurable with respect to $\Im_t^{(e)} := \sigma(e_l, l \leq t)$ and its probabilistic properties.

2.1.1 Strict and second order periodic stationarity

From (2.3) we have the following recursion

$$\underline{\epsilon}_t = A(t)\underline{\epsilon}_{t-s} + \underline{\xi}_t \tag{2.4}$$

where $A(t) := \prod_{i=0}^{s-1} A_{t-i}$ and where $\underline{\xi}_t := \sum_{k=1}^{s-1} \left\{\prod_{i=0}^{k-1} A_{t-i}\right\} \underline{e}_{t-k} + \underline{e}_t$. Define the top-Lyapunov exponent associated with the strictly stationary and ergodic sequence of random matrices $A = (A(t))_{t\in\mathbb{Z}}$ by

$$\gamma^{(s)}(A) := \inf_{t>0} \frac{1}{t} E\left\{\log \left\| \prod_{i=0}^{t-1} A(s(t-i)) \right\|\right\}$$

whenever $\sum_{v=1}^{s} E\left\{\log^+ \|A_v\|\right\} < \infty$ where $\log^+ x = \max\{\log x, 0\}$ for $x > 0$. Since (2.3) and (2.4) are valid for all integer t, by successive substitution we obtain the following formal series $\left(\underline{\epsilon}_t^{(1)}\right)_{t\in\mathbb{Z}}$ and $\left(\underline{\epsilon}_t^{(2)}\right)_{t\in\mathbb{Z}}$

$$\underline{\epsilon}_t^{(1)} = \sum_{k=1}^{\infty} \left\{\prod_{i=0}^{k-1} A_{t-i}\right\} \underline{e}_{t-k} + \underline{e}_t,\ \underline{\epsilon}_t^{(2)} = \sum_{k=1}^{\infty} \left\{\prod_{i=0}^{k-1} A(t-is)\right\} \underline{\xi}_{t-ks} + \underline{\xi}_t. \tag{2.5}$$

The usefulness of these series are examined in the next theorem.

Theorem 1 *If* $\gamma^{(s)}(A) < 0$, *then:*

1. *Equation* (2.4) *admits an unique, causal, SPS and periodically ergodic solution given by the series* $\left(\underline{\epsilon}_t^{(2)}\right)_{t\in\mathbb{Z}}$ *which converges a.s.*
2. *The Series* $\left(\underline{\epsilon}_t^{(1)}\right)_{t\in\mathbb{Z}}$ *converges a.s and constitute the unique, causal, SPS and periodically ergodic solution of Equation* (2.3).
3. $\underline{\epsilon}_t^{(1)} = \underline{\epsilon}_t^{(2)}$ *a.s.*

Proof. 1. Since $E\left\{\log^+ \|A(t)\|\right\}$ and $E\left\{\log^+ \left\|\underline{\xi}_t\right\|\right\}$ are finite, the proof follows from the Theorem 1.1 of Bougerol and Picard [16]. **2.** The proof fellows from standard arguments (c.f. Bibi and Aknouche [9]). **3.** By setting $\underline{\epsilon}_t^{(1)}(n) = \sum_{k=1}^{n} \left\{\prod_{i=0}^{k-1} A_{t-i}\right\} \underline{e}_{t-k} + \underline{e}_t$, $\underline{\epsilon}_t^{(2)}(n) = \sum_{k=1}^{n} \left\{\prod_{i=0}^{k-1} A(t-si)\right\} \underline{\xi}_{t-sk} + \underline{\xi}_t$, then, we can check after some tedious computations that for any $1 \leq m \leq s$, there is a constant $K > 0$ such that

$$\left\|\underline{\epsilon}_t^{(2)}(n) - \underline{\epsilon}_t^{(1)}(sn+m)\right\| \leq K \left\|\prod_{j=0}^{n} A(t-sj)\right\| \to 0$$

as $n \to \infty$, *a.s.* ■

Remark 2 *The need of the condition $\gamma^{(s)}(A) < 0$ for the existence of SPS solution can be shown by the same argument as in Bibi and Aknouche [9].*

Remark 3 *Similarly to the classical results on the GARCH processes theory (see for instance Berkes et al.[5]), if $\gamma^{(s)}(A) < 0$ then there exists $\delta > 0$ such that $E\left\{h_t^{\delta}\right\} < +\infty$ and $E\left\{\epsilon_t^{2\delta}\right\} < +\infty$ (see also [2]).*

Example 4 *For the $PGARCH(1,1)$ model, after some tedious algebra we find that the necessary and sufficient condition ensuring the existence of SPS solution is that $\sum_{v=1}^{s} E\left\{\log\left(a_1(v)e_0^2 + b_1(v)\right)\right\} < 0$. It is worth noting that the existence of regimes which satisfy $E\left\{\log\left(a_1(v)e_0^2 + b_1(v)\right)\right\} > 0$ does not preclude strict periodic stationarity.*

The top-Lyapunov exponent seems difficult to obtain explicitly; however it can easily be obtained by simulation using Equation (2.4). Hence, and for the estimation purpose, the SPS causal solutions need to belong to $\mathbb{L}_2$. The most one of characteristic of these solutions, is that the process $\left(\epsilon_t, \sqrt{h_t}\right)_{t\in\mathbb{Z}}$ defined by (2.1) has a periodic covariance structure in the sense that $Cov(\epsilon_{l+st}, \epsilon_{k+st}) = Cov(\epsilon_l, \epsilon_k)$ for all integers l, k and t. Such series are also called periodically correlated (PC) processes.

Theorem 5 *The $PGARCH$ process (2.1) admit a PC process solution if and only if*

$$\rho(A) < 1. \tag{2.6}$$

where $A := E\{A(t)\}$. Moreover, the solution process is unique, SPS, periodically ergodic, causal and given by the first component of one of the processes $\left(\underline{\epsilon}_t^{(1)}\right)_{t\in\mathbb{Z}}$ or $\left(\underline{\epsilon}_t^{(2)}\right)_{t\in\mathbb{Z}}$ defined in (2.5).

Proof. The condition is obviously necessary using (2.3). To show that (2.6) is also sufficient, we define the following $\mathbb{R}^r$−valued processes $(\underline{S}_n(t), \underline{\Delta}_n(t))_{(t,n)\in\mathbb{Z}\times\mathbb{Z}}$

$$\underline{S}_n(t) := \begin{cases} \underline{O}_{(r)} & \text{if } n < 0 \\ \underline{e}_t + A_t \underline{S}_{n-1}(t-1), & \text{if } n \geq 0 \end{cases}$$

and $\underline{\Delta}_n(t) := \underline{S}_n(t) - \underline{S}_{n-1}(t)$. It is easily seen that for all $n \geq 0$, $\underline{S}_n(t)$ and $\underline{\Delta}_n(t)$ are measurable functions of $e_t, e_{t-1}, ..., e_{t-n}$. Hence, for any fixed $n \geq 0$ the processes $(\underline{S}_n(t))_{t\in\mathbb{Z}}$ and $(\underline{\Delta}_n(t))_{t\in\mathbb{Z}}$ are SPS and periodically ergodic. From the definition of $\underline{S}_n(t)$ and $\underline{\Delta}_n(t)$ we have

$$\underline{\Delta}_n(t) := \begin{cases} \underline{O}_{(r)} & \text{if } n < 0 \\ \underline{e}_t & \text{if } n = 0 \\ A_t \underline{\Delta}_{n-1}(t-1), & \text{if } n > 0 \end{cases}$$

and thus for any $n \geq 1$, we have
$\underline{\Delta}_n(t) = \left\{ \prod_{i=0}^{n-1} A_{t-i} \right\} \underline{e}_{t-n}$ and $\underline{S}_n(t) = \sum_{k=0}^{n} \underline{\Delta}_k(t)$. Since $\left\{ \prod_{i=0}^{n-1} A_{t-i} \right\} \underline{e}_{t-n}$ has positive elements, we have $E \left\| \underline{\Delta}_n(t) \right\| = \left\| E \left\{ \prod_{i=0}^{n-1} A_{t-i} \right\} E \left\{ \underline{e}_{t-n} \right\} \right\| \leq K \left\| A^{\left[\frac{n}{s}\right]} \right\|$ for $n \geq 1$, where K is some positive constant and $[x]$ denotes the smallest integer greater than or equal to x. Thus $\limsup_{n\to\infty} E \left\| \underline{\Delta}_n(t) \right\| \leq K \limsup_{n\to\infty} \left(\rho\left(A\right)\right)^{\left[\frac{n}{s}\right]}$. Hence under (2.6), $\limsup_{n\to\infty} E \left\| \underline{\Delta}_n(t) \right\| = 0$ and hence $\underline{S}_n(t)$ converges in $\mathbb{L}_1$ to some limit say $\lim_{n\to\infty} \underline{S}_n(t)$ which satisfies the Equation (2.3) and thus the both series in (2.5). The rest of the assertions are immediate. ■

2.1.2 The existence of higher-order moments

In this subsection, we derive necessary and sufficient conditions for the finiteness of $E\left\{\epsilon_t^{2m}\right\}$, for any integer $m > 1$. By the $\mathbb{L}_m-$theory, $m \geq 1$ the problem of existence of $E\left\{\epsilon_t^{2m}\right\}$ is now reduces to the convergence of $\left(\underline{S}_n(t)\right)_{n\geq 0}$ in $\mathbb{L}_m$ for all $t \in \mathbb{Z}$. As it is shown in Theorem (5) $\left(\underline{S}_n(t)\right)_{n\geq 0}$ converges to $\underline{\epsilon}_t$ in $\mathbb{L}_1$. The key quantity of interest in determining $\mathbb{L}_m$ convergence is $\underline{V}_n := E\left\{\underline{\Delta}_n^{\otimes m}(t)\right\}$. For this purpose assuming that $E\left\{e_0^{2m}\right\} < +\infty$ and set $A^{(m)} := E\left\{A^{\otimes m}(t)\right\}$. From (2.5) there exists a constant $K > 0$ such that

$$
\begin{aligned}
\left\|\underline{\epsilon}_t\right\|_m &= E\left\{\left\|\underline{\epsilon}_t\right\|^m\right\}^{\frac{1}{m}} \leq \sum_{n\geq 0} \left\|\underline{\Delta}_n(t)\right\|_m \leq K \sum_{n\geq 0} \left\| \prod_{i=1}^{n-1} A_{t-i} \right\|_m \\
&\leq K \left\{ \sum_{n\geq 0} \left\| \left(A^{(m)}\right)^{\left[\frac{n}{s}\right]} \right\|^{\frac{1}{m}} \right\}.
\end{aligned}
$$

Hence, if $\rho\left(A^{(m)}\right) < 1$, then $\left\| \left(A^{(m)}\right)^{\left[\frac{n}{s}\right]} \right\|$ converges to 0 with exponential rate as $n \to \infty$. Since $\left\|\epsilon_t^2\right\|_m \leq \left\|\underline{\epsilon}_t^2\right\|_m$ thus a sufficient condition for the finiteness of $E\left\{\epsilon_t^{2m}\right\}$ is that $\rho\left(A^{(m)}\right) < 1$. Moreover, when $\rho\left(A^{(m)}\right) < 1$, the process $\sum_{k=0}^{n} \underline{\Delta}_k(t)$ is SPS and converges in $\mathbb{L}_m$ and $a.s.$ and that its limit is SPS and satisfies the Equations in (2.5). Now suppose that $\underline{\epsilon}_t \in \mathbb{L}_m$, then $E\left\{\underline{\epsilon}_t^{\otimes m}\right\} = E\left\{A(t)\underline{\epsilon}_{t-s} + \underline{\xi}_t\right\}^{\otimes m} \geq A^{(m)} E\left\{\underline{\epsilon}_{t-s}^{\otimes m}\right\}$ and thus

Theorem 6 *Assume that $E\left\{e_t^{2m}\right\} < +\infty$ and $\rho\left(A^{(m)}\right) < 1$ for any $m \geq 1$. Then, the $PGARCH(p,q)$ Model (2.1) has a SPS solution $\left(\epsilon_t, h_t\right)_{t\in\mathbb{Z}}$ such that $E\left\{\epsilon_t^{2m}\right\} < +\infty$. The solution process is unique, causal and periodically ergodic. Conversely, if $\rho\left(A^{\otimes m}\right) \geq 1$, then there is no SPS solution to Model (2.1) such that $E\left\{\epsilon_t^{2m}\right\} < +\infty$.*

Example 7 *The $PGARCH\,(1,1)$ process has an unique, SPS and causal solution in $\mathbb{L}_1$ given by $\epsilon_t = \sqrt{h_t} e_t$ and $h_t = \sum\limits_{k=0}^{\infty} \left\{ \prod\limits_{i=0}^{k-1} \left(a_1(t-i)e_{t-i}^2 + b_1(t-i) \right) \right\} a_0(t-k)$ if and only if $\prod\limits_{v=1}^{s} \left(a_1(v) + b_1(v) \right) < 1$. If $E\left\{e_t^{2m}\right\} < +\infty$, then $E\left\{\epsilon_t^{2m}\right\} < +\infty$ if and only if $\prod\limits_{v=1}^{s} E\left\{ \left(a_1(v)e_0^2 + b_1(v) \right)^m \right\} < 1$.*

2.2 Least squares estimation for $PGARCH(p,q)$ processes

In this section, the large sample properties of the least squares estimates (LSE) for $PGARCH$ model coefficients are studied. The process is thus described with the vector of parameters $\underline{\theta} = (\underline{\theta}'(1), ..., \underline{\theta}'(s))'$ where $\underline{\theta}(v) = (a_0(v), a_1(v), ..., a_q(v), b_1(v), ..., b_p(v))'$, $v = 1, ..., s$. The vector $\underline{\theta}$ belongs to a parameter space $\Theta_{\underline{\theta}} := \{ \underline{\theta} : \underline{\theta} \in \left( ]0, \infty[\times [0, \infty[^{(p+q)} \right)^s \}$. The orders p,q and the period s are supposed to be known and the true parameter value is unknown and is denoted by $\underline{\theta}_0$. Let $\{\epsilon_1, \epsilon_2, ..., \epsilon_N\}$ be a realization of length $N = sn$ of the unique, causal, SPS solution $(\epsilon_t)_{t\in\mathbb{Z}}$ to Model (2.1). Conditionally on initial values $\epsilon_0, \epsilon_{-1}, ..., \epsilon_{1-q}, \widehat{h}_0, \widehat{h}_{-1}, ..., \widehat{h}_{1-p}$ the LSE of $\underline{\theta}$ is defined as any measurable solution $\widehat{\underline{\theta}}_n$ of

$$\widehat{\underline{\theta}}_n = Arg \min_{\underline{\theta}\in\Theta_{\underline{\theta}}} \widehat{Q}_n\left(\underline{\theta}\right) \tag{2.7}$$

where $\widehat{Q}_n\left(\underline{\theta}\right) := \frac{1}{n} \sum\limits_{t=0}^{n-1} \widehat{l}_t\left(\underline{\theta}\right)$ with $\widehat{l}_t\left(\underline{\theta}\right) := \frac{1}{s} \sum\limits_{v=1}^{s} \widehat{\eta}_{st+v}^2\left(\underline{\theta}\right)$ and $\widehat{\eta}_{st+v}\left(\underline{\theta}\right) = z_{st+v} - \log \widehat{h}_{st+v}\left(\underline{\theta}\right)$ in which z_{st+v} is motivated by the regression relationship $z_{st+v} := \log \epsilon_{st+v}^2 - E\left\{\log e_{st+v}^2\right\}$ and $\widehat{h}_{st+v}(\theta)$ are defined recursively by $\widehat{h}_{st+v}\left(\underline{\theta}\right) = a_0(v) + \sum\limits_{i=1}^{q} a_i(v)\epsilon_{st+v-i}^2 + \sum\limits_{j=1}^{p} b_j(v)\widehat{h}_{st+v-j}\left(\underline{\theta}\right)$. For instance, the initial values can be chosen as $\epsilon_0^2 = \widehat{h}_0 = \epsilon_1^2, \epsilon_{-1}^2 = \widehat{h}_{-1} = \epsilon_1^2, \ldots, \epsilon_{1-p\vee q}^2 = \widehat{h}_{1-p\vee q} = \epsilon_1^2$. Noting that the choice of the initial values does not matter to the asymptotic properties of the LSE, it may have importance from a practical purpose as building h_t. Hence, and from a theoretical point of view, it is more convenient to work with $l_t\left(\underline{\theta}\right) := \frac{1}{s} \sum\limits_{v=1}^{s} \eta_{st+v}^2\left(\underline{\theta}\right)$ where $\eta_{st+v}\left(\underline{\theta}\right) := z_{st+v} - \log h_{st+v}(\underline{\theta})$ and where $h_{st+v}\left(\underline{\theta}\right) = a_0(v) + \sum\limits_{i=1}^{q} a_i(v)\epsilon_{st+v-i}^2 + \sum\limits_{j=1}^{p} b_j(v)h_{st+v-j}\left(\underline{\theta}\right)$ because $\left(l_t\left(\underline{\theta}\right)\right)_{t\in\mathbb{Z}}$ is SPS process whereas $\left(\widehat{l}_t\left(\underline{\theta}\right)\right)_{t\in\mathbb{Z}}$ is not due to the presence of initial values. However, instead of

$\widehat{Q}_n(\underline{\theta})$, we consider the minimization of the function $\widehat{O}_n(\underline{\theta}) := \frac{1}{n}\sum_{t=0}^{n-1} l_t(\underline{\theta})$ which constitute an approximation of $\widehat{Q}_n(\underline{\theta})$ in the sense that $\left|\widehat{O}_n(\underline{\theta}) - \widehat{Q}_n(\underline{\theta})\right|$ decays to zero uniformly $a.s.$ with geometric rate on the certain compact set. Noting here that the functions $\widehat{\zeta}_{st+v}(\underline{\theta}) = \frac{\epsilon^2_{st+v}}{\widehat{h}_{st+v}(\underline{\theta})} + \log\widehat{h}_{st+v}(\underline{\theta})$, $\widehat{\zeta}_{st+v}(\underline{\theta}) = \frac{\epsilon^2_{st+v}}{\widehat{h}_{st+v}(\underline{\theta})} - 1$ and $\widehat{\zeta}_{st+v}(\underline{\theta}) = \epsilon^2_{st+v} - \widehat{h}_{st+v}(\underline{\theta})$ can be used instead of $\widehat{\eta}_{st+v}(\underline{\theta})$ above.

Let $\mathcal{A}_v(z) = \sum_{j=1}^{q} a_{0j}(v)z^j$, $\mathcal{B}_v(z) = 1 - \sum_{j=1}^{p} b_{0j}(v)z^j$ and $\gamma^{(s)}(A^0)$ be the top-Lyapunov exponent associated with the sequence $(A^0_t)_{t\in\mathbb{Z}}$ where A^0_t is just the matrix A_t defined in Section 2.1 with $\underline{\theta}_0$ instead $\underline{\theta}$. Then to show the strong consistency, the following assumptions will be made.

A1. $\underline{\theta}_0 \in \Theta_{\underline{\theta}}$ and $\Theta_{\underline{\theta}}$ is compact

A2. $\gamma^{(s)}(A^0) < 0$ and $\sup_{\underline{\theta}\in\Theta} \rho\left(\prod_{v=0}^{s-1} B^1_{s-v}\right) < 1.$

A3. for all $v \in \{1, ..., s\}$, $\mathcal{A}_v(z)$ and $\mathcal{B}_v(z)$ have no common roots and $a_{0q}(v) + b_{0p}(v) \neq 0$.

A4. $(e^2_t)_{t\in\mathbb{Z}}$ has a non-degenerate distribution.

In Assumption **A1**, the compactness of Θ is assumed in order that several results from real analysis may be used. As seen in Remark 3, the first assumption in **A2** ensures the existence of some finite moments for the SPS solution of (2.1) which is the key for proving the strong consistency of LSE, and the second assumption is imposed in order to obtain $h_t(\underline{\theta})$ as a causal solution of $\{\epsilon_t, \epsilon_{t-1}, ...\}$, i.e. $h_{st+v}(\underline{\theta}) = \alpha_0(v) + \sum_{j=1}^{\infty} \alpha_j(v)\epsilon^2_{st+v-j}$ for all $v \in \{1, ..., s\}$ in which the weights $\alpha_j(v)$ satisfy $\max_{1\leq v\leq s} \alpha_j(v) = O(\rho^j)$ with $\rho \in]0, 1[$. While **A3** and **A4** are made to guarantee the identifiability of the parameters. The next theorem shows the strong consistency of LSE for $PGARCH$ processes.

Theorem 8 *Let* $\left(\widehat{\underline{\theta}}_n\right)$ *be the sequence of LSE satisfying* (2.7). *Then, under* ***A1-A4****, almost surely* $\widehat{\underline{\theta}}_n \to \underline{\theta}_0$ *as* $n \to \infty$.

In order to establish the asymptotic normality of LSE let $\kappa := Var\{\log e^2_t\}$ and consider the additional assumptions

A5. $\underline{\theta}_0 \in \overset{o}{\Theta}_{\underline{\theta}}$ where $\overset{o}{\Theta}_{\underline{\theta}}$ denotes the interior of $\Theta_{\underline{\theta}}$.

A6. $E\left\{e_t^4\right\} < \infty$

The second main result of this section is the following

Theorem 9 *Under **A1-A6**,* $\sqrt{n}\left(\widehat{\underline{\theta}}_n - \underline{\theta}_0\right) \rightsquigarrow \mathcal{N}\left(\underline{O}, \kappa\mathcal{I}^{-1}\right)$ *where* $\mathcal{I}^{-1} := diag\left\{\mathcal{I}_l^{-1}, l = 1, ..., s\right\}$ *and each block matrix is given by*

$$\mathcal{I}_l := E_{\underline{\theta}_0}\left\{\frac{\partial l_t\left(\underline{\theta}\right)}{\partial\underline{\theta}(l)}\frac{\partial l_t\left(\underline{\theta}\right)}{\partial\underline{\theta}(l)'}\right\} = \sum_{v=1}^{s} E_{\underline{\theta}_0}\left\{\frac{1}{h_{st+v}^2\left(\underline{\theta}\right)}\frac{\partial h_{st+v}\left(\underline{\theta}\right)}{\partial\underline{\theta}(l)}\frac{\partial h_{st+v}\left(\underline{\theta}\right)}{\partial\underline{\theta}(l)'}\right\},\ l = 1, ..., s.$$

Remark 10 *For Gaussian $QMLE$ we have* $\sqrt{n}\left(\widehat{\underline{\theta}}_n - \underline{\theta}_0\right) \rightsquigarrow \mathcal{N}\left(0, Var\left\{e_t^2\right\}\mathcal{I}^{-1}\right)$ *(see [2]). Hence, the performance of LSE with respect to $QMLE$ can be captured by* $\lambda := Var\left\{e_t^2\right\}/\kappa$ *which depends on the distribution of* $\left(e_t\right)_{t\in\mathbb{Z}}$.

Remark 11 *Since $\mathcal{I}$ is $s-$block diagonal matrices implies the asymptotic independency of the estimates for each regime $v \in \{1, ..., s\}$.*

Next, we establish the law of iterated logarithm (LIL) for $LSE - PGARCH$ estimator. This provide almost surely a flexible, completely consistent and bounds for $\widehat{\underline{\theta}}_n$.

Theorem 12 *Under Assumptions **A1-A6** we have*

$$\limsup_n \sqrt{\frac{n}{2\kappa\log\log n}}\mathcal{I}^{1/2}\left(\widehat{\underline{\theta}}_n - \underline{\theta}_0\right) \leq \underline{1}_{(s(p+q+1))}$$

where $\underline{1}_{(k)} = (1, ..., 1)' \in \mathbb{R}^k$.

Let us now apply the forgoing results to the first order $PARCH$ process given by $\epsilon_{st+v} = e_{st+v}\sqrt{h_{st+v}}$ with $h_{st+v} = a_0(v) + a_1(v)\epsilon_{st+v-1}^2$where $a_0(v) > 0$ and $a_1(v) \geq 0$. It is easily seen that the SPS condition for $PARCH\,(1)$ reduce to $0 \leq \prod_{v=1}^{s} a_{01}(v) < \exp\left(-sE\left\{\log e_0^2\right\}\right) := \alpha$ under which supposing that $\underline{\theta}_0 = \left(\underline{\theta}_0'(1), ..., \underline{\theta}_0'(s)\right)'$ with $\underline{\theta}_0'(v) := \left(a_0(v), a_{01}(v)\right)'$ belonging to a compact $\Theta_{\underline{\theta}}$ of the form $\Theta_{\underline{\theta}} = \left(\left[\epsilon, \frac{1}{\varepsilon}\right] \times \left[0, \alpha^{\frac{1}{s}} - \epsilon\right]\right)^s$ for any $\epsilon > 0$. The LSE is thus by Theorem 8 strongly consistent. Moreover, if $\Theta_{\underline{\theta}} = \overset{o}{\Theta}_{\underline{\theta}}$, then from Theorem 9 the LSE is also asymptotically $\mathcal{N}\left(\underline{O}, \kappa\mathcal{I}^{-1}\right)$ where $\mathcal{I}^{-1} := diag\left\{\mathcal{I}_l^{-1}, l = 1, ..., s\right\}$ with

$$\mathcal{I}_l := \sum_{v=1}^{s} E_{\underline{\theta}_0}\left\{\frac{1}{h_{st+v}^2\left(\underline{\theta}\right)}\begin{pmatrix} 1 & \epsilon_{st+v-1}^2 \\ \epsilon_{st+v-1}^2 & \epsilon_{st+v-1}^4 \end{pmatrix}\right\}.$$

2.3 Estimation of $PARMA-PGARCH$ processes

In this section, our aim is to extend the previous results to the case where the $PGARCH$ process is not directly observed. Since the process $(\epsilon_t)_{t\in\mathbb{Z}}$ solution of (2.1) is a martingale difference and thus can be used as the innovation of a periodic $ARMA$ ($PARMA$) process. The estimation of $PARMA-PGARCH$ models was considered by Aknouche and Bibi [2] using the QML approach. Here, we shall investigated the LSE method for estimating $PARMA-PGARCH$ processes. For this purpose we will consider a set of observations $\{X_1, ..., X_N; N = ns\}$ obtained from a SPS and causal $PARMA\,(P,Q)$-$PGARCH\,(p,q)$ process generated by the equations

$$\begin{cases} X_t - \mu(t) = \sum_{i=1}^{P} \phi_i(t)\left(X_{t-i} - \mu(t-i)\right) + \epsilon_t - \sum_{j=1}^{Q} \varphi_j(t)\epsilon_{t-j} \\ \epsilon_t = \sqrt{h_t} e_t \\ h_t = a_0(t) + \sum_{i=1}^{q} a_i(t)\epsilon_{t-i}^2 + \sum_{j=1}^{p} b_j(t)h_{t-j} \end{cases} \tag{2.8}$$

the coefficients $\mu(t)$, $(\phi_i(t))_{1\le i\le P}$ and $\left(\varphi_j(t)\right)_{1\le j\le Q}$ are periodic in t with known period s. The vector of parameters of interest is denoted by $\underline{\pi}:= \left(\underline{\beta}', \underline{\theta}'\right)'$ where $\underline{\beta}= \left(\underline{\beta}'(1), ..., \underline{\beta}'(s)\right)'$ with $\underline{\beta}(v) := \left(\mu(v), \phi_1(v), ..., \phi_P(v), \varphi_1(v), ..., \varphi_Q(v)\right)'$, $1 \le v \le s$ and the parameter space is $\Theta_{\underline{\pi}} \subset \Theta_{\underline{\beta}} \times \Theta_{\underline{\theta}}$ where $\Theta_{\underline{\beta}} := \mathbb{R}^{s(P+Q+1)}$. The true parameter value denoted by $\underline{\pi}_0 = \left(\underline{\beta}_0', \underline{\theta}_0'\right)'$ is supposed to belong to some Euclidian space Φ. If $q \ge Q$ the initial values $X_0, ..., X_{1-P-(q-Q)}, \widetilde{\epsilon}_{-(q-Q)}$, $..., \widetilde{\epsilon}_{-1-q}$, $\widetilde{h}_0, ..., \widetilde{h}_{1-p}$ allow to compute $\widetilde{\epsilon}_t\left(\underline{\beta}\right)$ for $t = 1 + Q - q, ..., N$ and $\widetilde{h}_t\left(\underline{\pi}\right)$ for $t = 1, ..., N$, from

$$\begin{cases} \widetilde{\epsilon}_t := \widetilde{\epsilon}_t\left(\underline{\beta}\right) = X_t - \mu(t) - \sum_{i=1}^{P} \varphi_i(t)\left(X_{t-i} - \mu(t-i)\right) + \sum_{j=1}^{Q} \phi_j(t)\widetilde{\epsilon}_{t-j} \\ \widetilde{h}_t := \widetilde{h}_t\left(\underline{\pi}\right) = a_0(t) + \sum_{i=1}^{q} a_i(t)\widetilde{\epsilon}_{t-i}^2 + \sum_{j=1}^{p} b_j(t)\widetilde{h}_{t-j}. \end{cases}$$

When $q < Q$ the required initial values are $X_0, ..., X_{1-(q-Q)}, \widetilde{\epsilon}_{-(q-Q)}, ..., \widetilde{\epsilon}_{1-Q}, \widetilde{h}_0, ...$ $, \widetilde{h}_{1-p}$.

The sequence of random vectors $\widehat{\underline{\pi}}_n = \left(\widehat{\underline{\beta}}_n', \widehat{\underline{\theta}}_n'\right)'$ is called two stages least squares estimator if it satisfies, almost surely

$$\widehat{\underline{\beta}}_n = Arg \min_{\underline{\beta}\in\Theta_{\underline{\beta}}} \widehat{Q}_{1,n}\left(\underline{\beta}\right), \; \widehat{\underline{\pi}}_n := Arg \min_{\underline{\theta}\in\Theta_{\underline{\theta}}} \widehat{Q}_{2,n}\left(\widehat{\underline{\beta}}_n, \underline{\theta}\right)$$

where $\widehat{Q}_{1,n}\left(\underline{\beta}\right) := \frac{1}{n}\sum_{t=0}^{n-1}\widehat{l}_{1,t}\left(\underline{\beta}\right)$, with $\widehat{l}_{1,t}\left(\underline{\beta}\right) := \frac{1}{s}\sum_{v=1}^{s}\widehat{\epsilon}^{2}_{st+v}\left(\underline{\beta}\right)$ and where $\widehat{Q}_{2,n}\left(\underline{\pi}\right) := \frac{1}{n}\sum_{t=0}^{n-1}\widehat{l}_{2,t}\left(\underline{\pi}\right)$, with $\widehat{l}_{2,t}\left(\underline{\pi}\right) := \frac{1}{s}\sum_{v=1}^{s}\widehat{\eta}^{2}_{st+v}\left(\underline{\pi}\right)$. For $v = 1, ..., s$, consider the polynomials $\Phi_v(z) = 1 - \sum_{i=1}^{P}\phi_{0i}(v)z^i$, $\ominus_v(z) = 1 - \sum_{i=1}^{Q}\varphi_{0i}(v)z^i$ and the matrices

$$\Phi_v := \begin{pmatrix} \phi_1(v) & \cdots & \phi_P(v) \\ I_{(P-1)} & & \underline{O}_{(P-1)\times 1} \end{pmatrix}, \Psi_v := \begin{pmatrix} \varphi_1(v) & \cdots & \varphi_Q(v) \\ I_{(Q-1)} & & \underline{O}_{(Q-1)\times 1} \end{pmatrix}.$$

and we introduce the following conditions

A.7 $\sup_{\underline{\beta}\in\Theta_{\underline{\beta}}} \rho\left(\prod_{v=0}^{s-1}\Phi_{s-v}\right) < 1, \sup_{\underline{\beta}\in\Theta_{\underline{\beta}}} \rho\left(\prod_{v=0}^{s-1}\Psi_{s-v}\right) < 1.$

A.8 The polynomials $\Phi_v(z)$ and $\Psi_v(z)$ have no common roots with $\phi_{0P}(v) \neq 0$ or $\varphi_{0Q}(v) \neq 0$ for all $v = 1, ..., s$.

The first inequality in Assumption **A.7** is the top-Lyapunov exponent associated with $PARMA$ model and thus implies the causality of a SPS solution. The second one, is the invertibility condition of the $PARMA$ model (2.8). Hence, under **A.7**, it follows that $(X_t)_{t\in\mathbb{Z}}$ and $(\epsilon_t)_{t\in\mathbb{Z}}$ can be related through the infinite order moving average and autoregressive expansions

$$X_{st+v} - \mu(v) = \sum_{i=0}^{\infty}\alpha_i\left(v\right)\epsilon_{st+v-i} \text{ and } \epsilon_{st+v} = \sum_{i=0}^{\infty}\beta_i\left(v\right)\left(X_{st+v-i} - \mu(v-i)\right) \tag{2.9}$$

In (2.9), the weights $\alpha_i\left(v\right)$ and $\beta_i\left(v\right)$ satisfy

$$\sup_{1\leq v\leq s}\left|\alpha_i\left(v\right)\right| = O(\rho^i) \text{ and } \sup_{1\leq v\leq s}\left|\beta_i\left(v\right)\right| = O(\rho^i) \text{ with } 0 < \rho < 1.$$

Theorem 13 *Let* $(X_t)_{t\in\mathbb{Z}}$ *be a SPS PARMA process satisfying* (2.8). *Then under **A1-A3** and **A7-A8**, almost surely* $\widehat{\underline{\pi}}_n \to \underline{\pi}_0$ *as* $n \to \infty$.

As the standard $ARMA - GARCH$ case (cf. [21] and [4]) we will prove asymptotic normality of $\widehat{\underline{\pi}}_n$ under the fourth order moment condition on the $(\epsilon_t)_{t\in\mathbb{Z}}$. From Theorem 6, such condition is expressed by $\rho\left(A^{(2)}\right) < 1$ (see Subsection 2.1.2). Thus we make the following assumptions

A.9 $\rho\left(A^{(2)}\right) < 1$.

A.10 $\underline{\pi}_0$ is in the interior of $\Theta_{\underline{\pi}}$.

Now, we are able to derive the limit distribution of $\widehat{\underline{\pi}}_n$.

Theorem 14 *Let* $(X_t)_{t\in\mathbb{Z}}$ *be a SPS PARMA process satisfying* (2.8). *Then under* ***A1-A10*** *we have*

$$\sqrt{n}\left(\widehat{\underline{\pi}}_n - \underline{\pi}_0\right) \rightsquigarrow \mathcal{N}\left(\begin{pmatrix} \underline{O} \\ \underline{O} \end{pmatrix}, \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}\right)$$

where $V_{11} = J_{11}^{-1} I_{11} J_{11}^{-1}$, $V_{12} = V_{21}' = J_{22}^{-1}\left(I_{21} + J_{21}J_{11}^{-1}I_{11}\right)J_{11}^{-1}$, $V_{22} = J_{22}^{-1}(I_{22} + J_{21}J_{11}^{-1}I_{11}J_{11}^{-1}J_{12} - I_{21}J_{11}^{-1}J_{12} - J_{21}J_{11}^{-1}I_{12})J_{22}^{-1}$ *with almost surly*

$$\begin{aligned}
I_{11} &= \lim_{n\to\infty} Var_{\underline{\beta}_0}\left\{\sqrt{n}\frac{\partial}{\partial\underline{\beta}}\widehat{Q}_{1,n}\left(\underline{\beta}\right)\right\},\ I_{22} = \lim_{n\to\infty} Var_{\underline{\pi}_0}\left\{\sqrt{n}\frac{\partial}{\partial\underline{\theta}}\widehat{Q}_{2,n}\left(\underline{\pi}\right)\right\}, \\
I_{12} &= \lim_{n\to\infty} E_{\underline{\pi}_0}\left\{n\frac{\partial}{\partial\underline{\beta}}\widehat{Q}_{1,n}\left(\underline{\beta}\right)\frac{\partial}{\partial\underline{\theta}'}\widehat{Q}_{2,n}\left(\underline{\pi}\right)\right\},\ I_{21} = I_{12}' \\
J_{11} &= \lim_{n\to\infty} Var_{\underline{\beta}_0}\left\{\frac{\partial^2}{\partial\underline{\beta}\partial\underline{\beta}'}\widehat{Q}_{1,n}\left(\underline{\beta}\right)\right\},\ J_{22}(v) = \lim_{n\to\infty} Var_{\underline{\beta}_0}\left\{\frac{\partial^2}{\partial\underline{\theta}\partial\underline{\theta}'}\widehat{Q}_{2,n}\left(\underline{\pi}\right)\right\} \\
J_{12} &= \lim_{n\to\infty} Var_{\underline{\beta}_0}\left\{\frac{\partial^2}{\partial\underline{\beta}\partial\underline{\theta}'}\widehat{Q}_{2,n}\left(\underline{\pi}\right)\right\},\ J_{21} = J_{12}'.
\end{aligned}$$

2.4 Appendix

The main aim here is to reveal the basic assumptions and to quantify the asymptotic properties of the LSE for $PGARCH$ and for $PARMA-PGARCH$ processes. The proof of Theorems 8 is by now standard and follows from similar arguments used in showing the strong consistency of the $QMLE-PGARCH$ models (cf. Aknouche and Bibi [2]) and hence, we do not detail the proof. Since there are several similarities between the standard $ARMA$ and $GARCH$ and its periodic versions $PARMA$ and $PGARCH$, certain steps of the proof for the LSE for $PGARCH$ and for $PARMA-PGARCH$ processes are similar in spirit to that of the standard $GARCH$ and $ARMA-GARCH$ one. Thus, we give details of proof only when it seems pertinent to us and refer to Aknouche and Bibi [2], Francq and Zakoïan [20], [21] or Straumann and Mikosch [46] for further details.

2.4.1 Proof of the Theorem 9

It is worth noting that the estimate $\widehat{\underline{\theta}}_n$ is a solution to the $s\times(p+q+1)$ -dimensional estimating equation $\sum_{t=0}^{n-1}\frac{\partial\widehat{l_t}(\underline{\theta})}{\partial\underline{\theta}} = \underline{O}$. The estimating equation above is asymptotically equivalent (see Lemma 15 below) to $\underline{S}_n(\underline{\theta}) = \sum_{t=0}^{n-1} \frac{\partial l_t(\underline{\theta})}{\partial \underline{\theta}} = \underline{O}$, so the asymptotic distribution of $\widehat{\underline{\theta}}_n$ can be obtained from $\frac{1}{\sqrt{n}}\underline{S}_n(\underline{\theta})$. Indeed; using Taylor-series expansion of the score vector around $\underline{\theta}_0$, we obtain

$$0 = \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} \frac{\partial l_t(\widehat{\underline{\theta}}_n)}{\partial \underline{\theta}} = \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} \frac{\partial l_t(\underline{\theta}_0)}{\partial \underline{\theta}} + \left(\frac{1}{n} \sum_{t=0}^{n-1} \frac{\partial^2 l_t(\underline{\theta}_*)}{\partial \underline{\theta} \partial \underline{\theta}'} \right) \sqrt{n} \left(\widehat{\underline{\theta}}_n - \underline{\theta}_0 \right) \quad (2.10)$$

where $\underline{\theta}_*$ is between $\widehat{\underline{\theta}}_n$and $\underline{\theta}_0$. Thus we will show that

$$\frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} \frac{\partial l_t(\underline{\theta}_0)}{\partial \underline{\theta}} \rightsquigarrow N\left(\underline{O}, \frac{4\kappa}{s^2} \mathcal{I}(\underline{\theta}_0) \right)$$

and $\frac{1}{n} \sum_{t=0}^{n-1} \frac{\partial^2 l_t(\underline{\theta}_*)}{\partial \underline{\theta} \partial \underline{\theta}'} \xrightarrow{P} \frac{2}{s} \mathcal{I}(\underline{\theta}_0)$ and the result fellows from Slutsky's theorem. For this purpose, we will split the proof into several intermediate results grouped in lemmas 15 and 21 below.

Lemme 15 *If the Assumptions **A1-A6** hold*

1. $E_{\underline{\theta}_0} \left\{ \sup_{\underline{\theta} \in \vartheta(\underline{\theta}_0)} \left\| \frac{\partial l_t(\underline{\theta})}{\partial \underline{\theta}} \frac{\partial l_t(\underline{\theta})}{\partial \underline{\theta}'} \right\| \right\} < +\infty$ *and* $E_{\underline{\theta}_0} \left\{ \sup_{\underline{\theta} \in \vartheta(\underline{\theta}_0)} \left\| \frac{\partial^2 l_t(\underline{\theta})}{\partial \underline{\theta} \partial \underline{\theta}'} \right\| \right\} < +\infty$, $E_{\underline{\theta}_0} \left\{ \sup_{\underline{\theta} \in \vartheta(\underline{\theta}_0)} \left| \frac{\partial^3 l_t(\underline{\theta})}{\partial \theta_i \partial \theta_j \partial \theta_k} \right| \right\} < +\infty$ *for any neighborhood* $\vartheta(\underline{\theta}_0)$ *of* $\underline{\theta}_0$ *and for all* $i, j, k \in \{1, ..., s(p+q+1)\}$.

2.
$$\left\| \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} \left\{ \frac{\partial l_t(\underline{\theta})}{\partial \underline{\theta}} - \frac{\partial \widehat{l}_t(\underline{\theta})}{\partial \underline{\theta}} \right\} \right\| \xrightarrow{P} 0 \text{ as } n \to \infty,$$
$$\sup_{\underline{\theta} \in \vartheta(\underline{\theta}_0)} \left\| \frac{1}{n} \sum_{t=0}^{n-1} \left\{ \frac{\partial^2 l_t(\underline{\theta})}{\partial \underline{\theta} \partial \underline{\theta}'} - \frac{\partial^2 \widehat{l}_t(\underline{\theta})}{\partial \underline{\theta} \partial \underline{\theta}'} \right\} \right\| \xrightarrow{P} 0 \text{ as } n \to \infty$$

Proof. See Aknouche and Bibi [2]. ■

Lemme 16 *Under the Assumptions **A1-A6** we have*

1. $Var_{\underline{\theta}_0} \left\{ \frac{\partial l_t(\underline{\theta})}{\partial \underline{\theta}} \right\} = \frac{4\kappa}{s^2} \mathcal{I}(\underline{\theta}_0)$ *where* $\mathcal{I}(\underline{\theta}_0)$ *is positive definite matrix.*

2. $\frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} \frac{\partial l_t(\underline{\theta}_0)}{\partial \underline{\theta}} \rightsquigarrow N\left(\underline{O}, \frac{4\kappa}{s^2} \mathcal{I}(\underline{\theta}_0) \right)$

3. $\frac{1}{n}\sum_{t=0}^{n-1}\frac{\partial^2 l_t(\underline{\theta}_*)}{\partial\underline{\theta}\partial\underline{\theta}'}\overset{P}{\to}\frac{2}{s}\mathcal{I}(\underline{\theta}_0)$

Proof.

1. It is not hard to see that $E_{\underline{\theta}_0}\left\{\frac{\partial l_t(\underline{\theta}_0)}{\partial\underline{\theta}}\right\}=\underline{O}$ and $Var_{\underline{\theta}_0}\left\{\frac{\partial l_t(\underline{\theta}_0)}{\partial\underline{\theta}}\right\}=E_{\underline{\theta}_0}\left\{\frac{\partial l_t(\underline{\theta}_0)}{\partial\underline{\theta}}\frac{\partial l_t(\underline{\theta}_0)}{\partial\underline{\theta}'}\right\}=\frac{4\kappa}{s^2}\mathcal{I}(\underline{\theta}_0)$. Now, using the same arguments as in Francq and Zakoïan [20], we can show that for any $v=1,...,s$, the vth block $\mathcal{I}_v(\underline{\theta}_0)$ of $\mathcal{I}(\underline{\theta}_0)$ is positive definite matrix and thus $\mathcal{I}(\underline{\theta}_0)$ it is.

2. Notice that $E_{\underline{\theta}_0}\left\{\frac{\partial l_t(\underline{\theta})}{\partial\underline{\theta}}|\Im_{t-1}\right\}=\underline{O}$ where $\Im_t:=\sigma\{\epsilon_{t-i},i\geq 0\}$ and that $Var_{\underline{\theta}_0}\left\{\frac{\partial l_t(\underline{\theta}_0)}{\partial\underline{\theta}}\right\}$ exists. Hence for any $\underline{\lambda}\in\mathbb{R}^{s(p+q+1)}$, the sequence $\left(\left(\underline{\lambda}'\frac{\partial l_t(\underline{\theta})}{\partial\underline{\theta}},\Im_t\right)\right)_{t\in\mathbb{Z}}$ is a square integrable martingale difference. Then by Theorem 3.1 of Billingsley [12] and the Wald-Cramèr device,

$$\frac{1}{\sqrt{n}}\sum_{t=0}^{n-1}\frac{\partial l_t(\underline{\theta}_0)}{\partial\underline{\theta}}\rightsquigarrow N\left(\underline{O},\frac{4\kappa}{s^2}\mathcal{I}(\underline{\theta}_0)\right).$$

3. The convergence follows from the $a.s$ convergence of $\underline{\theta}_*$ to $\underline{\theta}_0$, $\frac{1}{n}\sum_{t=0}^{n-1}\left\{\frac{\partial^2 l_t(\underline{\theta}_*)}{\partial\underline{\theta}\partial\underline{\theta}'}-\frac{\partial^2 l_t(\underline{\theta})}{\partial\underline{\theta}\partial\underline{\theta}'}\right\}$ to O and an application of the ergodic theorem to $\frac{1}{n}\sum_{t=0}^{n-1}\frac{\partial^2 l_t(\underline{\theta})}{\partial\underline{\theta}\partial\underline{\theta}'}$.

■

2.4.2 Proof of Theorem 12

Let $\underline{Y}_n=\frac{1}{\sqrt{2n\log\log n}}\sum_{t=0}^{n-1}\left|\frac{\partial l_t(\underline{\theta})}{\partial\underline{\theta}}-\frac{\partial\widehat{l_t}(\underline{\theta})}{\partial\underline{\theta}}\right|$. By Assertion 2 of Lemma 15, it can be shown that almost surly $\underline{Y}_n\to\underline{O}$ as $n\to\infty$. In other hand from (2.10) we have

$$\sqrt{\frac{n}{2\kappa\log\log n}}\left(\widehat{\underline{\theta}}_n-\underline{\theta}_0\right) \leq \left|\frac{s}{2}\mathcal{I}^{-1}-\left(\frac{1}{n}\sum_{t=0}^{n-1}\frac{\partial^2 l_t(\underline{\theta}_*)}{\partial\underline{\theta}\partial\underline{\theta}'}\right)^{-1}\right| \times\frac{1}{\sqrt{2\kappa n\log\log n}}\sum_{t=0}^{n-1}\frac{\partial l_t(\underline{\theta}_0)}{\partial\underline{\theta}} +\left|\frac{s}{2}\mathcal{I}^{-1}-\left(\frac{1}{n}\sum_{t=0}^{n-1}\frac{\partial^2 l_t(\underline{\theta}_*)}{\partial\underline{\theta}\partial\underline{\theta}'}\right)^{-1}\right|\left|\underline{Y}_n\right| +\mathcal{I}^{-1/2}\frac{1}{\sqrt{2\kappa n\log\log n}}\sum_{t=0}^{n-1}\frac{s}{2}\mathcal{I}^{1/2}\frac{\partial\widehat{l}_t(\underline{\theta}_0)}{\partial\underline{\theta}}$$

Since almost surly, $\frac{s}{2}\mathcal{I}^{-1}-\left(\frac{1}{n}\sum_{t=0}^{n-1}\frac{\partial^2 l_t(\underline{\theta}_*)}{\partial\underline{\theta}\partial\underline{\theta}'}\right)^{-1} \to O$, $\underline{Y}_n \to \underline{O}$, then LIL for martingale difference stationary and ergodic sequence can be applied here to show that $\limsup_n \frac{1}{\sqrt{2\kappa n\log\log n}}\sum_{t=0}^{n-1}\frac{s}{2}\mathcal{I}^{1/2}\frac{\partial\widehat{l}_t(\underline{\theta}_0)}{\partial\underline{\theta}} \to \underline{1}_{(s(p+q+1))}$.

2.4.3 Proof of Theorem 13 [Consistency of LSE PARMA-PGARCH]

The proof of Theorem 13 relies on a set of intermediate results presented below. It will convenient to consider the functions $\widehat{O}_{1,n}, \widehat{O}_{2,n}$, corresponding to the SPS processes $l_{1,t}$ and $l_{2,t}$.

Lemme 17 *Under* ***A1-A3*** *and* ***A7-A8*** *almost surely*

1. $\widehat{\underline{\beta}}_n \to \underline{\beta}_0$ *as* $n\to\infty$.
2. $\left\{\exists t\in\mathbb{Z} : \epsilon_t\left(\underline{\beta}\right)=\epsilon_t\left(\underline{\beta}_0\right) \text{ and } h_t\left(\underline{\pi}\right)=h_t\left(\underline{\pi}_0\right)\right\} \Longrightarrow \underline{\pi}=\underline{\pi}_0$

Proof. The proof follows from the same arguments as in Aknouche and Bibi [2]. ■

Lemme 18 *Under* ***A1-A3*** *and* ***A7-A8*** *we have*

1. $\lim_n \sup_{\underline{\theta}\in\Theta_{\underline{\theta}}}\left|\widehat{O}_{2,n}\left(\widehat{\underline{\beta}}_n,\underline{\theta}\right)-\widehat{O}_{2,n}\left(\underline{\beta}_0,\underline{\theta}\right)\right|=0$
2. $\lim_n \sup_{\underline{\pi}\in\Theta_{\underline{\pi}}}\left|\widehat{Q}_{2,n}\left(\underline{\pi}\right)-\widehat{O}_{2,n}\left(\underline{\pi}\right)\right|=0$
3. $\sigma^2 := E_{\underline{\pi}_0}\left\{l_{2,t}\left(\underline{\beta}_0,\underline{\theta}_0\right)\right\} < E_{\underline{\pi}_0}\left\{l_{2,t}\left(\underline{\beta}_0,\underline{\theta}\right)\right\}$ *for any* $\underline{\theta}\neq\underline{\theta}_0$ *and* $\underline{\theta}_0\in\Theta_{\underline{\theta}}$.

4. *For any* $\underline{\theta}_* \in \Theta_{\underline{\theta}}$, $\underline{\theta}_* \neq \underline{\theta}_0$ *there exists a neighborhood* $V(\underline{\theta}_*) \subset \Theta_{\underline{\theta}}$ *of* $\underline{\theta}_*$ *such that* $\liminf_n \inf_{\underline{\theta} \in V(\underline{\theta}_*)} \widehat{Q}_{2,n}\left(\widehat{\underline{\beta}}_n, \underline{\theta}\right) > \sigma^2$ *a.s.*

Proof. The proof of these assertions follows essentially the same arguments as in [21], therefore we do not detail the proof. ■

Lemme 19 *Under the Conditions* ***A2*** *and* ***A7*** *we have for all* $v \in \{1, ..., s\}$ *and any* $\underline{\pi} \in \Theta$

$$\begin{aligned} h_{st+v}(\underline{\pi}) &= \alpha_0(v) + \sum_{j=1}^{\infty} \alpha_j(v) \epsilon^2_{st+v-j}(\underline{\beta}), \\ \epsilon_{st+v}(\underline{\beta}) &= (X_{st+v} - \mu(v)) + \sum_{j=1}^{\infty} c_j(v)(X_{st+v-j} - \mu(v-j)). \end{aligned}$$

Moreover, $\forall \underline{\beta} \in \Theta_{\underline{\beta}}$, *the functions* $\epsilon_t(.)$ *and* $h_t(\underline{\beta}, .)$ *are continuously differentiable with*

$$\begin{aligned} \frac{\partial h_{st+v}(\underline{\beta}, \underline{\theta})}{\partial \theta_j(v)} &= \widetilde{\alpha}_0(v) + \sum_{j=1}^{\infty} \widetilde{\alpha}_j(v) \epsilon^2_{st+v-j}(\underline{\beta}), \\ \frac{\partial \epsilon_{st+v}(\underline{\beta})}{\partial \beta_j(v)} &= \sum_{j=1}^{\infty} \widetilde{c}_j(v)(X_{st+v-j} - \mu(v-j)). \end{aligned}$$

where the weights $\alpha_j(v), c_j(v), \widetilde{\alpha}_j(v)$ *and* $\widetilde{c}_j(v)$, *satisfy* $\max_{1 \le v \le s} \alpha_j(v) = O(\rho^j)$, $\max_{1 \le v \le s} \widetilde{\alpha}_j(v) = O(\rho^j)$, $\max_{1 \le v \le s} |c_j(v)| = O(\rho^j)$, $\max_{1 \le v \le s} |c_j(v)| = O(\rho^j)$ *with* $\rho \in]0, 1[$.

Proof. This is straightforward consequence from the Remark 3 and the invertibility assumption in **A7**. ■

In order to complete the proof of Theorem 13, let $\mathcal{V}(\underline{\theta}_0)$ be any neighborhood of $\underline{\theta}_0$. By the compactness, $\Theta_{\underline{\theta}}$ is covered by $\mathcal{V}(\underline{\theta}_0), \mathcal{V}(\underline{\theta}_1), ..., \mathcal{V}(\underline{\theta}_k)$ where $\underline{\theta}_j \in \Theta_{\underline{\theta}} \backslash \mathcal{V}(\underline{\theta}_0)$ and the $\mathcal{V}(\underline{\theta}_j)$, $1 \le j \le k$ are defined in Lemma 18, Assertion 4. Using Lemma 18, Assertions 1, 2 and 4 and the periodic ergodicity we have almost surely $\lim_{n \to \infty} \widehat{Q}_{2,n}\left(\widehat{\underline{\beta}}_n, \underline{\theta}\right) = \sigma^2$ and

$$\inf_{\underline{\theta} \in \Theta_{\underline{\theta}}} \widehat{Q}_{2,n}\left(\widehat{\underline{\beta}}_n, \underline{\theta}\right) = \min_{0 \le i \le k} \inf_{\underline{\theta} \in \mathcal{V}(\underline{\theta}_i) \cap \Theta_{\underline{\theta}}} \widehat{Q}_{2,n}\left(\widehat{\underline{\beta}}_n, \underline{\theta}\right) = \inf_{\underline{\theta} \in \mathcal{V}(\underline{\theta}_0) \cap \Theta_{\underline{\theta}}} \widehat{Q}_{2,n}\left(\widehat{\underline{\beta}}_n, \underline{\theta}\right)$$

for n large enough. This proves that $\widehat{\underline{\theta}}_n \in \mathcal{V}(\underline{\theta}_0)$, $a.s$ for n large enough, giving the required consistency result.

2.4.4 Proof of Theorem 14[Asymptotic normality of LSE PARMA-PGARCH]

The proof rests classically on the Taylor-series expansion of score vectors around the true parameter values

$$\begin{aligned}
0 &= \frac{1}{\sqrt{n}}\sum_{t=0}^{n-1}\frac{\partial l_{1,t}(\widehat{\underline{\beta}}_n)}{\partial\underline{\beta}} = \frac{1}{\sqrt{n}}\sum_{t=0}^{n-1}\frac{\partial l_{1,t}(\underline{\beta}_0)}{\partial\underline{\beta}} + \left(\frac{1}{n}\sum_{t=0}^{n-1}\frac{\partial^2 l_{1,t}(\underline{\beta}_*)}{\partial\underline{\beta}\partial\underline{\beta}'}\right)\sqrt{n}\left(\widehat{\underline{\beta}}_n-\underline{\beta}_0\right)\\
0 &= \frac{1}{\sqrt{n}}\sum_{t=0}^{n-1}\frac{\partial l_{2,t}(\widehat{\underline{\beta}}_n,\widehat{\underline{\theta}}_n)}{\partial\underline{\theta}} = \frac{1}{\sqrt{n}}\sum_{t=0}^{n-1}\frac{\partial l_{2,t}(\widehat{\underline{\beta}}_n,\underline{\theta}_0)}{\partial\underline{\theta}}\\
&\quad+\left(\frac{1}{n}\sum_{t=0}^{n-1}\frac{\partial^2 l_{2,t}(\widehat{\underline{\beta}}_n,\underline{\theta}_*)}{\partial\underline{\theta}\partial\underline{\theta}'}\right)\sqrt{n}\left(\widehat{\underline{\theta}}_n-\underline{\theta}_0\right)\\
&= \frac{1}{\sqrt{n}}\sum_{t=0}^{n-1}\frac{\partial l_{2,t}(\underline{\beta}_0,\underline{\theta}_0)}{\partial\underline{\theta}} + \left(\frac{1}{n}\sum_{t=0}^{n-1}\frac{\partial^2 l_{2,t}(\underline{\beta}_{**},\underline{\theta}_0)}{\partial\underline{\theta}\partial\underline{\beta}'}\right)\sqrt{n}\left(\widehat{\underline{\beta}}_n-\underline{\beta}_0\right)\\
&\quad+\left(\frac{1}{n}\sum_{t=0}^{n-1}\frac{\partial^2 l_{2,t}(\widehat{\underline{\beta}}_n,\underline{\theta}_*)}{\partial\underline{\theta}\partial\underline{\theta}'}\right)\sqrt{n}\left(\widehat{\underline{\theta}}_n-\underline{\theta}_0\right).
\end{aligned}$$

The above equations leads to $-\frac{1}{\sqrt{n}}\sum_{t=0}^{n-1}\frac{\partial \underline{l}_t(\underline{\pi}_0)}{\partial\underline{\pi}} = \frac{1}{n}\sum_{t=0}^{n-1}\frac{\partial^2 \underline{l}_t(\underline{\pi}_*)}{\partial\underline{\pi}\partial\underline{\pi}'}\sqrt{n}\left(\widehat{\underline{\pi}}_n-\underline{\pi}_0\right)$ where $\frac{\partial \underline{l}_t(\underline{\pi}_0)}{\partial\underline{\pi}} := \left(\frac{\partial l_{1,t}(\underline{\beta}_0)}{\partial\underline{\beta}'},\frac{\partial l_{2,t}(\underline{\pi}_0)}{\partial\underline{\theta}'}\right)'$ and where $\underline{\beta}_*$'s (resp. $\underline{\beta}_{**}$, $\underline{\theta}_*$, $\underline{\pi}_*$) are between $\widehat{\underline{\beta}}_n$ and $\underline{\beta}_0$, (resp. $\widehat{\underline{\beta}}_n$ and $\underline{\beta}_0$, $\widehat{\underline{\theta}}_n$ and $\underline{\theta}_0$ and between $\widehat{\underline{\pi}}_n$ and $\underline{\pi}_0$). Thus we will show that

$$\frac{1}{\sqrt{n}}\sum_{t=0}^{n-1}\frac{\partial \underline{l}_t(\underline{\pi}_0)}{\partial\underline{\pi}} \rightsquigarrow N\left(\underline{O},I\right) \text{ and } \frac{1}{n}\sum_{t=0}^{n-1}\frac{\partial^2 \underline{l}_t(\underline{\pi}_*)}{\partial\underline{\pi}\partial\underline{\pi}'} \longrightarrow J \text{ in probability with}$$

$I=\begin{pmatrix} I_{11} & I_{12}\\ I_{21} & I_{22}\end{pmatrix}$, $J=\begin{pmatrix} J_{11} & O\\ J_{21} & J_{22}\end{pmatrix}$ where the matrices $(I_{ij})_{1\le i,j\le 2}$ and $(J_{ij})_{1\le i,j\le 2}$ are given in Theorem 14. The theorem will straightforwardly follow. For this purpose, we show an analogue Lemma 15 for $\frac{\partial \underline{l}_t(\underline{\pi}_0)}{\partial\underline{\pi}}$, i.e.,

Lemme 20 *If the Assumptions **A1-A10** hold*

1. *For any* $\underline{\pi}\in\Theta$, *the random vectors* $\frac{\partial}{\partial\beta}l_{1,n}\left(\underline{\beta}\right)$, $\frac{\partial}{\partial\theta}l_{2,n}\left(\underline{\pi}\right)$ *exist and belong to* $\mathbb{L}_2$
2. $E_{\underline{\theta}_0}\left\{\sup_{\underline{\pi}\in\vartheta(\underline{\pi}_0)}\left\|\frac{\partial \underline{l}_t(\underline{\pi})}{\partial\underline{\pi}}\frac{\partial \underline{l}_t(\underline{\pi})}{\partial\underline{\pi}'}\right\|\right\}<+\infty$ *and* $E_{\underline{\theta}_0}\left\{\sup_{\underline{\pi}\in\vartheta(\underline{\pi}_0)}\left\|\frac{\partial^2 \underline{l}_t(\underline{\pi})}{\partial\underline{\pi}\partial\underline{\pi}'}\right\|\right\}<+\infty$, $E_{\underline{\theta}_0}\left\{\sup_{\underline{\pi}\in\vartheta(\underline{\pi}_0)}\left|\frac{\partial^3 \underline{l}_t(\underline{\pi})}{\partial\pi_i\partial\pi_j\partial\pi_k}\right|\right\}<+\infty$ *for any neighborhood* $\vartheta\left(\underline{\pi}_0\right)$ *of* $\underline{\pi}_0$ *and for all* $i,j,k\in\{1,...,s(p+P+q+Q+1)\}$.

3.

$$\left\| \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} \left\{ \frac{\partial \underline{l}_t(\underline{\pi})}{\partial \underline{\pi}} - \frac{\partial \widehat{\underline{l}}_t(\underline{\pi})}{\partial \underline{\pi}} \right\} \right\| \xrightarrow{P} 0 \ as\ n \ \rightarrow \ \infty,$$

$$\sup_{\underline{\theta} \in \vartheta(\underline{\theta}_0)} \left\| \frac{1}{n} \sum_{t=0}^{n-1} \left\{ \frac{\partial^2 \underline{l}_t(\underline{\pi})}{\partial \underline{\pi} \partial \underline{\pi}'} - \frac{\partial^2 \widehat{\underline{l}}_t(\underline{\pi})}{\partial \underline{\pi} \partial \underline{\pi}'} \right\} \right\| \xrightarrow{P} 0 \ as\ n \ \rightarrow \ \infty$$

Proof.

1. Under the Assumption **A9**, it follow that $E\{X_t^4\} < +\infty$. By Lemma 19 and Cauchy-Schwartz inequality we can see that

$$\frac{\partial \epsilon^2_{st+v}\left(\underline{\beta}\right)}{\partial \underline{\beta}} = 2\epsilon_{st+v}\left(\underline{\beta}\right) \frac{\partial \epsilon_{st+v}\left(\underline{\beta}\right)}{\partial \underline{\beta}} \text{ and } \frac{\partial \eta^2_{st+v}\left(\underline{\pi}\right)}{\partial \underline{\theta}} = 2\eta_{st+v}\left(\underline{\pi}\right) \frac{\partial \eta_{st+v}\left(\underline{\pi}\right)}{\partial \underline{\theta}}$$

belong to $\mathbb{L}_2$.

2. The Assertions 2 and 3 follows similarly as proving Lemma 15.

■

Lemme 21 *Under **A1-A10**, we have*

1. $\frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} \frac{\partial \underline{l}_t(\underline{\pi}_0)}{\partial \underline{\pi}} \rightsquigarrow N\left(\underline{O}, I\left(\underline{\pi}_0\right)\right)$, *where the sub-matrices* I_{11}, I_{12} *and* I_{22} *exist and are strictly positive definite.*

2. $\frac{1}{n} \sum_{t=0}^{n-1} \frac{\partial^2 \underline{l}_t\left(\underline{\pi}_*\right)}{\partial \underline{\pi} \partial \underline{\pi}'} \xrightarrow{P} J\left(\underline{\pi}_0\right)$.

Proof.

1. We will apply a central limit theorem for the martingale difference. Let $\Im_t^{(\epsilon)} := \sigma\left\{\epsilon_{t-i}, i \geq 0\right\}$, then by Assumption **A7**, we have $\Im_t^{(\epsilon)} = \Im_t^{(X)}$. Notice that $E_{\underline{\beta}_0}\left\{\frac{\partial l_{1,t}(\underline{\beta}_0)}{\partial \underline{\beta}} | \Im_{t-1}^{(X)}\right\} = \underline{O}$, $E_{\underline{\pi}_0}\left\{\frac{\partial l_{2,t}(\underline{\pi}_0)}{\partial \underline{\theta}} | \Im_{t-1}^{(X)}\right\} = \underline{O}$ and that in view of Assertion 1 in Lemma 20, $Var_{\underline{\beta}_0}\left\{\frac{\partial l_{1,t}(\underline{\beta}_0)}{\partial \underline{\beta}}\right\}$ and $Var_{\underline{\pi}_0}\left\{\frac{\partial l_{2,t}(\underline{\pi}_0)}{\partial \underline{\theta}}\right\}$ exists and not singular matrices. Hence, for any $\left(\underline{\lambda}', \underline{\mu}'\right)' \in \mathbb{R}^{s(P+Q+1)} \times \mathbb{R}^{s(p+q+1)}$, the sequence $\left\{\left(\underline{\lambda}', \underline{\mu}'\right) \frac{\partial \underline{l}_t\left(\underline{\pi}\right)}{\partial \underline{\pi}}, \Im_t^{(X)}\right\}_t$ is a square integrable SPS matringale diference. The central limite theorem of Billingsley [12] and the Wold-Cramèr devoce allow to derive the asymptotic normality result.

2. The convergence follows from the $a.s$ convergence of $\underline{\pi}_*$ to $\underline{\pi}_0$, an application of the ergodic theorem to $\frac{1}{n}\sum_{t=0}^{n-1} \frac{\partial^2 \underline{l}_t(\underline{\pi}_*)}{\partial \underline{\pi} \partial \underline{\pi}'}$ and the fact that almost surely as $n \to \infty$

$$\left\| \frac{1}{n} \sum_{t=0}^{n-1} \left(\frac{\partial^2 l_{1,t}(\underline{\beta}_*)}{\partial \underline{\beta} \partial \underline{\beta}'} - \frac{\partial^2 l_{1,t}(\underline{\beta}_0)}{\partial \underline{\beta} \partial \underline{\beta}'} \right) \right\| \to 0,$$

$$\left\| \frac{1}{n} \sum_{t=0}^{n-1} \left(\frac{\partial^2 l_{2,t}(\underline{\beta}_{**}, \underline{\theta}_0)}{\partial \underline{\theta} \partial \underline{\beta}'} - \frac{\partial^2 l_{2,t}(\underline{\pi}_0)}{\partial \underline{\theta} \partial \underline{\beta}'} \right) \right\| \to 0,$$

$$\left\| \frac{1}{n} \sum_{t=0}^{n-1} \left(\frac{\partial^2 l_{2,t}(\widehat{\underline{\beta}}_n, \underline{\theta}_*)}{\partial \underline{\theta} \partial \underline{\theta}'} - \frac{\partial^2 l_{2,t}(\underline{\pi}_0)}{\partial \underline{\theta} \partial \underline{\theta}'} \right) \right\| \to 0.$$

■

Chapter 3

CLS approach for $PGARCH$ models

Abstract: This chapter studies the strong consistency and asymptotic normality (CAN) of the conditional least squares estimates (LSE) for periodic $GARCH$ ($PGARCH$) models with martingale difference centered squared innovations. The approach is extended to the $PARMA-PGARCH$ models. The results are obtained under mild conditions, in particular, no restrictions on the conditional mean are imposed. Our proofs closely follow those in Francq and Zakoïan [20] for independent and identically distributed innovations.

3.1 Introduction

In the process of attempting to model the conditional variance in financial time series $(\epsilon_n)_{n\in\mathbb{Z}}$ exhibiting structural changes, Bollerslev and Ghysels [15] have proposed a $GARCH\,(p,q)$ model with time-varying coefficient which has the form of

$$\forall n \in \mathbb{Z}\text{: } \epsilon_n = e_n\sqrt{h_n} \text{ and } h_n = a_0(n)+\sum\nolimits_{i=1}^{q} a_i(n)\epsilon_{n-i}^2+\sum\nolimits_{j=1}^{p} b_j(n)h_{n-j} \quad (3.1)$$

where $(e_n)_{n\in\mathbb{Z}}$ is a sequence of random variables (its characteristics are specified below), the coefficients $(a_i(n))_{0\leq i\leq q}$ and $(b_j(n))_{1\leq j\leq p}$ are positive except that $a_0(n) > 0$. The Model (3.1) is called periodic $GARCH$ ($PGARCH$) when the functions $(a_i(n))_{0\leq i\leq q}$ and $(b_j(n))_{1\leq j\leq p}$ are periodic in n with period $s > 0$, *i.e.*, $a_i(n) = a_i(n+sk)$ and $b_j(n) = b_j(n+sk)$ for all integers $n, k \in \mathbb{Z}$. So, by setting

$n = st + v$, $1 \le v \le s$, Model (3.1) may be equivalently written as

$$\epsilon_{st+v} = e_{st+v}\sqrt{h_{st+v}} \text{ and } h_{st+v} = a_0(v) + \sum_{i=1}^{q} a_i(v)\epsilon^2_{st+v-i} + \sum_{j=1}^{p} b_j(v)h_{st+v-j}. \tag{3.2}$$

The *PGARCH* models are generally nonstationary but are stationary within each period. They are becoming an appealing tool for investigating both volatility and distinct seasonal patterns and continue to gain popularity in various disciplines (see, e.g., Bibi and Lescheb [10],[11] and the references therein for further discussions). Unfortunately, its probabilistic and statistical properties remain unexplored compared with respect to the other structures (for instance standard and Markovian switching *GARCH* models). The main raison is certainly, that the lack of stationarity and thus the ergodicity in such models result in enormous technical difficulties. Since the seminal paper by Pagano [41], with periodic coefficients, Model (3.2) may be connected with multivariate model with time-invariant coefficient. More precisely $\underline{\epsilon}_t = (\epsilon_{st+1}, ..., \epsilon_{st+s})'$ is a $s-$variate $GARCH\,(p^*, q^*)$ model in the sense that

$$\underline{\epsilon}_t = \{diag\,\{\underline{h}_t\}\}^{\frac{1}{2}}\,\underline{e}_t \text{ and } \underline{h}_t = \underline{a}_0 + \sum_{i=0}^{q^*} A_i\underline{\epsilon}^2_{t-i} + \sum_{j=0}^{p^*} B_j\underline{h}_{t-j} \tag{3.3}$$

where $\underline{\epsilon}^2_t = \left(\epsilon^2_{st+1}, ..., \epsilon^2_{st+s}\right)'$, $\underline{h}_t = (h_{st+1}, ..., h_{st+s})'$ and where $\underline{e}_t = (e_{st+1}, ..., e_{st+s})'$. The model orders in (3.3) are $p^* = \left[\frac{p}{s}\right]$ and $q^* = \left[\frac{q}{s}\right]$ where $[x]$ denotes the smallest integer greater than or equal to x. The $s \times s$ matrices $(A_i)_{0\le i\le q^*}$ and $(B_i)_{0\le i\le p^*}$ are computed as follows (see Basawa and Lund [4]). A_0, B_0 have $(i,j)\,th$ entries $(B_0)_{i,j} = b_{i-j}(i)\mathbb{I}_{\{i>j\}}$, $(A_0)_{i,j} = a_{i-j}(i)\mathbb{I}_{\{i>j\}}$ and $(B_m)_{i,j} = b_{ms+i-j}(i)$ for $1 \le m \le p^*$, $(A_m)_{i,j} = a_{ms+i-j}(i)$ for $1 \le m \le q^*$ and the intercept vector $\underline{a}_0 = (a_0(1), ..., a_0(s))'$. In the sequel, $I_{(k)}$ denotes the identity matrix of order k, O (resp. $\underline{O}$) denotes the matrix (resp. vector) whose entries are zeros. The norm of a matrix $M = (m_{ij})$ is defined by $\|M\|$. This chapter investigates the strong consistency and asymptotic normality of the least squares estimator (*LSE*) in *PGARCH* and extends those asymptotic results to *PARMA* $-$ *PGARCH* models.

3.2 Conditional least squares estimation for *PGARCH* models

Consider the *PGARCH* model (3.2) described with the vector of parameters $\underline{\theta} = (\underline{\theta}'(1), ..., \underline{\theta}'(s))'$ where $\underline{\theta}(v) = (a_0(v), a_1(v), ..., a_q(v), b_1(v), ..., b_p(v))'$, $v =$

$1, ..., s$. The vector $\underline{\theta}$ belongs to a parameter space $\Theta_{\underline{\theta}} := \{\underline{\theta} : \underline{\theta} \in \left(]0, \infty[\times [0, \infty[^{(p+q)}\right)^s\}$. The orders p and q and the period s are supposed to be known, whereas the true parameter value $\underline{\theta}_0$ is unknown. Let $\{\epsilon_1, \epsilon_2, ..., \epsilon_N\}$ be a realization of length $N = sn$ to estimate the parameter $\underline{\theta}$. Conditionally on initial values $\epsilon_0, \epsilon_{-1}, ..., \epsilon_{1-q}$, $\widehat{h}_0, \widehat{h}_{-1}, ..., \widehat{h}_{1-p}$ properly chosen, the LSE of $\underline{\theta}$ is defined as any measurable solution $\widehat{\underline{\theta}}_n$ of

$$\widehat{\underline{\theta}}_n = Arg \min_{\underline{\theta} \in \Theta_{\underline{\theta}}} \widehat{Q}_n(\underline{\theta}), \ \widehat{Q}_n(\underline{\theta}) := \frac{1}{n} \sum_{t=0}^{n-1} \widehat{l}_t(\underline{\theta}), \ \widehat{l}_t(\underline{\theta}) := \frac{1}{s} \sum_{v=1}^{s} \widehat{\eta}_{st+v}^2(\underline{\theta}) \tag{3.4}$$

with $\widehat{\eta}_{st+v}(\underline{\theta}) = \epsilon_{st+v}^2 - \widehat{h}_{st+v}(\underline{\theta})$ where $\widehat{h}_{st+v}(\theta)$ are defined recursively by $\widehat{h}_{st+v}(\underline{\theta}) = a_0(v) + \sum_{i=1}^{q} a_i(v)\epsilon_{st+v-i}^2 + \sum_{j=1}^{p} b_j(v)\widehat{h}_{st+v-j}(\underline{\theta})$. For the strong consistency of $\widehat{\underline{\theta}}_n$ we need the following regularity conditions. First define the local polynomials

$$\mathcal{A}_v(z) = \sum_{j=1}^{q} a_{0j}(v) z^j, \mathcal{B}_v(z) = 1 - \sum_{j=1}^{p} b_{0j}(v) z^j$$

and assume that

A1. $\underline{\theta}_0 \in \Theta_{\underline{\theta}}$ and $\Theta_{\underline{\theta}}$ is compact.

A2. $(e_n)_{n\in\mathbb{Z}}$ is a sequence of of strictly stationary and ergodic random variables satisfying almost surely $(a.s)$ $E\left\{e_n^2 | \Im_{n-1}^{(\epsilon)}\right\} = 1$ here $\Im_n^{(\epsilon)}$ refers to the $\sigma-$field generated by $\{\epsilon_t, t \leq n\}$. Moreover, $(e_t^2)_{t\in\mathbb{Z}}$ has a non-degenerate distribution.

A3. $(\epsilon_t)_{t\in\mathbb{Z}}$ is strictly periodically stationary (SPS) and periodically ergodic process in the sense that $(\underline{\epsilon}_t)_{t\in\mathbb{Z}}$ is strictly stationary and ergodic process with $E\{\epsilon_t^4\} < \infty$.

A4. for all $v \in \{1, ..., s\}$ and $\underline{\theta} \in \Theta_{\underline{\theta}}$, the local polynomials $\mathcal{A}_v(z)$ and $\mathcal{B}_v(z)$ have no common roots and the polynomial $\det\left(I_{(s)} - \sum_{j=0}^{p^*} B_j z^j\right)$ has its roots outside the unit circle. Moreover, $\mathcal{A}_v(1) \neq 0$ and $a_{0q}(v) + b_{0p}(v) \neq 0$.

Noting that if $\det\left(I_{(s)} - \sum_{j=0}^{\max(p^*+q^*)} (A_j + B_j) z^j\right)$ has its roots outside the unit circle, then Equation (3.3) has a strict stationarity, $\Im_n^{(e)}-$measurable, ergodic solution and $\det\left(I_{(s)} - \sum_{j=0}^{p^*} B_j z^j\right) \neq 0$ for all z such that $|z| \leq 1$. The last condition is imposed in order to obtain $h_t(\underline{\theta})$ as a causal solution of $\{\epsilon_t, \epsilon_{t-1}, ...\}$, i.e., $h_{st+v}(\underline{\theta}) = \alpha_0(v) + \sum_{j=1}^{\infty} \alpha_j(v)\epsilon_{st+v-j}^2$ for all $v \in \{1, ..., s\}$ in which the weights $\alpha_j(v)$ satisfy $\max_{1\leq v\leq s} \alpha_j(v) = O(\rho^j)$ with $\rho \in]0, 1[$.

Theorem 1 *Under **A1-A4**, almost surely* $\widehat{\underline{\theta}}_n \to \underline{\theta}_0$ *as* $n \to \infty$.

In order to establish the asymptotic normality of $LSE - PGARCH$, let $\kappa_t := E\left\{e_t^4|\Im_{t-1}^{(\epsilon)}\right\}$ and consider the additional assumptions

A5. $\underline{\theta}_0 \in \overset{o}{\Theta}_{\underline{\theta}}$ where $\overset{o}{\Theta}_{\underline{\theta}}$ denotes the interior of $\Theta_{\underline{\theta}}$.

A6. $E\left\{|e_t|^{4(1+\delta)}\right\} < \infty$ for some $\delta > 0$ and $E\left\{\epsilon_t^8\right\} < \infty$.

Theorem 2 *Under **A1-A6**,* $\sqrt{n}\left(\widehat{\underline{\theta}}_n - \underline{\theta}_0\right) \rightsquigarrow \mathcal{N}\left(\underline{O}, \mathcal{J}^{-1}\mathcal{I}\mathcal{J}^{-1}\right)$ *where* $\mathcal{J}^{-1} := diag\left\{\mathcal{J}_l^{-1}, l = 1, ..., s\right\}$, $\mathcal{I} := diag\left\{\mathcal{I}_l, l = 1, ..., s\right\}$ *and each block matrix is given by*

$$\begin{aligned}
\mathcal{J}_l &: = \sum_{v=1}^{s} E_{\underline{\theta}_0}\left\{\frac{\partial h_{st+v}(\underline{\theta})}{\partial \underline{\theta}(l)}\frac{\partial h_{st+v}(\underline{\theta})}{\partial \underline{\theta}(l)'}\right\}, \\
\mathcal{I}_l &: = \sum_{v=1}^{s} E_{\underline{\theta}_0}\left\{(\kappa_{st+v} - 1)\, h_{st+v}^2(\underline{\theta})\frac{\partial h_{st+v}(\underline{\theta})}{\partial \underline{\theta}(l)}\frac{\partial h_{st+v}(\underline{\theta})}{\partial \underline{\theta}(l)'}\right\}.
\end{aligned}$$

Moreover, under constant conditional kurtosis , i.e., $\kappa_t := E\left\{e_t^4|\Im_{t-1}^{(\epsilon)}\right\} = \kappa$, *then* $\sqrt{n}\left(\widehat{\underline{\theta}}_n - \underline{\theta}_0\right) \rightsquigarrow \mathcal{N}\left(\underline{O}, (\kappa - 1)\,\mathcal{J}^{-1}\right)$.

The LSE is not efficient due to the conditional heteroskedasticity of the innovations. To designe a more efficient estimators of $\underline{\theta}$, we weight approprietely the non linear innovations $\widehat{\eta}_t(\underline{\theta})$. Consider therefore

$$\widehat{l}_t^{(\tau)}(\underline{\theta}) := \frac{1}{s}\sum_{v=1}^{s} \tau_{st+v}\widehat{\eta}_{st+v}^2(\underline{\theta})$$

where $\tau := (\tau_t)_t$ is a sequence of positive weights and τ_t is $\Im_{t-1}^{(\epsilon)}-$measurable. Hence it is easy to show that the asymptotic variance of weighted LSE is minimal when $\tau_t = \sigma_t^{-2}$ where $\sigma_t^2 = Var_{\underline{\theta}_0}\left\{\widehat{\eta}_t|\Im_{t-1}^{(\epsilon)}\right\}$. For this asymptotically optimal sequence of weights, the corresponding estimator is called the generalized least square (GLS) estimator denoted by $\widehat{\underline{\theta}}_n^G$. In most of practical situations, σ_t is unknown and depends on a nuissance parameters

3.3 Estimation of $PARMA-PGARCH$ processes

Consider a set of observations $\{X_1, ..., X_N; N = ns\}$ obtained from a centred $PARMA(P, Q) - PGARCH(p, q)$ process $(X_t, t \in \mathbb{Z})$ satisfying

$$\begin{cases} X_t = \sum_{i=1}^{P}\phi_i(t)X_{t-i} + \epsilon_t - \sum_{j=1}^{Q}\varphi_j(t)\epsilon_{t-j} \\ \epsilon_t = \sqrt{h_t}e_t, h_t = a_0(t) + \sum_{i=1}^{q}a_i(t)\epsilon_{t-i}^2 + \sum_{j=1}^{p}b_j(t)h_{t-j} \end{cases} \tag{3.5}$$

the coefficients $(\phi_i(t))_{1\leq i\leq P}$ and $\left(\varphi_j(t)\right)_{1\leq j\leq Q}$ are periodic in t with period s. Assume that the process $(X_t, t\in\mathbb{Z})$ is described by a vector of parameters of interest $\underline{\pi}:=\left(\underline{\beta}',\underline{\theta}'\right)'$ where $\underline{\beta}=\left(\underline{\beta}'(1),...,\underline{\beta}'(s)\right)'$ with $\underline{\beta}(v):=(\phi_1(v),...,\phi_P(v),$ $\varphi_1(v),...,\varphi_Q(v))'$, $1\leq v\leq s$ and the parameter space is $\Theta_{\underline{\pi}}\subset\Theta_{\underline{\beta}}\times\Theta_{\underline{\theta}}$ where $\Theta_{\underline{\beta}}:=\mathbb{R}^{s(P+Q)}$. The orders P, Q, p and q and the period s are supposed to be known, unlike the true parameter value $\underline{\pi}_0=\left(\underline{\beta}_0',\underline{\theta}_0'\right)'$ is unknown. The corresponding vectorial version is

$$\begin{cases} \underline{X}_t=\sum_{i=0}^{P^*}\Phi_i\underline{X}_{t-i}+\underline{\epsilon}_t-\sum_{j=0}^{Q^*}\Psi_j\underline{\epsilon}_{t-j} \\ \underline{\epsilon}_t=\left\{diag\left\{\underline{h}_t\right\}\right\}^{\frac{1}{2}}\underline{e}_t \text{ and } \underline{h}_t=\underline{a}_0+\sum_{i=0}^{q^*}A_i\underline{\epsilon}_{t-i}^2+\sum_{j=0}^{p^*}B_j\underline{h}_{t-j}\end{cases}$$

where the matrices $(\Phi_i)_{0\leq i\leq P^*}$, $(\Psi_i)_{0\leq i\leq Q^*}$ and the orders P^*, Q^* may be computed as for $PGARCH(p,q)$. Conditionally on initial values $X_0,...,X_{1-P-(q-Q)}$, $\widetilde{\epsilon}_{-(q-Q)}$, $...,\widetilde{\epsilon}_{-1-q}$, $\widetilde{h}_0,...,\widetilde{h}_{1-p}$ properly chosen (cf. Aknouche and Bibi [2]), the sequence of random vectors $\widehat{\underline{\pi}}_n=\left(\widehat{\underline{\beta}}_n',\widehat{\underline{\theta}}_n'\right)'$ is called two stages least squares estimator if it satisfies, almost surely

$$\widehat{\underline{\beta}}_n=Arg\min_{\underline{\beta}\in\Theta_{\underline{\beta}}}\widehat{Q}_{1,n}\left(\underline{\beta}\right),\ \widehat{\underline{\pi}}_n:=Arg\min_{\underline{\theta}\in\Theta_{\underline{\theta}}}\widehat{Q}_{2,n}\left(\widehat{\underline{\beta}}_n,\underline{\theta}\right)$$

where $\widehat{Q}_{1,n}\left(\underline{\beta}\right):=\dfrac{1}{n}\sum_{t=0}^{n-1}\widehat{l}_{1,t}\left(\underline{\beta}\right)$ with $\widehat{l}_{1,t}\left(\underline{\beta}\right):=\dfrac{1}{s}\sum_{v=1}^{s}\widehat{\epsilon}_{st+v}^2\left(\underline{\beta}\right)$ and where $\widehat{Q}_{2,n}\left(\underline{\pi}\right):=\dfrac{1}{n}\sum_{t=0}^{n-1}\widehat{l}_{2,t}\left(\underline{\pi}\right)$ with $\widehat{l}_{2,t}\left(\underline{\pi}\right):=\dfrac{1}{s}\sum_{v=1}^{s}\widehat{\eta}_{st+v}^2\left(\underline{\pi}\right)$. For $v=1,...,s$, consider the local polynomials $\Phi_v(z)=1-\sum_{i=1}^{P}\phi_{0i}(v)z^i$, $\Psi_v(z)=1-\sum_{i=1}^{Q}\varphi_{0i}(v)z^i$ and we introduce the following conditions

A7. $(X_t)_{t\in\mathbb{Z}}$ is strictly periodically stationary (SPS) and periodically invertible process in the sense that $\left(\underline{X}_t\right)_{t\in\mathbb{Z}}$ is strictly stationary and invertible process.

A8. The polynomials $\Phi_v(z)$ and $\Psi_v(z)$ have its roots outside the unit circle and no common roots with $\phi_{0P}(v)\neq 0$ or $\varphi_{0Q}(v)\neq 0$ for all $v=1,...,s$.

Noting that if det $\left(I_{(s)}-\sum_{j=0}^{P^*}\Phi_jz^j\right)$ has its roots outside the unit circle, then $(X_t)_{t\in\mathbb{Z}}$ is SPS. Moreover, if det $\left(I_{(s)}-\sum_{j=0}^{Q^*}\Psi_jz^j\right)$ has its roots outside the unit circle, then $(X_t)_{t\in\mathbb{Z}}$ is invertible. However, under **A8.**, it follows that $(X_t)_{t\in\mathbb{Z}}$ and $(\epsilon_t)_{t\in\mathbb{Z}}$ can be related through the infinite order moving average and autoregressive expansions

$$X_{st+v}=\sum_{i=0}^{\infty}\alpha_i\left(v\right)\epsilon_{st+v-i}\text{ and }\epsilon_{st+v}=\sum_{i=0}^{\infty}\beta_i\left(v\right)X_{st+v-i} \tag{3.6}$$

In (3.6), the weights $\alpha_i(v)$ and $\beta_i(v)$ satisfy

$$\sup_{1\le v\le s}|\alpha_i(v)| = O(\rho^i) \text{ and } \sup_{1\le v\le s}|\beta_i(v)| = O(\rho^i) \text{ with } 0<\rho<1.$$

Theorem 3 *Let* $(X_t)_{t\in\mathbb{Z}}$ *be a* SPS $PARMA$ *process satisfying* (3.5). *Then under **A1-A4** and **A7-A8**, almost surely* $\widehat{\underline{\pi}}_n \to \underline{\pi}_0$ *as* $n\to\infty$.

As in the standard $ARMA-GARCH$ case (cf. [21]) we will prove asymptotic normality of $\widehat{\underline{\pi}}_n$ under the fourth order moment condition on the $(\epsilon_t)_{t\in\mathbb{Z}}$.

A9. $E_{\underline{\pi}}\{\epsilon_t^4\} < +\infty$.

A10. $\underline{\pi}_0$ is in the interior of $\Theta_{\underline{\pi}}$.

Now, we are able to derive the limit distribution of $\widehat{\underline{\pi}}_n$.

Theorem 4 *Let* $(X_t)_{t\in\mathbb{Z}}$ *be a* SPS $PARMA$ *process satisfying* (3.5). *Then under **A1-A10** we have*

$$\sqrt{n}\left(\widehat{\underline{\pi}}_n - \underline{\pi}_0\right) \rightsquigarrow \mathcal{N}\left(\begin{pmatrix} \underline{O} \\ \underline{O} \end{pmatrix}, \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}\right)$$

where $V_{11} = J_{11}^{-1}I_{11}J_{11}^{-1}$, $V_{12} = V_{21}' = J_{22}^{-1}\left(I_{21} + J_{21}J_{11}^{-1}I_{11}\right)J_{11}^{-1}$, $V_{22} = J_{22}^{-1}(I_{22} + J_{21}J_{11}^{-1}I_{11}J_{11}^{-1}J_{12} - I_{21}J_{11}^{-1}J_{12} - J_{21}J_{11}^{-1}I_{12})J_{22}^{-1}$ *with*

$$\begin{aligned}
I_{11} &= \lim_{n\to\infty} Var_{\underline{\beta}_0}\left\{\sqrt{n}\frac{\partial}{\partial\underline{\beta}}\widehat{Q}_{1,n}\left(\underline{\beta}\right)\right\}, \\
I_{22} &= \lim_{n\to\infty} Var_{\underline{\pi}_0}\left\{\sqrt{n}\frac{\partial}{\partial\underline{\theta}}\widehat{Q}_{2,n}\left(\underline{\pi}\right)\right\}, \\
I_{12} &= \lim_{n\to\infty} E_{\underline{\pi}_0}\left\{n\frac{\partial}{\partial\underline{\beta}}\widehat{Q}_{1,n}\left(\underline{\beta}\right)\frac{\partial}{\partial\underline{\theta}'}\widehat{Q}_{2,n}\left(\underline{\pi}\right)\right\}, I_{21} = I_{12}', \\
J_{11} &= \lim_{n\to\infty} Var_{\underline{\beta}_0}\left\{\frac{\partial^2}{\partial\underline{\beta}\partial\underline{\beta}'}\widehat{Q}_{1,n}\left(\underline{\beta}\right)\right\}, \\
J_{22} &= \lim_{n\to\infty} Var_{\underline{\beta}_0}\left\{\frac{\partial^2}{\partial\underline{\theta}\partial\underline{\theta}'}\widehat{Q}_{2,n}\left(\underline{\pi}\right)\right\}, \\
J_{12} &= \lim_{n\to\infty} Var_{\underline{\beta}_0}\left\{\frac{\partial^2}{\partial\underline{\beta}\partial\underline{\theta}'}\widehat{Q}_{2,n}\left(\underline{\pi}\right)\right\}, J_{21} = J_{12}'.
\end{aligned}$$

3.4 Proofs

The proof of Theorems 1 and 3 are by now standard and follows from similar arguments used in showing the strong consistency of the $QMLE-PGARCH$ and $QMLE-PARMA-PGARCH$ models (cf. Aknouche and Bibi [2] and Bibi and Lescheb [10]) and hence, we do not detail the proof. Thus, we give only a sketch of proof for the asymptotic normality and refer to Aknouche and Bibi [2], Francq and Zakoïan [20], [21] or Bibi and Lescheb [10] for further details. Because $\left(\widehat{l}_t(\underline{\theta})\right)_{t\in\mathbb{Z}}$ (resp. $\left(\widehat{l}_{1,t}(\underline{\beta})\right)_{t\in\mathbb{Z}}$, $\left(\widehat{l}_{2,t}(\underline{\pi})\right)_{t\in\mathbb{Z}}$) is not a SPS process due to the presence of initial values, we shall replace it by its SPS version $(l_t(\underline{\theta}))_{t\in\mathbb{Z}}$ (resp. $\left(l_{1,t}(\underline{\beta})\right)_{t\in\mathbb{Z}}$, $(l_{2,t}(\underline{\pi}))_{t\in\mathbb{Z}}$) in which no constraint on the initial values were imposed.

3.4.1 Proof of the Theorem 2 [Asymptotic normality of LSE PGARCH]

Using Taylor-series expansion around $\underline{\theta}_0$, we obtain

$$\underline{O}=\frac{1}{\sqrt{n}}\sum_{t=0}^{n-1}\frac{\partial l_t(\widehat{\underline{\theta}}_n)}{\partial\underline{\theta}}=\frac{1}{\sqrt{n}}\sum_{t=0}^{n-1}\frac{\partial l_t(\underline{\theta}_0)}{\partial\underline{\theta}}+\left(\frac{1}{n}\sum_{t=0}^{n-1}\frac{\partial^2 l_t(\underline{\theta}_*)}{\partial\underline{\theta}\partial\underline{\theta}'}\right)\sqrt{n}\left(\widehat{\underline{\theta}}_n-\underline{\theta}_0\right) \quad (3.7)$$

where $\underline{\theta}_*$ is between $\widehat{\underline{\theta}}_n$and $\underline{\theta}_0$. Thus we need to show that $\frac{1}{\sqrt{n}}\sum\limits_{t=0}^{n-1}\frac{\partial l_t(\underline{\theta}_0)}{\partial\underline{\theta}}\rightsquigarrow N\left(\underline{O},\frac{4}{s^2}\mathcal{I}\left(\underline{\theta}_0\right)\right)$ and $\frac{1}{n}\sum\limits_{t=0}^{n-1}\frac{\partial^2 l_t(\underline{\theta}_*)}{\partial\underline{\theta}\partial\underline{\theta}'}\overset{P}{\longrightarrow}\frac{2}{s}\mathcal{J}\left(\underline{\theta}_0\right)$ and hence the result fellows from Slutsky's theorem and the following intermediate results grouped in the following lemma

Lemme 5 *Under* **A1**. – **A6**. *we have*

1. $E_{\underline{\theta}_0}\left\{\left\|\frac{\partial l_t(\underline{\theta})}{\partial\underline{\theta}}\frac{\partial l_t(\underline{\theta})}{\partial\underline{\theta}'}\right\|\right\}<+\infty$, $E_{\underline{\theta}_0}\left\{\left\|\frac{\partial^2 l_t(\underline{\theta})}{\partial\underline{\theta}\partial\underline{\theta}'}\right\|\right\}<+\infty$ *and* $E_{\underline{\theta}_0}\left\{\sup\limits_{\underline{\theta}\in\vartheta(\underline{\theta}_0)}\left|\frac{\partial^3 l_t(\underline{\theta})}{\partial\theta_i\partial\theta_j\partial\theta_k}\right|\right\}<+\infty$ *for some neighborhood* $\vartheta\left(\underline{\theta}_0\right)$ *of* $\underline{\theta}_0$ *and for all* $i,j,k\in\{1,...,s(p+q+1)\}$.
2. $\left\|\frac{1}{\sqrt{n}}\sum\limits_{t=0}^{n-1}\left\{\frac{\partial l_t(\underline{\theta})}{\partial\underline{\theta}}-\frac{\partial\widehat{l}_t(\underline{\theta})}{\partial\underline{\theta}}\right\}\right\|$ *and* $\sup\limits_{\underline{\theta}\in\vartheta(\underline{\theta}_0)}\left\|\frac{1}{n}\sum\limits_{t=0}^{n-1}\left\{\frac{\partial^2 l_t(\underline{\theta})}{\partial\underline{\theta}\partial\underline{\theta}'}-\frac{\partial^2\widehat{l}_t(\underline{\theta})}{\partial\underline{\theta}\partial\underline{\theta}'}\right\}\right\|$ *converges in probability to* 0 *as* $n\to\infty$.
3. $Var_{\underline{\theta}_0}\left\{\frac{\partial l_t\left(\underline{\theta}\right)}{\partial\underline{\theta}}\right\}=\frac{4}{s^2}\mathcal{I}\left(\underline{\theta}_0\right)$.

4. $\frac{1}{\sqrt{n}}\sum_{t=0}^{n-1}\frac{\partial l_t(\underline{\theta}_0)}{\partial\underline{\theta}}$ *converges in distribution to* $N\left(\underline{O},\frac{4}{s^2}\mathcal{I}\left(\underline{\theta}_0\right)\right)$.

5. $\frac{1}{n}\sum_{t=0}^{n-1}\frac{\partial^2 l_t\left(\underline{\theta}_*\right)}{\partial\underline{\theta}\partial\underline{\theta}'}$ *converges in probability to* $\frac{2}{s}\mathcal{J}\left(\underline{\theta}_0\right)$ *and* $\mathcal{J}\left(\underline{\theta}_0\right)$ *is non singular matrix.*

Proof. The proof follows from standard arguments (c.f. Aknouche and Bibi [2], Francq and Zakoïan [20] and Bibi and Lescheb [10]). ■

3.4.2 Proof of Theorem 4[Asymptotic normality of LSE PARMA-PGARCH]

The proof rests classically on the Taylor-series expansion around the true parameters values

$$\begin{aligned}
\underline{O} &= \frac{1}{\sqrt{n}}\sum\nolimits_{t=0}^{n-1}\frac{\partial l_{1,t}(\widehat{\underline{\beta}}_n)}{\partial\underline{\beta}}=\frac{1}{\sqrt{n}}\sum\nolimits_{t=0}^{n-1}\frac{\partial l_{1,t}(\underline{\beta}_0)}{\partial\underline{\beta}}\\
&+\left(\frac{1}{n}\sum\nolimits_{t=0}^{n-1}\frac{\partial^2 l_{1,t}(\underline{\beta}_*)}{\partial\underline{\beta}\partial\underline{\beta}'}\right)\sqrt{n}\left(\widehat{\underline{\beta}}_n-\underline{\beta}_0\right)\\
\underline{O} &= \frac{1}{\sqrt{n}}\sum\nolimits_{t=0}^{n-1}\frac{\partial l_{2,t}(\widehat{\underline{\beta}}_n,\widehat{\underline{\theta}}_n)}{\partial\underline{\theta}}=\frac{1}{\sqrt{n}}\sum\nolimits_{t=0}^{n-1}\frac{\partial l_{2,t}(\widehat{\underline{\beta}}_n,\underline{\theta}_0)}{\partial\underline{\theta}}\\
&+\left(\frac{1}{n}\sum\nolimits_{t=0}^{n-1}\frac{\partial^2 l_{2,t}(\widehat{\underline{\beta}}_n,\underline{\theta}_*)}{\partial\underline{\theta}\partial\underline{\theta}'}\right)\sqrt{n}\left(\widehat{\underline{\theta}}_n-\underline{\theta}_0\right)\\
&= \frac{1}{\sqrt{n}}\sum\nolimits_{t=0}^{n-1}\frac{\partial l_{2,t}(\underline{\beta}_0,\underline{\theta}_0)}{\partial\underline{\theta}}+\left(\frac{1}{n}\sum\nolimits_{t=0}^{n-1}\frac{\partial^2 l_{2,t}(\underline{\beta}_{**},\underline{\theta}_0)}{\partial\underline{\theta}\partial\underline{\beta}'}\right)\sqrt{n}\left(\widehat{\underline{\beta}}_n-\underline{\beta}_0\right)\\
&+\left(\frac{1}{n}\sum\nolimits_{t=0}^{n-1}\frac{\partial^2 l_{2,t}(\widehat{\underline{\beta}}_n,\underline{\theta}_*)}{\partial\underline{\theta}\partial\underline{\theta}'}\right)\sqrt{n}\left(\widehat{\underline{\theta}}_n-\underline{\theta}_0\right),
\end{aligned}$$

where $\underline{\beta}_*$'s (resp. $\underline{\beta}_{**}$, $\underline{\theta}_*$, $\underline{\pi}_*$) are between $\widehat{\underline{\beta}}_n$ and $\underline{\beta}_0$, (resp. $\widehat{\underline{\beta}}_n$ and $\underline{\beta}_0$, $\widehat{\underline{\theta}}_n$ and $\underline{\theta}_0$ and between $\widehat{\underline{\pi}}_n$ and $\underline{\pi}_0$). The above equations leads to

$$-\frac{1}{\sqrt{n}}\sum\nolimits_{t=0}^{n-1}\frac{\partial \underline{l}_t\left(\underline{\pi}_0\right)}{\partial\underline{\pi}}=\frac{1}{n}\sum\nolimits_{t=0}^{n-1}\frac{\partial^2 \underline{l}_t\left(\underline{\pi}_*\right)}{\partial\underline{\pi}\partial\underline{\pi}'}\sqrt{n}\left(\widehat{\underline{\pi}}_n-\underline{\pi}_0\right)$$

where $\frac{\partial \underline{l}_t\left(\underline{\pi}_0\right)}{\partial\underline{\pi}}:=\left(\frac{\partial l_{1,t}(\underline{\beta}_0)}{\partial\underline{\beta}'},\frac{\partial l_{2,t}(\underline{\pi}_0)}{\partial\underline{\theta}'}\right)'$. Thus we need to show that

$$\frac{1}{\sqrt{n}}\sum\nolimits_{t=0}^{n-1}\frac{\partial \underline{l}_t\left(\underline{\pi}_0\right)}{\partial\underline{\pi}}\rightsquigarrow N\left(\underline{O},I\right)\text{ and }\frac{1}{n}\sum\nolimits_{t=0}^{n-1}\frac{\partial^2 \underline{l}_t\left(\underline{\pi}_*\right)}{\partial\underline{\pi}\partial\underline{\pi}'}\longrightarrow J\text{ in probability}$$

with $I = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix}, J = \begin{pmatrix} J_{11} & O \\ J_{21} & J_{22} \end{pmatrix}$ where the matrices $(I_{ij})_{1\leq i,j\leq 2}$ and $(J_{ij})_{1\leq i,j\leq 2}$ are given in Theorem 4. The theorem will straightforwardly follow. For this purpose, we show an analogue Lemma 5 for $\dfrac{\partial \underline{l}_t(\underline{\pi}_0)}{\partial \underline{\pi}}$.

Lemme 6 *If the Assumptions **A1-A10** hold, then*

1. *For any* $\underline{\pi}\in\Theta$*, the random vectors* $\dfrac{\partial}{\partial \underline{\beta}} l_{1,n}(\underline{\beta})$, $\dfrac{\partial}{\partial \underline{\theta}} l_{2,n}(\underline{\pi})$ *exist and belong to* $\mathbb{L}_2$

2. $E_{\underline{\pi}_0}\left\{\left\|\dfrac{\partial \underline{l}_t(\underline{\pi})}{\partial \underline{\pi}}\dfrac{\partial \underline{l}_t(\underline{\pi})}{\partial \underline{\pi}'}\right\|\right\} < +\infty$, $E_{\underline{\pi}_0}\left\{\left\|\dfrac{\partial^2 \underline{l}_t(\underline{\pi})}{\partial \underline{\pi}\partial \underline{\pi}'}\right\|\right\} < +\infty$ *and* $E_{\underline{\pi}_0}\left\{\sup\limits_{\underline{\theta}\in\vartheta(\underline{\theta}_0)}\left|\dfrac{\partial^3 \underline{l}_t(\underline{\pi})}{\partial \pi_i\partial \pi_j\partial \pi_k}\right|\right\} < +\infty$ *for some neighborhood* $\vartheta(\underline{\pi}_0)$ *of* $\underline{\pi}_0$ *and for all* $i,j,k\in\{1,...,s(p+q+P+Q+1)\}$.

3. $\left\|\frac{1}{\sqrt{n}}\sum_{t=0}^{n-1}\left\{\dfrac{\partial \underline{l}_t(\underline{\pi})}{\partial \underline{\pi}} - \dfrac{\partial \widehat{\underline{l}}_t(\underline{\pi})}{\partial \underline{\pi}}\right\}\right\|$ *and* $\sup\limits_{\underline{\pi}\in\vartheta(\underline{\pi}_0)}\left\|\frac{1}{n}\sum_{t=0}^{n-1}\left\{\dfrac{\partial_t^2 \underline{l}(\underline{\pi})}{\partial \underline{\pi}\partial \underline{\pi}'} - \dfrac{\partial^2 \widehat{\underline{l}}_t(\underline{\pi})}{\partial \underline{\pi}\partial \underline{\pi}'}\right\}\right\|$ *converges in probability to* 0 *as* $n\to\infty$.

4. $\frac{1}{\sqrt{n}}\sum_{t=0}^{n-1}\dfrac{\partial \underline{l}_t(\underline{\pi}_0)}{\partial \underline{\pi}}$ *converges in distribution to* $N(\underline{O}, I(\underline{\pi}_0))$ *where the sub-matrices* I_{11}, I_{12} *and* I_{22} *exist and are strictly positive definite.*

5. $\frac{1}{n}\sum_{t=0}^{n-1}\dfrac{\partial_t^2 \underline{l}(\underline{\pi}_*)}{\partial \underline{\pi}\partial \underline{\pi}'}$ *converges in probability to* $J(\underline{\pi}_0)$ *and* $(J_{ii}(\underline{\pi}_0))_{1\leq i\leq 2}$ *are non singular matrices.*

Proof.

1. Noting that under the Assumptions **A7-A9** $E\{X_t^4\} < +\infty$. By Cauchy-Schwartz inequality we can see that $\dfrac{\partial \epsilon_{st+v}^2(\underline{\beta})}{\partial \underline{\beta}} = 2\epsilon_{st+v}(\underline{\beta})\dfrac{\partial \epsilon_{st+v}(\underline{\beta})}{\partial \underline{\beta}}$ and $\dfrac{\partial \eta_{st+v}^2(\underline{\pi})}{\partial \underline{\theta}} = 2\eta_{st+v}(\underline{\pi})\dfrac{\partial \eta_{st+v}(\underline{\pi})}{\partial \underline{\theta}}$ belong to $\mathbb{L}_2$.

2. The statements in Assertions 2 and 3 follows similarly as proving Lemma 5.

3. By Assumption **A7-A9**, we have $\Im_t^{(\epsilon)} = \Im_t^{(X)}$, $E_{\underline{\beta}_0}\left\{\dfrac{\partial l_{1,t}(\underline{\beta}_0)}{\partial \underline{\beta}}|\Im_{t-1}^{(X)}\right\} = \underline{O}$, $E_{\underline{\pi}_0}\left\{\dfrac{\partial l_{2,t}(\underline{\pi}_0)}{\partial \underline{\theta}}|\Im_{t-1}^{(X)}\right\} = \underline{O}$ and $Var_{\underline{\beta}_0}\left\{\dfrac{\partial l_{1,t}(\underline{\beta}_0)}{\partial \underline{\beta}}\right\}$ and $Var_{\underline{\pi}_0}\left\{\dfrac{\partial l_{2,t}(\underline{\pi}_0)}{\partial \underline{\theta}}\right\}$ exists and not singular matrices. Hence, for any $(\underline{\lambda}', \underline{\mu}')' \in \mathbb{R}^{s(P+Q)} \times$

$\mathbb{R}^{s(p+q+1)}$, the sequence $\left\{ \left(\underline{\lambda}', \underline{\mu}'\right) \dfrac{\partial \underline{l}_t\left(\underline{\pi}\right)}{\partial \underline{\pi}}, \Im_t^{(X)} \right\}_t$ is a square integrable SPS martingale difference. The central limit theorem and the Wold-Cramèr device allow to derive the asymptotic normality result.

4. The convergence follows from the $a.s$ convergence of $\underline{\pi}_*$ to $\underline{\pi}_0$, an application of the ergodic theorem to $\frac{1}{n}\sum\limits_{t=0}^{n-1} \dfrac{\partial^2 \underline{l}_t\left(\underline{\pi}_*\right)}{\partial \underline{\pi} \partial \underline{\pi}'}$ and the fact that almost surely as $n \to \infty$

$$\left\| \frac{1}{n} \sum\nolimits_{t=0}^{n-1} \left(\frac{\partial^2 l_{1,t}(\underline{\beta}_*)}{\partial \underline{\beta} \partial \underline{\beta}'} - \frac{\partial^2 l_{1,t}(\underline{\beta}_0)}{\partial \underline{\beta} \partial \underline{\beta}'} \right) \right\| \to 0,$$

$$\left\| \frac{1}{n} \sum\nolimits_{t=0}^{n-1} \left(\frac{\partial^2 l_{2,t}(\underline{\beta}_{**}, \underline{\theta}_0)}{\partial \underline{\theta} \partial \underline{\beta}'} - \frac{\partial^2 l_{2,t}(\underline{\pi}_0)}{\partial \underline{\theta} \partial \underline{\beta}'} \right) \right\| \to 0,$$

$$\left\| \frac{1}{n} \sum\nolimits_{t=0}^{n-1} \left(\frac{\partial^2 l_{2,t}(\widehat{\underline{\beta}}_n, \underline{\theta}_*)}{\partial \underline{\theta} \partial \underline{\theta}'} - \frac{\partial^2 l_{2,t}(\underline{\pi}_0)}{\partial \underline{\theta} \partial \underline{\theta}'} \right) \right\| \to 0.$$

■

Chapter 4

Yule-Waker equations for $PGARCH(1,1)$ models

Abstract: This chapter studies the probabilistic structure and asymptotic inference of the first order periodic generalized autoregressive conditional heteroscedasticity ($PGARCH(1,1)$) models in which the parameters in volatility process are allowed to switch between different regimes. First, we establish necessary and sufficient conditions for a $PGARCH(1,1)$ process to have a unique stationary solution (in periodic sense) and for the existence of moments of any order. Second, using the representation of squared $PGARCH(1,1)$ model as a $PARMA(1,1)$ model, we then consider Yule-Walker type estimators for the parameters in $PGARCH(1,1)$ model and derives their consistency and asymptotic normality. The estimator can be surprisingly efficient for quite small numbers of autocorrelations and, in some cases can be more efficient than the least squares estimate (LSE). We use a residual bootstrap to define bootstrap estimators for the Yule-Walker estimates and prove the consistency of this bootstrap method. A set of numerical experiments illustrates the practical relevance of our theoretical results.

4.1 Introduction

In this chapter, we continuous to investigate the asymptotic behavior of empirical studies of $PGARCH$ models. Since $PGARCH(p,q)$ models with $p,q \geq 2$ are rare in practice, we restrict ourselves to one particular model which has very often used in applications: the $PGARCH\,(1,1)$ model in which

$$h_n^2 = \omega(s_n) + \alpha(s_n)X_{n-1}^2 + \beta(s_n)h_{n-1}^2 \text{ with } s_n = \sum\nolimits_{\upsilon=1}^{s} \upsilon \mathbb{I}_{\Delta(\upsilon)}(n), \tag{4.1}$$

where $\Delta(\upsilon) := \{st + \upsilon, 1 \leq \upsilon \leq s, t \in \mathbb{Z}\}$ denotes the set of indices corresponding to regime υ. So by setting $n = st + \upsilon$, $e_t(\upsilon) := e_{st+\upsilon}, h_t(\upsilon) := h_{st+\upsilon}$ and $X_t(\upsilon) := X_{st+\upsilon}$, $1 \leq \upsilon \leq s$, and the convenience $X_t(\upsilon) := X_{t-1}(s-\upsilon)$ (respectively $h_t(\upsilon) := h_{t-1}(s-\upsilon)$ and $e_t(\upsilon) := e_{t-1}(s-\upsilon)$) if $\upsilon \leq 0$, Models (4.1) may be equivalently written as

$$\begin{cases} X_t(\upsilon) = h_t(\upsilon)e_t(\upsilon), \\ h_t^2(\upsilon) = \omega(\upsilon) + \alpha(\upsilon)X_t^2(\upsilon-1) + \beta(\upsilon)h_t^2(\upsilon-1),\ \upsilon = 1, ..., s. \end{cases} \tag{4.2}$$

In (4.2) $X_t(\upsilon)$ (respectively $h_t(\upsilon)$, $e_t(\upsilon)$) refers to X_t (respectively h_t, e_t) during 'season' υ, $1 \leq \upsilon \leq s$, of 'year' t. In the sequel, the notations X_t, h_t, etc. are used in preference to $X_t(\upsilon)$, $h_t(\upsilon)$, etc. whenever emphasis on seasonality is not paramount.

4.2 Probability structure

In this section, we are interested in conditions ensuring the existence of causal solutions, i.e., solutions such that X_t is measurable with respect to $\Im_t^{(e)} := \sigma\,\{e_{t-k}, k \geq 0\}$. For this purpose, letting $g_x(y) = \alpha(x)y + \beta(x)$, then we have $X_t(\upsilon) := h_t(\upsilon)e_t(\upsilon)$ and $h_t^2(\upsilon) = \omega(\upsilon) + g_\upsilon(e_t^2(\upsilon-1))\ h_t^2(\upsilon-1)$, $\upsilon = 1, ..., s$. This formulation shows that h_t can be viewed as a separable Markov chain and thus one can uses the theory of Markov chains to study properties of either the joint process $(X_t, h_t)_{t\in\mathbb{Z}}$ or of $(h_t)_{t\in\mathbb{Z}}$ in isolation from $(X_t)_{t\in\mathbb{Z}}$. By recursing the last equation we obtain

$$\begin{aligned} h_t^2(s) \;&=\; \left\{\prod\nolimits_{\upsilon=0}^{s-1} g_{s-\upsilon}\left(e_t^2(s-\upsilon-1)\right)\right\} h_{t-1}^2(s) \\ &\quad + \sum\nolimits_{k=0}^{s-1}\left\{\prod\nolimits_{\upsilon=0}^{k-1} g_{s-\upsilon}\left(e_t^2(s-\upsilon-1)\right)\right\}\omega(s-k). \end{aligned} \tag{4.3}$$

where, as usual, empty products are set equal to one. Now, set $Z_t = h_t^2(s)$, $G(\underline{e}_t^2) = \prod_{\upsilon=0}^{s-1} g_{s-\upsilon}\left(e_t^2(s-\upsilon-1)\right)$ and

$$\xi(\underline{e}_t^2) = \sum\nolimits_{k=0}^{s-1}\left\{\prod\nolimits_{\upsilon=0}^{k-1} g_{s-\upsilon}\left(e_t^2(s-\upsilon-1)\right)\right\}\omega(s-k)$$

where $\underline{e}_t^2 = (e_t^2(1), .., e_t^2\,(s-1))$ and rewrite (4.1) as

$$Z_t = G(\underline{e}_t^2)Z_{t-1} + \xi(\underline{e}_t^2). \tag{4.4}$$

The Representation (4.4) is potentially useful for deriving probabilistic properties for $(X_t, h_t)_{t\in\mathbb{Z}}$. Note that$(G(\underline{e}_t^2), \xi(\underline{e}_t^2))$ being *i.i.d.* pairs of random variables, independent of Z_k for any $k < t$. This process is clearly Markovian with state space $\mathbb{R}_+$ and transition probability measure $P(z, .)$ equal to the distribution of $G(\underline{e}_t^2)z + \xi(\underline{e}_t^2)$. However, since the probabilistic properties of Models (4.2) and (4.4) are the same (cf. Bibi and Lescheb [10] and Lee and Shin [34]), we shall restrict ourselves by studying the latter one. Hence, the solutions of (4.2) are called to be a strictly periodically stationary (SPS) (resp. periodically ergodic) whenever the version (4.4) has a strictly stationary (resp. ergodic) solutions. The important results on SPS solutions of (4.2) are summarized in the following theorem

Theorem 1 *Let $(X_t, h_t)_{t\in\mathbb{Z}}$ be the $PGARCH\,(1,1)$ process defined by (4.2). Additionally, assume that*

$$-\infty \leq \gamma_L := \inf_{n>0} \frac{1}{n} E\left\{\prod_{j=0}^{n} G(\underline{e}_{t-j}^2)\right\} = \sum\nolimits_{v=1}^{s} E\left\{\log\left(\ g_v(e_0^2)\right)\right\} < 0. \quad (4.5)$$

Then under the Condition (4.5), a causal, SPS solution for (4.2) is given by

$$\begin{cases} X_t\,(v) = h_t(v)e_t(v) \\ h_t^2\,(v) = \omega(v) + \sum_{k\geq 1}\left\{\prod_{i=0}^{k-1} g_{v-i}\left(e_t^2(v-i-1)\right)\right\}\omega(v-k) \end{cases} \quad (4.6)$$

with the above series converging almost surely (a.s.). Moreover, the solution process is unique and periodically ergodic.

Conversely, if $\gamma_L \geq 0$, there is no a SPS solution $(X_t, h_t)_{t\in\mathbb{Z}}$ to model (4.2). More precisely $h_t \to +\infty$, a.s. as $t \to +\infty$ whenever $\gamma_L > 0$, otherwise $h_t \to +\infty$ in probability as $t \to +\infty$.

Proof. The proof rests classically by Theorem 1 of Brandt [17] and Theorem 1.1 of Bougerol and Picard [16] using (4.4). ■

Remark 2 *The condition $E\{\log\{g_v(e_0^2(v))\}\} < 0$ for all $v = 1, ..., s$ (local stationarity) implies the existence of SPS solution for Model (4.2). The converse is not true, i.e., the Condition (4.5) (global stationarity) does not entail local stationarity of all regimes. This mean that the existence of some explosive regimes (i.e., $E\{\log\{g_v(e_0^2(v))\}\} \geq 0$) does not preclude the existence of SPS solution.*

Remark 3 *For the $PARCH(1)$ model, we obtain*

$$\gamma_L = E\left\{\log\left\{\prod\nolimits_{v=1}^{s} \alpha(v)\, e_0^2(v)\right\}\right\}$$

and hence $\gamma_L < 0$ if and only if $\prod_{v=1}^{s} \alpha(v) < \exp\{-sE\{\log e_0^2\}\}$.

In order to make an estimation theory possible, the process solution need to have some moments. For instance, though the Criterion (4.5) could be used as a sufficient condition ensuring the finiteness of $E\left\{X_t^{2r}\right\}$ for some $r \in]0,1]$ (cf. Aknouche and Bibi [2]), it is of little use in practice, and it may have importance from theoretical point of view. Therefore, we have to search for conditions based on parameters of model ensuring the existence of second order moments for the strict stationary solution. Under such conditions $(X_t)_{t\in\mathbb{Z}}$ is called periodically correlated (PC) process characterized by $E\left\{X_t\right\} = E\left\{X_{t+s}\right\}$ and $Cov\left(X_{t+s}, X_{r+s}\right) = Cov\left(X_t, X_r\right)$ for all $t, r \in \mathbb{Z}$.

Theorem 4 *Let* $(X_t, h_t)_{t\in\mathbb{Z}}$ *be the* $PGARCH\,(1,1)$ *process defined by* (4.2). *Then*

1. *if*
$$\lambda_{(1)} := E\left\{G(\underline{e}_t^2)\right\} = \prod_{i=1}^{s} \left(\alpha\left(i\right) + \beta(i)\right) < 1. \tag{4.7}$$
the $PGARCH\,(1,1)$ *model* (4.2) *has an unique,* PC*, causal, periodically ergodic solution given by* (4.6) *in which the series converges a.s. and in* $\mathbb{L}_1$*. Moreover, the solution process is* SPS *and satisfies* $E\left\{X_t\right\} = 0$ *and* $Cov\left(X_t, X_r\right) = 0$ *for all* $t \neq r$*.*

2. *Conversely, if* $\lambda_{(1)} \geq 1$*, then there is no a* SPS *solution* $(X_t, h_t)_{t\in\mathbb{Z}}$ *to model* (4.2) *such that* $E\left\{X_t^2\right\} < \infty$*.*

Proof. The proof follows essentially the same arguments as in Bibi and Aknouche [9]. ∎

Remark 5 *Since the Conditions* (4.5) *and* (4.7) *are necessary and sufficient, we have necessarily* $\left[\lambda_{(1)} < 1\right] \Longrightarrow \left[\gamma_L < 0\right]$*.*

Remark 6 *If* e_0 *has a positive and continuous density* g*, then under the Condition* (4.7) *the process* $(X_t, h_t)_{t\in\mathbb{Z}}$ *is geometrically ergodic (cf. Bibi and Aknouche [8]) and if it is initialized from its* SPS *distribution, then the process* $(X_t, h_t)_{t\in\mathbb{Z}}$ *is* $\beta-$*mixing with the* $\beta-$*mixing coefficient satisfy* $\beta_k \leq c\rho^k$*,* $k \in \mathbb{Z}_+$ *for some constants* $0 < \rho < 1$ *and* $c > 0$*.*

The third exploration of the Representation (4.4) is for the existence of higher-order moments.

Theorem 7 *Let* $(X_t, h_t)_{t\in\mathbb{Z}}$ *be the* $PGARCH\,(1,1)$ *process defined by* (4.2) *and assume that* $\kappa_m = E\left\{e_0^{2m}\right\} < +\infty$*, for some integer* $m \in [1, \infty[$*. Then the following statements are equivalent*

1. $E\left\{X_t^{2m}\right\} < +\infty$

2. $\lambda_{(m)} := E\left\{G^m(\underline{e}_t^2)\right\} = \prod_{i=1}^{s} E\left\{\left(\alpha\left(i\right)e_0^2 + \beta(i)\right)^m\right\} < 1.$

Proof. The proof follows by induction using the development

$$Z_t^m = G^m(\underline{e}_t^2)Z_{t-1}^m + \xi^m\left(\underline{e}_t^2\right) + \sum_{\ell=1}^{m-1}\frac{m!}{\ell!\left(m-\ell\right)!}G^{m-\ell}(\underline{e}_t^2)\xi^\ell\left(\underline{e}_t^2\right)Z_{t-1}^{m-\ell},\ m\geq 1.$$

■

We conclude this section with a periodic $ARMA$ representation of the process $(X_t^2)_{t\in\mathbb{Z}}$ which will be used in the next section. For this purpose assume that $(X_t, h_t)_{t\in\mathbb{Z}}$ is a SPS process and defining the martingale difference sequence $\eta_t := h_t^2\left(e_t^2-1\right)$, so we have the following periodic $ARMA\left(1,1\right)$ ($PARMA\left(1,1\right)$) representation

$$X_t^2\left(\upsilon\right) = \omega(\upsilon) + \left(\alpha\left(\upsilon\right)+\beta\left(\upsilon\right)\right)X_t^2\left(\upsilon-1\right) + \eta_t\left(\upsilon\right) - \beta\left(\upsilon\right)\eta_t\left(\upsilon-1\right). \quad (4.8)$$

The $PARMA$ models are not only of interest in their own right, but, because of their connection with multivariate stationary $ARMA$ models. Indeed, by stack the s regimes in vector $\underline{X}_t^2 := \left(X_t^2\left(1\right), ..., X_t^2\left(s\right)\right)'$ the Equation (4.8) has the $s-$variate $ARMA\left(1,1\right)$ representation

$$A_0\underline{X}_t^2 = \underline{\omega} + A_1\underline{X}_{t-1}^2 + B_0\underline{\eta}_t + B_1\underline{\eta}_{t-1} \quad (4.9)$$

where $\left(\underline{\eta}_t\right)_{t\in\mathbb{Z}}$ is the vector of innovation process containing the stacked η_t variables. The precise expressions of the $s\times s$ matrices $\left(A_i\right)_{0\leq i\leq 1}$ and $\left(B_i\right)_{0\leq i\leq 1}$ can be found in Bibi and Lescheb [10]. The $VARMA\left(1,1\right)$ Model (4.9) is causal (and hence stationary) provided that

$$\det\left(A_0 - A_1 z\right) \neq 0 \text{ for all complex } z \text{ satisfying } |z|\leq 1. \quad (4.10)$$

It is straightforward to verify that the causality condition (4.10) reduce to (4.7). Furthermore, the process $\left(X_t, h_t\right)_{t\in\mathbb{Z}}$ is said to be periodically integrated ($IPGARCH\left(1,1\right)$) when $\prod_{i=1}^{s}\left(\alpha\left(i\right)+\beta(i)\right) = 1$. The latter include the usual $IGARCH\left(1,1\right)$ process $\alpha\left(i\right)+\beta(i) = 1$ for all i nested within general $IPGARCH$ process. It is worth noting here that the $IGARCH$ models can be strictly stationary (unlike to $ARIMA$ models) and geometrically $\beta-$mixing processes (cf. Meitz and Saikonnen [38]). Since, from (4.9) we have

$$\Delta\underline{X}_t^2 = \underline{X}_t^2 - \underline{X}_{t-1}^2 = A_0^{-1}\underline{\omega} + \Pi\underline{X}_{t-1}^2 + A_0^{-1}B_0\underline{\eta}_t + A_0^{-1}B_1\underline{\eta}_{t-1}$$

where $\Pi = A_0^{-1}(A_1 - A_0)$. Then, periodic stationarity implies that $A_1 - A_0$ and hence Π is non singular. Periodic integration, on the other hand, implies that $rank(\Pi) = s - 1$ so that $\left(\underline{X}_t^2\right)_{t\in\mathbb{Z}}$ is cointegrated of order $(1,1)$. Hence, the associated $PARMA(1,1)$ with the parametrization $\beta(v) = 1 - \alpha(v)$ for all v is an $PARIMA(1,1,1)$

$$\Delta\underline{X}_t^2 = A_0^{-1}\underline{\omega} + A_0^{-1}B_0\underline{\eta}_t + A_0^{-1}B_1\underline{\eta}_{t-1} \tag{4.11}$$

showing that $\left(\underline{X}_t^2\right)_{t\in\mathbb{Z}}$ is an integrated process with moving average error term. For the theoretical development, $\left(\Delta\underline{X}_t^2\right)_{t\in\mathbb{Z}}$ must be a weak stationary process even when the process $\left(\underline{X}_t^2\right)_{t\in\mathbb{Z}}$ does not. In spite of the right side in (4.11) is a moving average, we cannot conclude anything about the weak stationarity (even in the Gaussian strong $IPGARCH$) of $\left(\Delta\underline{X}_t^2\right)_{t\in\mathbb{Z}}$ that requires that $\lim_{t\to\infty} E\left\{\eta_t^2|\Im_{t-1}^{(e)}\right\} < +\infty$ (See Kim and Linton [31] for further discussion).

4.3 Asymptotic properties of empirical mean and covariance of squared process

In this section, we are concerned with the problem of asymptotic behavior of empirical mean and covariance of $(X_t^2)_{t\in\mathbb{Z}}$ which are needs laters. Let $\{X_1^2,\ldots,X_n^2\}$ be a realization of length $n = sN$ of the unique PC solution for the Equation (4.8), or equivalently a realization $\left\{\underline{X}_1^2,...,\underline{X}_N^2\right\}$ from a second order stationary process $\left(\underline{X}_t^2\right)_{t\in\mathbb{Z}}$ defined by (4.9). Let $\mu_2(v) = E\{X_t^2(v)\}$ and $\gamma_v(h) = Cov(X_t^2(v), X_t^2(v-h))$ be the season v means and covariances functions at lag $h \geq 0$ and their samples estimates $\hat{\mu}_2(v) := \frac{1}{N}\sum_{t=0}^{N-1} X_t^2(v)$ and $\widehat{\gamma}_v(h) := \frac{1}{N}\sum_{t=0}^{N-1} X_t^2(v)X_t^2(v-h) - \widehat{\mu}_2(v)\widehat{\mu}_2(v-h)$. Define the vectors $\underline{\hat{\mu}}_2 := (\hat{\mu}_2(1),...,\hat{\mu}_2(s))'$, $\underline{\mu}_2 := (\mu_2(1),\ldots, \mu_2(s))'$, $\underline{\widehat{\gamma}}(h) := (\widehat{\gamma}_1(h),...,\widehat{\gamma}_s(h))'$ and $\underline{\gamma}(h) := (\gamma_1(h),...,\gamma_s(h))'$. Noting that the dependence of the above estimates on N is generally suppressed hereafter for notational convenience. The following results characterize the asymptotic behavior of the empirical mean $\hat{\mu}_2(v)$.

Proposition 8 *Consider the $PARMA(1,1)$ representation (4.8) of $PGARCH$ (1, 1) Model (4.2) and let $(\underline{X}_t^2)_{t\in\mathbb{Z}}$ be the associated $VARMA$ Representation (4.9). Under the Conditions of Theorem 7 and if $(\underline{X}_t^2)_{t\in\mathbb{Z}}$ admits a moments up to $2-th$ order, then for each $v, v' \in \{1,...,s\}$*

1. *$\hat{\mu}_2(v)$ converges to $\mu_2(v)$ a.s. as $N \to \infty$.*

2. $\lim_{N\to\infty} NCov\left(\hat{\mu}_2(\upsilon), \hat{\mu}_2(\upsilon')\right) = (V_{as})_{\upsilon,\upsilon'} := \sum_{k\in\mathbb{Z}} \gamma_{\upsilon'}(\upsilon' - \upsilon + sk)$ *where* $V_{as} := \sum_{h\in\mathbb{Z}} Cov\left(\underline{X}_t^2, \underline{X}_{t-h}^2\right)$

3. $\lim_{N\to\infty} E\left\{\hat{\mu}_2(\upsilon) - \mu_2(\upsilon)\right\}^2 = 0$

4. *The vector* $\sqrt{N}\left(\underline{\hat{\mu}}_2 - \underline{\mu}_2\right)$ *converges in distribution to* $\mathcal{N}\left(\underline{0}, V_{as}\right)$.

Proof. Since $(\underline{X}_t^2)_{t\in\mathbb{Z}}$ is stationary and ergodic process, then the almost surely convergence is immediate. On the other hand, since

$$\lim_{N\to\infty} NCov(\hat{\mu}_2(\upsilon), \hat{\mu}_2(\upsilon')) = \lim_{N\to\infty} \sum_{|h|<N} \left(1 - \frac{|h|}{N}\right)\left(Cov\left(\underline{X}_t^2, \underline{X}_{t-h}^2\right)\right)_{\upsilon,\upsilon'},$$

then using the dominated convergence theorem, the second and the third assertions follows. To show the Assertion 4 it is not difficult to see from (4.9) that

$$\underline{X}_t^2 - \underline{\mu}_2 = \underline{U}_t + \underline{W}_t \tag{4.12}$$

where for any integer $m \geq 1$, $\underline{U}_t = \sum_{k=0}^{m} \Phi^k A_0^{-1}\left(B_0\underline{\eta}_{t-k} + B_1\underline{\eta}_{t-k-1}\right)$, $\underline{W}_t = \sum_{k=m+1}^{\infty} \Phi^k A_0^{-1}\left(B_0\underline{\eta}_{t-k} + B_1\underline{\eta}_{t-k-1}\right)$ with $\Phi = A_0^{-1}A_1$. Let $\widehat{\underline{Q}} = \frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}\underline{U}_t$ and $\widehat{\underline{V}} = \frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}\underline{W}_t$, then $\sqrt{N}\left(\underline{\hat{\mu}}_2 - \underline{\mu}_2\right) = \widehat{\underline{Q}} + \widehat{\underline{V}}$. Since $(\underline{U}_t)_{t\in\mathbb{Z}}$ is an $(m+1)-$dependent stationary process and $Var(\widehat{\underline{V}})$ tends to 0 as $m\to\infty$ uniformly in N, and hence $\widehat{\underline{V}}$ converges in distribution to $\underline{0}$ as $m\to\infty$ uniformly in N, then, the asymptotic distribution of $\sqrt{N}\left(\widehat{\underline{\mu}}_2 - \underline{\mu}_2\right)$ is the same as that of $\widehat{\underline{Q}}$. Moreover, for m fixed $\widehat{\underline{Q}} \rightsquigarrow \mathcal{N}\left(\underline{0}, V\right)$ where $V := \sum_{h=-m}^{m} Cov\left(\underline{U}_t, \underline{U}_{t-h}\right)$ (see Jiming [29], Chapter 8). As $m\to\infty$, $\underline{U}_t$ converges to $\underline{X}_t^2$ in probability and V converge to $V_{as} := \sum_{k\in\mathbb{Z}} Cov\left(\underline{X}_t^2, \underline{X}_{t-k}^2\right) < \infty$. ■

Similar results can be addressed for the empirical covariance function $\widehat{\gamma}_\upsilon(h)$.

Proposition 9 *Consider the* $PARMA\,(1,1)$ *Representation* (4.8) *of* $PGARCH\,(1,1)$ *model* (4.2) *and let* $(\underline{X}_t^2)_{t\in\mathbb{Z}}$ *be the associated* $VARMA$ *Version* (4.9). *Under the Conditions of Theorem 7 and if* $(\underline{X}_t^2)_{t\in\mathbb{Z}}$ *admits a moments up to* $4-th$ *order then for any* $h \geq 0$ *and for each* $\upsilon, \upsilon' \in \{1, ..., s\}$

1. $\widehat{\gamma}_\upsilon(h)$ *converges a.s. to* $\gamma_\upsilon(h)$ *as* $N\to\infty$

2. $\lim_{N\to\infty} NCov(\widehat{\gamma}_\upsilon(h), \widehat{\gamma}_\upsilon(k)) = (W_{as}(h,k))_{\upsilon,\upsilon'}$ *where* $W_{as}(h,k) := \sum_{l\in\mathbb{Z}} Cov\left(\underline{X}_t^2 \odot \underline{X}_t^2(h), \underline{X}_{t-l}^2 \odot \underline{X}_{t-l}^2(k)\right)$

3. $\lim_{N\to\infty} E\left\{\widehat{\gamma}_\upsilon(h) - \gamma_\upsilon(h)\right\}^2 = 0$ *for all* $\upsilon \in \{1, ..., s\}$

4. *The vector* $\sqrt{N}\left(\widehat{\underline{\gamma}}(h)-\underline{\gamma}(h)\right)$ *converges in distribution to* $\mathcal{N}\left(\underline{0},W_{as}(h)\right)$, *where* $W_{as}(h):=W_{as}(h,h)$.

Proof. For any integer $h\geq 0$, we rewrite the vector $\widehat{\underline{\gamma}}(h)$ as $\widehat{\underline{\gamma}}(h)=\frac{1}{N}\sum_{t=0}^{N-1}\underline{X}_t^2\odot\underline{X}_t^2(h)-\widehat{\underline{\mu}}_2\odot\widehat{\underline{\mu}}_2(h)$ where $\underline{X}_t^2(h):=\left(X_t^2\left(1-h\right),...,X_t^2(s-h)\right)'$ and $\widehat{\underline{\mu}}_2(h)$ its sample estimate. By the ergodicity, the first, second and third assertions follows. Since $\underline{\mu}_2\odot\underline{\mu}_2(h)=-\underline{\gamma}(h)+E\left\{\underline{X}_t^2\odot\underline{X}_t^2(h)\right\}$, then the asymptotic distribution of $\sqrt{N}\left(\widehat{\underline{\gamma}}(h)-\underline{\gamma}(h)\right)$ is the same as that of

$$\frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}\left(\underline{X}_t^2\odot\underline{X}_t^2(h)-E\left\{\underline{X}_t^2\odot\underline{X}_t^2(h)\right\}\right) \tag{4.13}$$

Simple computation using (4.12) shows that the asymptotic distribution of (4.13) is the same as that of

$$\frac{1}{\sqrt{N}}\left\{\sum_{t=0}^{N-1}\left(\underline{U}_t\odot\underline{U}_t(h)-E\left\{\underline{U}_t\odot\underline{U}_t(h)\right\}\right)\right\}$$

as $m\to\infty$. Now, for any $s\times 1$ vectors $\underline{\lambda}$ let $\widehat{P}(h)=\frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}Y_t(h)$ where $Y_t(h):=\underline{\lambda}'\left(\underline{U}_t\odot\underline{U}_t(h)-E\left\{\underline{U}_t\odot\underline{U}_t(h)\right\}\right)$. Clearly $\left(Y_t(h)\right)_{t\in\mathbb{Z}}$ is also a stationary $(m+1)$-dependent process with $E\left\{Y_t(h)Y_{t-k}(h)\right\}=\underline{\lambda}'W_k(h)\underline{\lambda}<+\infty$ where $W_k(h):=Cov\left(\underline{U}_t\odot\underline{U}_t(h),\underline{U}_{t-k}\odot\underline{U}_{t-k}(h)\right)$. Therefore, we have $\frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}Y_t(h)\rightsquigarrow\mathcal{N}\left(0,\underline{\lambda}'W(h)\underline{\lambda}\right)$ where $W(h):=\sum_{k=-m}^{m}W_k(h)$. As $m\to\infty$, $W(h)$ converges to $W_{as}\left(h\right)=\sum_{k\in\mathbb{Z}}Cov\left(\underline{X}_t^2\odot\underline{X}_t^2(h),\underline{X}_{t-k}^2\odot\underline{X}_{t-k}^2(h)\right)$ hence the proof follows. ■

Remark 10 *The asymptotic distribution of* $\widehat{\gamma}_v(h)$ *has been examined by many authors (see for instance [39] for an extensive discussion of this problem) without requiring the 8th moment but the limit distribution is not Gaussian.*

4.4 Yule-Walker estimation in $PGARCH\left(1,1\right)$ processes and its asymptotic properties

One of the most commonly used estimation procedures for $PGARCH$ models is the $QMLE$ approach. (see for the instance [2]). In this approach, the estimator is obtained as a minimizer of a Gaussian likelihood function. However, and in spite of its strongly consistency and its asymptotic normality, this estimator does not admit a closed-form expression. The main aim of this paper is thus to propose an estimator of $PGARCH\left(1,1\right)$ based on the Yule-Walker equations of $PARMA$ representation (4.8) which has a closed-form expression, computationally easy and which, compares favourably with the $QMLE$ and also with LSE (see [32]). As is well established in $PARMA$ models (see Lund and Basawa [35]), the empirical covariances of the process can be used to obtain Yule-Walker type estimator for the parameter $\underline{\theta} = (\underline{\theta}'(1), ..., \underline{\theta}'(s))'$ where $\underline{\theta}(v) = (\alpha(v), \beta(v), \omega(v))'$, $v = 1, ..., s$. Indeed, considering the centered squared process

$$X_t^2(v) - \mu_2(v) = (\alpha(v) + \beta(v)) \left(X_t^2(v-1) - \mu_2(v-1)\right) + \eta_t(v) - \beta(v)\,\eta_t(v-1), \tag{4.14}$$

and assume that $E\{X_t^4\} < \infty$. Conditions for the existence of moments are given in Theorem 7. Set $\sigma_\eta^2(v) = Var\{\eta_t(v)\}$ and noting that $E\{X_t^2(v)\eta_t(v)\} = \sigma_\eta^2(v)$, $E\{X_t^2(v)\eta_t(v-1)\} = \alpha(v)\,\sigma_\eta^2(v-1)$ and $E\{X_t^2(v-h)\eta_t(v)\} = 0$ for all $h > 0$. Then, by multiplying both sides of Equation (4.14) with $X_t^2(v-h)$, $h \geq 0$ and computing the expectations, we obtain the following identities

$$\begin{aligned}
\gamma_v(0) - (\alpha(v) + \beta(v))\,\gamma_v(1) &= \sigma_\eta^2(v) - \alpha(v)\beta(v)\,\sigma_\eta^2(v-1) \\
\gamma_v(1) - (\alpha(v) + \beta(v))\,\gamma_{v-1}(0) &= -\beta(v)\,\sigma_\eta^2(v-1), \\
\gamma_v(h) - (\alpha(v) + \beta(v))\,\gamma_{v-1}(h-1) &= 0,\ h \geq 2.
\end{aligned}$$

Elimination of $\sigma_\eta^2(.)$ gives the equations

$$\begin{cases}
\alpha(v) + \beta(v) = \dfrac{\gamma_v(2)}{\gamma_{v-1}(1)} \\
\alpha(v) - \beta^{-1}(v)\,\pi(v) = \dfrac{(\alpha(v) + \beta(v))\,\gamma_v(1) - \gamma_v(0)}{\gamma_v(1) - (\alpha(v) + \beta(v))\,\gamma_{v-1}(0)}
\end{cases} \tag{4.15}$$

where $\pi(v) = \dfrac{\mu_4(v)}{\mu_4(v-1)}$ with $\mu_4(v) = E\{X_t^4(v)\}$. Now, setting $\beta(v) + \beta^{-1}(v)\,\pi(v) = \delta(v)$ where $\delta(v) = (\alpha(v) + \beta(v)) + \left(\beta^{-1}(v)\,\pi(v) - \alpha(v)\right)$, then we have $\beta^2(v) - \delta(v)\,\beta(v) + \pi(v) = 0$. Hence, if $\delta(v) \geq 2\sqrt{\pi(v)}$ we set

$$\beta(v) = \frac{\delta(v)}{2} - \sqrt{\frac{\delta^2(v)}{4} - \pi(v)} \tag{4.16}$$

so that $0 < \beta(v)$ and $\prod_{v=1}^{s} \beta(v) < 1$ (which corresponding to the invertibility condition of the $PARMA(1,1)$ model) because (4.7). The above expressions can now be used to obtain estimators of the parameter $\underline{\theta}(v)$. First, we can estimate $\alpha(v) + \beta(v)$ by $(\widehat{\alpha(v) + \beta}(v)) := \dfrac{\widehat{\gamma}_v(2)}{\widehat{\gamma}_{v-1}(1)}$ and $\mu_4(v)$ by $\widehat{\mu}_4(v) = \frac{1}{N}\sum_{t=0}^{N-1} X_{st+v}^4$. Second, substituting these estimators into (4.15), (4.16), we obtain $\left(\widehat{\alpha(v) - \beta^{-1}}(v)\,\pi(v)\right) := \dfrac{(\widehat{\alpha(v) + \beta}(v))\widehat{\gamma}_v(1) - \widehat{\gamma}_v(0)}{\widehat{\gamma}_v(1) - (\widehat{\alpha(v) + \beta}(v))\widehat{\gamma}_{v-1}(0)}$ and

$$\begin{cases} \widehat{\beta}(\upsilon) := \frac{\widehat{\delta}(\upsilon)}{2} - \sqrt{\frac{\widehat{\delta}^2(\upsilon)}{4} - \widehat{\pi}(\upsilon)} \\ \widehat{\alpha}(\upsilon) := (\widehat{\alpha(\upsilon) + \beta}(\upsilon)) - \widehat{\beta}(\upsilon) \\ \widehat{\omega}(\upsilon) := \widehat{\mu}_2(\upsilon) - (\widehat{\alpha(\upsilon) + \beta}(\upsilon))\widehat{\mu}_2(\upsilon - 1). \end{cases} \tag{4.17}$$

As already mentioned by Kristensen and Linton [32], this method may lead to $(\widehat{\alpha(\upsilon) + \beta}(\upsilon)) < 0$ or $(\widehat{\alpha(\upsilon) + \beta}(\upsilon)) > 1$, $\upsilon \in \{1, ..., s\}$. However, the estimators can be censored at zero and one or at ϵ and $1 - \epsilon$ for small positive ϵ.

In order to derive the asymptotic properties of our estimators $\underline{\widehat{\theta}}(\upsilon) := \left(\widehat{\alpha}(\upsilon), \widehat{\beta}(\upsilon), \widehat{\omega}(\upsilon)\right)'$ and to construct confidence intervalls for $\underline{\theta}$, we assume that

A1. $(X_t)_{t\in\mathbb{Z}}$ is a SPS process and $\lambda_{(2)} := \prod_{i=1}^{s} E\left\{(\alpha(i)\, e_0^2 + \beta(i))^2\right\} < 1$

A2. $\lambda_{(4)} := \prod_{i=1}^{s} E\left\{(\alpha(i)\, e_0^2 + \beta(i))^4\right\} < 1.$

As already seen in Theorem 7, the moment condition in **A1** is necessary and sufficient for the $PGARCH$ model (4.2) to have a SPS solution with a $4th$ order moment (this rules out mildly the $IPGARCH$ models). The moment condition in **A2** is imposed in order that $\sqrt{N}\underline{\widehat{\theta}}(\upsilon)$ is asymptotically normal distributed.

Lemme 11 *Consider the $PARMA$ representation (4.8) of the PGARCH model (4.2) and let $(\underline{X}_t^2)_{t\in\mathbb{Z}}$ be the associated vectorial representation satisfying (4.9). Then, under **A1**-**A2** we have for each $\upsilon \in \{1, ..., s\}$*

1. *The estimator $\underline{\widehat{\theta}}(\upsilon)$ of $\underline{\theta}(\upsilon)$ is strongly consistent.*

2. $\sqrt{N}\left(\underline{\widehat{\theta}}(\upsilon) - \underline{\theta}(\upsilon)\right) \rightsquigarrow N\left(\underline{0}, \Sigma_{as}(\upsilon)\right)$ *where (if $\Sigma_{as}(\upsilon)$ is positive definite)* $\Sigma_{as}(\upsilon) := A(\upsilon)B(\upsilon)\widetilde{\Sigma}_{as}(\upsilon)B'(\upsilon)A'(\upsilon)$ *with*

$$B(\upsilon) := \begin{pmatrix} 1 & 0 & 0 \\ \dfrac{\gamma_\upsilon^2(1) - \gamma_{\upsilon-1}(0)\gamma_\upsilon(0)}{\beta^2(\upsilon)\left(\sigma_\eta^2(\upsilon-1)\right)^2} & 1 & 0 \\ -\mu_2(\upsilon-2) & 0 & 1 \end{pmatrix},$$

$$A(\upsilon) := \begin{pmatrix} 1 - a(\upsilon) & -a(\upsilon) & 0 \\ a(\upsilon) & a(\upsilon) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and $\widetilde{\Sigma}_{as}(\upsilon) := E\left\{(\eta_t(\upsilon) - \beta(\upsilon)\eta_t(\upsilon-1))^2 \underline{Z}_t(\upsilon)\underline{Z}_t'(\upsilon)\right\}$ *where*

$$\underline{Z}_t(\upsilon) = \left(\begin{array}{c} \dfrac{(X_t^2(\upsilon-2) - \mu_2(\upsilon-2))}{\gamma_{\upsilon-1}(1)} \\ \dfrac{X_t^2(\upsilon) - \mu_2(\upsilon) - \left(\alpha(\upsilon) - \beta^{-1}(\upsilon)\pi(\upsilon)\right)(X_t^2(\upsilon-1) - \mu_2(\upsilon-1))}{\gamma_\upsilon(1) - (\alpha(\upsilon) + \beta(\upsilon))\gamma_{\upsilon-1}(0)} \\ 1 \end{array}\right),$$

$$a(\upsilon) = \frac{1}{2} - \frac{\beta(\upsilon) + \beta^{-1}(\upsilon)\pi(\upsilon)}{4\sqrt{\dfrac{\left(\beta(\upsilon) + \beta^{-1}(\upsilon)\pi(\upsilon)\right)^2}{4} - \pi(\upsilon)}}$$

Proof. The strong consistency follows from Propositions 8 and 9. To show the asymptotic normality we use the same approach as Maercker and Moser [36]. We split the proof in two steps, in the first step, we will prove joint asymptotic normality of

$$\begin{aligned} (\widehat{\alpha(\upsilon) + \beta}(\upsilon)) &= \frac{\widehat{\gamma}_\upsilon(2)}{\widehat{\gamma}_{\upsilon-1}(1)} \\ \left(\alpha(\upsilon) - \widetilde{\beta^{-1}}(\upsilon)\pi(\upsilon)\right) &= \frac{(\alpha(\upsilon) + \beta(\upsilon))\widehat{\gamma}_\upsilon(1) - \widehat{\gamma}_\upsilon(0)}{\widehat{\gamma}_\upsilon(1) - (\alpha(\upsilon) + \beta(\upsilon))\widehat{\gamma}_{\upsilon-1}(0)} \\ \widetilde{\omega}(\upsilon) &= \widehat{\mu}_2(\upsilon) - (\alpha(\upsilon) + \beta(\upsilon))\widehat{\mu}_2(\upsilon-1). \end{aligned}$$

From (4.14) we have

$$\begin{aligned} &\sqrt{N}(\widetilde{\omega}(\upsilon) - \omega(\upsilon)) \\ =\ &\sqrt{N}(\widehat{\mu}_2(\upsilon) - \mu_2(\upsilon) - (\alpha(\upsilon) + \beta(\upsilon))(\widehat{\mu}_2(\upsilon-1) - \mu_2(\upsilon-1))) \\ =\ &\frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}(\eta_t(\upsilon) - \beta(\upsilon)\eta_t(\upsilon-1)). \end{aligned}$$

Furthermore

$$\begin{aligned} &\sqrt{N}\left((\widehat{\alpha(\upsilon) + \beta}(\upsilon)) - (\alpha(\upsilon) + \beta(\upsilon))\right) \\ =\ &\frac{\widehat{\gamma}_{\upsilon-1}^{-1}(1)}{\sqrt{N}}\sum_{t=0}^{N-1} X_t^2(\upsilon-2) \\ &\times\left(X_t^2(\upsilon) - \widehat{\mu}_2(\upsilon) - (\alpha(\upsilon) + \beta(\upsilon))\left(X_t^2(\upsilon-1) - \widehat{\mu}_2\ (\upsilon-1)\right)\right) \\ =\ &\frac{\widehat{\gamma}_{\upsilon-1}^{-1}(1)}{\sqrt{N}}\sum_{t=0}^{N-1} X_t^2(\upsilon-2) \\ &\times\left(\eta_t(\upsilon)\ - \beta(\upsilon)\eta_t(\upsilon-1) - \frac{1}{N}\ \sum_{t=0}^{N-1}(\eta_t(\upsilon) - \beta(\upsilon)\eta_t(\upsilon-1))\right) \\ =\ &\frac{\widehat{\gamma}_{\upsilon-1}^{-1}(1)}{\sqrt{N}}\sum_{t=0}^{N-1}(\eta_t(\upsilon) - \beta(\upsilon)\eta_t(\upsilon-1))\left(X_t^2(\upsilon-2) - \widehat{\mu}_2(\upsilon-2)\right) \end{aligned}$$

In a similar way we get

$$\widehat{\gamma}_v(1) - (\alpha(v) + \beta(v))\,\widehat{\gamma}_{v-1}(0) = \frac{1}{N}\sum_{t=0}^{N-1} (\eta_t(v) - \beta(v)\,\eta_t(v-1)) \times \left(X_t^2(v-1) - \widehat{\mu}_2(v-1)\right)$$

$$\widehat{\gamma}_v(0) - (\alpha(v) + \beta(v))\,\widehat{\gamma}_v(1) = \frac{1}{N}\sum_{t=0}^{N-1} (\eta_t(v) - \beta(v)\,\eta_t(v-1)) \left(X_t^2(v) - \widehat{\mu}_2(v)\right)$$

and therefore

$$\begin{aligned} &\widehat{\gamma}_v(0) - (\alpha(v) + \beta(v))\,\widehat{\gamma}_v(1) - (\widehat{\gamma}_v(1) - (\alpha(v) + \beta(v))) \\ &\times\widehat{\gamma}_{v-1}(0)\left(\alpha(v) - \beta^{-1}(v)\,\pi(v)\right) \\ =\; &\frac{1}{N}\sum_{t=0}^{N-1} (\eta_t(v) - \beta(v)\,\eta_t(v-1)) \\ &\times\left(X_t^2(v) - \widehat{\mu}_2(v) - \left(\alpha(v) - \beta^{-1}(v)\,\pi(v)\right)\left(X_t^2(v-1) - \widehat{\mu}_2(v-1)\right)\right) \end{aligned}$$

we conclude that

$$\begin{aligned} &\sqrt{N}\left(\left(\alpha(v) - \widetilde{\beta^{-1}}(v)\,\pi(v)\right) - \left(\alpha(v) - \beta^{-1}(v)\,\pi(v)\right)\right) \\ =\; &\left(\widehat{\gamma}_v(1) - (\alpha(v) + \beta(v))\,\widehat{\gamma}_{v-1}(0)\right)^{-1}\frac{1}{\sqrt{N}}\sum_{t=0}^{N-1} (\eta_t(v) - \beta(v)\,\eta_t(v-1)) \\ &\left(X_t^2(v) - \widehat{\mu}_2(v) - \left(\alpha(v) - \beta^{-1}(v)\,\pi(v)\right)\left(X_t^2(v-1) - \widehat{\mu}_2(v-1)\right)\right) \end{aligned}$$

By ergodicity, we have almost surely $\widehat{\mu}_2(v)$, $\widehat{\mu}_4(v)$ and $\widehat{\gamma}_v(.)$ converges to $\mu_2(v)$, $\mu_4(v)$ and $\gamma_v(.)$ respectively. In order to prove
$\sqrt{N}\left((\widehat{\alpha(v) + \beta}(v)) - (\alpha(v) + \beta(v)),\right.$
$\left.\left(\alpha(v) - \widetilde{\beta^{-1}}(v)\,\pi(v)\right) - \left(\alpha(v) - \beta^{-1}(v)\,\pi(v)\right), \widetilde{\omega}(v) - \omega(v)\right) \rightsquigarrow N\left(0, \widetilde{\Sigma}_{as}(v)\right)$
it is therfore sufficient to show that for any $\underline{\lambda} \in \mathbb{R}^3$

$$\frac{1}{\sqrt{N}}\sum_{t=0}^{N-1} (\eta_t(v) - \beta(v)\,\eta_t(v-1))\,\underline{\lambda}'\underline{Z}_t(v) \rightsquigarrow N\left(\underline{0}, \underline{\lambda}'\widetilde{\Sigma}_{as}(v)\underline{\lambda}\right) \tag{4.18}$$

using the Cramer-Wold device and an application of the CLT for martingale difference (see [29]) now gives (4.18). The next step we note

$$\begin{aligned} &\widehat{\omega}(v) - \widetilde{\omega}(v) \\ =\; &-\left((\widehat{\alpha(v) + \beta}(v)) - (\alpha(v) + \beta(v))\right)\widehat{\mu}_2(v-1), \\ &\left(\alpha(v) - \widehat{\beta^{-1}}(v)\,\pi(v)\right) - \left(\alpha(v) - \widetilde{\beta^{-1}}(v)\,\pi(v)\right) \\ =\; &\frac{\left((\widehat{\alpha(v) + \beta}(v)) - (\alpha(v) + \beta(v))\right)\left(\widehat{\gamma}_v^2(1) - \widehat{\gamma}_{v-1}(0)\,\widehat{\gamma}_v(0)\right)}{\left(\widehat{\gamma}_v(1) - (\widehat{\alpha(v) + \beta}(v))\widehat{\gamma}_{v-1}(0)\right)\left(\widehat{\gamma}_v(1) - (\alpha(v) + \beta(v))\,\widehat{\gamma}_{v-1}(0)\right)}. \end{aligned}$$

Hence, the ergodic theorem and Slutsky's theorem implies
$\sqrt{N}\left(\left(\widehat{\alpha(v)+\beta}(v)\right)-\left(\alpha(v)+\beta(v)\right),\left(\alpha(v)-\widehat{\beta^{-1}}(v)\,\pi(v)\right)\right.$
$\left.-\left(\alpha(v)-\beta^{-1}(v)\,\pi(v)\right),\widehat{\omega}(v)-\omega(v)\right) \rightsquigarrow N\left(0,B(v)\widetilde{\Sigma}_{as}(v)B(v)'\right)$. Since
$\left(\alpha(v),\beta(v),\omega(v)\right)=T_{\pi(v)}\left(\alpha(v)+\beta(v),\beta^{-1}(v)\,\pi(v)-\alpha(v),\omega(v)\right)$ where

$$T_{\pi}(x,y,z):=\begin{pmatrix} x-\dfrac{x+y}{2}+\sqrt{\dfrac{(x+y)^2}{4}-\pi} \\ \dfrac{x+y}{2}-\sqrt{\dfrac{(x+y)^2}{4}-\pi} \\ z \end{pmatrix},$$

then by application of the delta method, the result follows. ■

Theorem 12 *Consider the PGARCH $(1,1)$ Model (4.2) and let $(\underline{X}_t^2)_{t\in\mathbb{Z}}$ be the associated VARMA Representation (4.8). Then, under **A1-A2** $\sqrt{N}\left(\widehat{\underline{\theta}}-\underline{\theta}\right)\rightsquigarrow N\left(\underline{0},\Sigma_{as}(\underline{\theta})\right)$ where $\Sigma_{as}(\underline{\theta}):=A(\underline{\theta})B(\underline{\theta})\widetilde{\Sigma}_{as}(\underline{\theta})B'(\underline{\theta})A'(\underline{\theta})$ with $\widetilde{\Sigma}_{as}(\underline{\theta})$ (respectively $A(\underline{\theta})$ and $B(\underline{\theta})$)) is $3s\times 3s$ symmetric covariance diagonal bloc matrices with the v-th bloc being $\widetilde{\Sigma}_{as}(v)$ (respectively $A(v)$ and $B(v)$).*

Proof. The proof follows essentially from the vectorial representation of $\left(\widehat{\underline{\theta}}(v)-\underline{\theta}(v)\right)_{1\leq v\leq s}$ and the Lemma 11. ■

Remark 13 *The fact that the matrices $A(\underline{\theta})$, $B(\underline{\theta})$, and $\widetilde{\Sigma}_{as}(\underline{\theta})$ are $s-$block diagonal implies the asymptotic independence of the estimates for each regime $1\leq v\leq s$. This is not surprising result in periodic time-varying models as in PARMA processes.*

Remark 14 *For the Gaussian PARMA models, it is well known that the Yule-Walker estimates are asymptotically most efficient, because their asymptotic covariance matrices are the inverse of the corresponding Fisher information matrices. In PGARCH models, this property is not true, in spite of that, the last admits a PARMA representation, this is due to Heteroscedasticity of the process.*

4.4.1 The Wald test statistic

As an application of Theorem 12, we consider the problem of testing a null hypothesis $H0$ against an alternative hypothesis $H1$ of the form $H0: R\underline{\theta}=\underline{\theta}_0$, $H1: R\underline{\theta}\neq\underline{\theta}_0$, where R is a given $r\times 3s$ matrix of rank $r\leq 3s$, and $\underline{\theta}_0$ is a given

$r \times 1$ vector. Under the null hypothesis $H0$ and the conditions of Theorem 12, $\sqrt{N}\left(R\widehat{\underline{\theta}} - \underline{\theta}_0\right) \rightsquigarrow \mathcal{N}(\underline{0}, R\Sigma_{as}(\underline{\theta})R')$ and if the matrix $\Sigma_{as}(\underline{\theta})$ is nonsingular then the asymptotic variance matrix involved is nonsingular and thus we have

Theorem 15 *Under the conditions of Theorem 12, the additional condition that $\Sigma_{as}(\underline{\theta})$ is nonsingular with R of rank r, then under $H0$*

$$\widehat{W}(\underline{\theta}) := N\left(R\widehat{\underline{\theta}} - \underline{\theta}_0\right)'\left(R\ \Sigma_{as}(\underline{\theta})\ R'\right)^{-1}\left(R\widehat{\underline{\theta}} - \underline{\theta}_0\right) \rightsquigarrow \chi^2_{(r)}. \tag{4.19}$$

On the other hand, under the alternative hypothesis $H1$ we have in probability

$$\lim_{N\to\infty} \frac{\widehat{W}(\underline{\theta})}{N} := (R\underline{\theta} - \underline{\theta}_0)'\left(R\Sigma_{as}(\underline{\theta})\ R'\right)^{-1}(R\underline{\theta} - \underline{\theta}_0) > 0. \tag{4.20}$$

We first note that the statistics $\widehat{W}(\underline{\theta})$ and $\widehat{W}\left(\widehat{\underline{\theta}}\right)$ have asymptotically the same distribution as $N \to \infty$ i.e., $\widehat{W}\left(\widehat{\underline{\theta}}\right) \rightsquigarrow \chi^2_{(r)}$. Now, the statistic $\widehat{W}\left(\widehat{\underline{\theta}}\right)$ is the test statistic of the Wald test of the null hypothesis $H0$. Given the size $\alpha \in]0,1[$, choose a critical value β, so that under the null hypothesis $H0$, $P\left(\widehat{W}\left(\widehat{\underline{\theta}}\right) > \beta\right) \longrightarrow \alpha$. Then the null hypothesis is accepted if $\widehat{W}\left(\widehat{\underline{\theta}}\right) \leq \beta$ and rejected in favor of the alternative hypothesis if $\widehat{W}\left(\widehat{\underline{\theta}}\right) > \beta$. This test is consistent due to (4.20). In the case when R is a raw vector, so $\underline{\theta}_0$ is a scalar, we can modify (4.19) to $\widehat{t} = \sqrt{N}\left(R\ \Sigma_{as}\left(\widehat{\underline{\theta}}\right) R'\right)^{-1/2}\left(R\widehat{\underline{\theta}} - \underline{\theta}_0\right)$ so $\widehat{t} \rightsquigarrow \mathcal{N}(0,1)$, whereas under the alternative hypothesis $H1$ we have in probability $\lim_{N\to\infty} \frac{\widehat{t}}{\sqrt{N}} = \left(R\Sigma_{as}(\underline{\theta})\ R'\right)^{-1/2}(R\underline{\theta} - \underline{\theta}_0) \neq 0$. These results can be used to construct a two-sided or a one sided tests.

4.5 Numerical illustrations and bootstrap comparison

In this section, we examine the performance of the finite sample properties of the Yule-Walker type estimators by comparing it with the LSE using the Monte Carlo study.

1. First, we simulate a periodic $GARCH(1,1)$ process with period $s = 2$ given by (4.2) where $(e_t)_{t\in\mathbb{Z}}$ is an *i.i.d.* Gaussian process with zero mean and variance $\sigma^2 = 1$ for four different sets of parameter values. For each choice of parameter values, we simulate 1000 data sets with length $N \in$

$\{1000, 1500, 2000\}$. For each trajectory, $\underline{\theta}$ has been estimated by Yule-Walker $\widehat{\underline{\theta}}^{(YW)}$ and by the least squares $\widehat{\underline{\theta}}^{(LS)}$ methods. Replacing the unknown parameters by their estimates, we obtain an estimate $\widehat{\Sigma}_{as}(\upsilon)$ for $\Sigma_{as}(\upsilon)$, $\upsilon \in \{1, ..., s\}$. We denote by

$$\sqrt{Var_{as}(\widehat{\underline{\theta}}^{(YW)}(\upsilon))_i} = \frac{1}{\sqrt{N}}\sqrt{\left(\widehat{\Sigma}_{as}^{(YW)}(\upsilon)\right)_{ii}} i = 1, 2, 3$$

the estimate of the standard deviation of $(\underline{\theta}(\upsilon))_i$. In order to demonstrate that this estimate, although based on the asymptotic theory, can be successfully applied to finite samples of reasonable size, the mean of $Var_{as}((\widehat{\underline{\theta}}^{(YW)}(\upsilon))_i$ over 1000 simulations has been compared with the mean of $\left((\underline{\theta}(\upsilon))_i - (\widehat{\underline{\theta}}^{(YW)}(\upsilon))_i\right)^2$ over 1000 simulations, denoted by $MSE^{(WY)}$. Now, let us consider the hypothesis $H_0^{(i)}(\upsilon, \upsilon') : (\underline{\theta}(\upsilon))_i = (\underline{\theta}(\upsilon'))_i$ for $\upsilon, \upsilon' \in \{1, .., s\}$ and $i = 1, 2, 3$. Then, if a Wald test is used, the hypothesis $H_0(\upsilon, \upsilon')$ is rejected when

$$N\left((\widehat{\underline{\theta}}^{(YW)}(\upsilon))_i - (\widehat{\underline{\theta}}^{(YW)}(\upsilon'))_i\right)^2 \left(R\Sigma_{as}\left(\widehat{\underline{\theta}}^{(YW)}\right)R'\right)^{-1}$$

is greater than 95% quantile of χ_1^2 distribution. Similar notations for the least squares estimate $\widehat{\underline{\theta}}^{(LS)}$.

2. Once the parameter $\underline{\theta}$ is estimated, we naturally want to know how efficient it is as an estimator of θ. For this purpose, the so-called residuals bootstrap method (see [45]) can be used as an alternative to the conventional method of finding sampling distribution. The residuals bootstrap replicates can be obtained (briefly) from the following. Define the residual

$$\widetilde{e}_t(\upsilon) = \frac{X_t(\upsilon)}{\widetilde{h}_t(\upsilon)} \text{ and } \widetilde{h}_t^2(\upsilon) := \widehat{\omega}(\upsilon) + \widehat{\alpha}(\upsilon)X_t^2(\upsilon-1) + \widehat{\beta}(\upsilon)\widetilde{h}_t^2(\upsilon-1),$$

$t = 1, ..., N, \upsilon = 1, ..., s$ and let $\widehat{e}_t$ be the standardized version of $\widetilde{e}_t$ such that $\frac{1}{N}\sum_{t=0}^{N-1}\widetilde{e}_t = 0$ and $\frac{1}{N}\sum_{t=0}^{N-1}\widetilde{e}_t^2 = 1$. Now we draw e_t^*, $t = 1, ..., N$ with replacement from $\widehat{e}_t$ and define

$$X_t^*(\upsilon) = h_t^*(\upsilon)e_t^*(\upsilon) \text{ and } h_t^{*2}(\upsilon) := \widehat{\omega}(\upsilon) + \widehat{\alpha}(\upsilon)X_t^{*2}(\upsilon-1) + \widehat{\beta}(\upsilon)h_t^{*2}(\upsilon-1),$$

$t = 1, ..., N$ with starting values $X_t^*(1) = h_t^*(1) = e_t^*(1)$. Noting here that the choice of initial values does not matter for the asymptotic properties. However, it may have importance from a practical point of view. Once we have the bootstrap replicate, we need to estimate its parameter $\underline{\widehat{\theta}} = \left(\underline{\widehat{\theta}}'(1), ..., \underline{\widehat{\theta}}'(s)\right)'$ by Yule-Walker estimator $\underline{\widehat{\theta}}_B$ as solution to the moment-type equations (4.17). We repeat the simulation process several $L-$times to estimate the distribution of $\underline{\widehat{\theta}}_B$. For the purpose of comparison, the columns of the next tables have been bisected. Hence, the column $\underline{\widehat{\theta}}^{(YW)}$ corresponds to the Yule-Walker inference results, the next corresponds to the bootstrap inference results. For purpose of comparison with the least squares estimate, we have executed the same procedure.

Table 1 (respectively 2 and 3) reports the squared biais (respectively variance and MSE) of $\underline{\widehat{\theta}}^{(YW)}$ and $\underline{\widehat{\theta}}^{(LS)}$ over N simulations and the bootstrap approximation $\underline{\widehat{\theta}}_B^{(YW)}$ and $\underline{\widehat{\theta}}_B^{(LS)}$ over 1000 replications. The results reported in Tables $1, 2$ and 3 are in accordance with the asymptotic theory. It is clear that the bootstrap estimates $\underline{\widehat{\theta}}_B^{(YW)}$ and $\underline{\widehat{\theta}}_B^{(LS)}$ are very close to the corresponding estimates $\underline{\widehat{\theta}}^{(YW)}$ and $\underline{\widehat{\theta}}^{(LS)}$ and the $MSE^{(YW)}$ and $MSE^{(LS)}$ is almost equal to the bootstrap estimate $MSE_B^{(YW)}$ and $MSE_B^{(LS)}$. It is worth noting that the LSE clearly outperforms the YW estimator, indeed, Table 2 shows that $MSE^{(LS)}$ and $MSE_B^{(LS)}$ are smaller than $MSE^{(YW)}$ and $MSE_B^{(YW)}$ respectively. The asymptotic validity for the bootstrap can also be verified numerically by looking at how the approximate distribution $\sqrt{N}\left(\underline{\widehat{\theta}}_B - \underline{\widehat{\theta}}\right)$ behaves for the bootstrap estimates. The χ^2 test supports the observation that the bootstrap histograms are normally distributed. This result shows numerically that the bootstrap method is asymptotically valid for the Yule-Walker

N	YW		LSE	
	$10^{-3}\widehat{\theta}^{(YW)}$	$10^{-3}\ \widehat{\theta}_B^{(YW)}$	$10^{-3}\widehat{\underline{\theta}}^{(LS)}$	$10^{-3}\widehat{\underline{\theta}}_B^{(LS)}$
1000	0.00, 13.0, 13.0	0.00, 12.6, 12.6	0.00, 10.7, 10.7	1.50, 8.40, 9.90
	0.10, 28.5, 28.5	0.00, 27.5, 27.5	0.20, 30.0, 30.2	2.30, 33.9, 36.1
	0.36, 23.3, 23.6	0.00, 22.3, 22.3	0.00, 3.10, 3.10	0.80, 2.20, 3.00
	6.20, 115.2, 121.3	0.00, 12.0, 12.0	0.00, 7.90, 7.90	2.90, 5.30, 8.20
	0.10, 35.0, 35.1	0.00, 35.9, 35.9	0.00, 41.2, 41.2	0.10, 58.2, 58.2
	2.40, 100.5, 102.8	0.00, 100.6, 100.5	1.80, 192.9, 194.6	2.30, 391.9, 394.7
1500	0.00, 8.60, 8.60	0.00, 8.40, 8.40	0.00, 7.20, 7.20	1.30, 5.10, 6.40
	0.00, 17.6, 17.6	0.00, 17.3, 17.3	0.10, 16.8, 16.9	1.60, 12.9, 14.5
	0.00, 8.90, 8.90	0.00, 8.70, 8.60	0.00, 2.20, 2.20	0.70, 1.60, 2.30
	2.20, 64.3, 66.4	2.00, 62.7, 62.7	0.00, 5.90, 5.90	2.50, 4.00, 6.50
	0.00, 29.2, 29.2	0.00, 29.3, 29.2	0.00, 27.9, 27.9	0.10, 34.5, 34.5
	1.00, 78.1, 79.0	0.00, 79.2, 79.1	0.50, 106.9, 107.3	1.20, 138.8, 139.9
2000	0.00, 6.60, 6.60	0.00, 6.50, 6.50	0.00, 5.60, 5.60	1.10, 3.80, 5.00
	0.00, 12.7, 12.7	0.00, 12.6, 12.6	0.00, 11.7, 11.8	1.50, 8.70, 10.2
	0.00, 0.70, 0.70	0.00, 6.90, 6.90	0.00, 1.80, 1.80	0.70, 1.20, 1.90
	0.60, 36.2, 36.7	0.00, 32.4, 32.4	0.00, 4.70, 4.70	2.20, 3.20, 5.40
	0.00, 23.3, 23.2	0.00, 23.2, 23.2	0.00, 20.9, 20.8	0.00, 24.7, 24.7
	0.30, 67.9, 68.2	0.00, 69.4, 69.3	0.20, 75.3, 75.4	0.90, 92.0, 92.8

Table 1 : MSE of Yule-Walker and LSE estimator for $(\underline{\omega}, \underline{\alpha}, \underline{\beta})' = ((0.2, 0.2), (0.15, 0.25), (0.25, 0.50))$

Notes : The nine line of each cellare,from left to right squared biais, variance and MSE

N	YW		LSE	
	$10^{-3}\widehat{\theta}^{(YW)}$	$10^{-3}\ \widehat{\theta}_B^{(YW)}$	$10^{-3}\widehat{\underline{\theta}}^{(LS)}$	$10^{-3}\widehat{\underline{\theta}}_B^{(LS)}$
1000	3.90, 2596, 2599 0.60, 2465, 2465 0.00, 129.2, 129.1 19.8, 198.9, 218.6 0.30, 72.3, 72.6 20.5, 70.2, 90.6	2.30, 2229, 2229 0.70, 1878, 1877 0.20, 116.7, 116.8 0.00, 187.1, 187.0 0.00, 76.3, 76.3 0.00, 67.0, 67.0	2.20, 2031, 2031 1.20, 1934, 1934 0.10, 55.6, 55.7 0.00, 28.1, 28.2 0.10, 153.0, 153.9 0.00, 131.7, 131.6	57.0, 585.6, 642.6 36.6, 774.2, 810.8 6.80, 23.8, 30.6 2.20, 14.2, 16.4 0.10, 69.4, 69.4 0.30, 124.0, 124.2
1500	1.20, 2251, 2250 1.40, 1635, 1635 0.00, 110.7, 110.6 15.2, 144.7, 159.7 0.00, 71.6, 71.6 16.9, 63.6, 80.4	0.80, 1981, 1980 0.60, 1345, 1345 0.10, 106.4, 106.3 0.00, 140.5, 140.4 0.00, 69.6, 69.5 0.00, 60.7, 60.6	0.00, 2306, 2306 2.10, 1673, 1675 0.10, 51.6, 51.6 0.10, 25.9, 26.0 0.10, 106.7, 106.7 0.00, 111.7, 111.6	46.4, 516.2, 562.2 42.3, 429.8, 471.7 8.00, 19.1, 27.0 2.20, 11.4, 13.6 0.10, 59.0, 59.0 0.20, 72.5, 72.6
2000	0.40, 1882, 1881 4.10, 1636, 1639 0.00, 101.1, 101.1 11.9, 144.5, 156.2 0.10, 60.0, 60.0 14.9, 62.9, 77.7	0.80, 1647, 1646 0.50, 1265, 1264 0.10, 95.3, 95.3 0.20, 128.7, 128.8 0.00, 58.6, 58.6 0.10, 61.5, 61.5	0.60, 1787, 1786 4.60, 1629, 1631 0.00, 48.2, 48.2 0.20, 23.8, 23.9 0.30, 109.3, 109.5 0.00, 107.1, 107.0	39.5, 463.8, 502.9 41.7, 360.9, 402.2 7.60, 18.2, 25.8 2.10, 9.90, 12.0 0.00, 48.2, 48.1 0.20, 57.0, 57.2

Table 2 : MSE of Yule-Walker and LSE estimator for $(\underline{\omega}, \underline{\alpha}, \underline{\beta})' = ((1.0, 1.0), (0.50, 0.25), (0.25, 0.50))$

Notes : The nine line of each cellare,from left to right squared biais, variance and MSE

N	YW		LSE	
	$10^{-3}\widehat{\theta}^{(YW)}$	$10^{-3}\ \widehat{\theta}_B^{(YW)}$	$10^{-3}\widehat{\underline{\theta}}^{(LS)}$	$10^{-3}\widehat{\underline{\theta}}_B^{(LS)}$
1000	1.10, 159.5, 160.6	0.00, 138.4, 138.4	0.10, 68.4, 68.4	0.20, 12.9, 13.1
	0.10, 206.2, 206.2	0.00, 165.1, 165.0	0.00, 46.2, 46.2	0.30, 1034, 1034
	0.00, 165.2, 165.0	0.00, 159.4, 159.4	0.70, 74.1, 74.7	2.00, 43.0, 45.0
	22.6, 256.1, 287.5	0.00, 252.9, 252.8	0.20, 60.0, 60.1	1.40, 36.3, 37.7
	1.00, 58.5, 59.5	0.00, 58.7, 58.6	0.10, 220.0, 219.9	0.10, 83.5, 83.5
	24.6, 92.6, 117.1	0.00, 91.4, 91.4	0.10, 217.4, 217.3	5.30, 6526, 6526
1500	0.40, 159.1, 159.4	0.00, 124.9, 124.8	0.00, 131.8, 131.7	0.10, 13.7, 13.8
	0.20, 148.1, 148.3	0.00, 110.6, 110.6	0.10, 42.7, 42.7	7.10, 5024, 5026
	0.10, 151.9, 151.9	0.10, 142.2, 142.2	0.60, 73.3, 73.8	2.10, 40.30, 42.4
	12.4, 222.2, 234.6	0.40, 204.9, 205.1	0.40, 57.3, 57.7	1.30, 33.2, 34.5
	0.80, 57.9, 58.7	0.00, 55.9, 55.9	1.10, 469.6, 470.3	0.60, 76.4, 76.9
	26.1, 76.4, 102.5	0.10, 74.9, 74.9	0.10, 196.7, 196.5	58.4, 5621, 5621
2000	0.20, 137.4, 137.5	0.00, 111.4, 111.3	0.20, 383.3, 383.5	0.00, 10.8, 10.8
	0.30, 187.9, 188.1	0.00, 127.8, 127.8	0.10, 34.6, 34.8	0.30, 10.4, 10.7
	0.00, 127.3, 127.3	0.00, 119.2, 119.2	0.30, 68.2, 68.5	2.70, 35.4, 38.1
	17.8, 255.3, 273.0	0.20, 218.2, 218.2	0.50, 53.8, 54.3	1.50, 28.4, 30.0
	0.50, 57.3, 57.8	0.00, 57.0, 55.0	1.90, 759.5, 760.7	1.10, 63.3, 64.4
	22.5, 75.7, 98.2	0.00, 72.1, 72.0	0.00, 159.3, 159.3	0.00, 100.9, 100.8

Table 3 : MSE of Yule-Walker and LSE estimator for $(\underline{\omega}, \underline{\alpha}, \underline{\beta})' = ((0.2, 0.2), (0.50, 0.35), (0.25, 0.50))$

Notes: The nine line of each cellare,from left to right squared biais, variance and MSE.

For this designe neither the $4th$ nor $8th$ moment exist

Chapter 5

The LAN properities for $PARCH$ processes

Abstract: In this chapter, we continue to investigate in asymptotic inference for $PARCH$ processes by considering the asymptotic efficiency of the conditional least squares (CLS) estimators based on LAN approach.

5.1 Introduction

We consider a time series $(\epsilon_t, t \in \mathbb{Z} = \{0, \pm 1, \pm 2, ...\})$ exhibiting changes in regimes at known dates. Suppose that we have s regimes. Let $s_n := \sum\limits_{v=1}^{s} v \mathbb{I}_{\Delta(v)}(n)$ where $\mathbb{I}_{\Delta}$ denotes the indicator function of a set Δ and $\Delta(v) := \{n | n = st + v\}$ be the regime corresponding to index n, so $s_n = v$ when the time series is in regime v at time n for $v = 1, ..., s$. Given s_n, it is supposed that the dynamics in each regime can be described by an $ARCH\,(q)$ equation. Thus we have

$$\epsilon_n = \sqrt{h_n} e_n \text{ and } h_n = w\,(s_n) + \sum_{i=1}^{q} \alpha_i\,(s_n)\,\epsilon_{n-i}^2,\ n \in \mathbb{Z}. \tag{5.1}$$

where $(e_n, n \in \mathbb{Z})$ is a sequence of independent and identically distributed $(i.i.d.)$ random variables defined on a probability space $(\Omega, \mathcal{A}, P)$ such that $E\,\{e_n\} = 0, E\,\{e_n^2\} = 1$ and e_k is independent of ϵ_n for all $k > n$. The functions $w\,(s_n)$ and $\alpha_i\,(s_n)$ are such that $w\,(s_n) > 0, \alpha_i\,(s_n) \geq 0, i = 1, ..., q$, for all $n \in \mathbb{Z}$. By setting $n = st + v$, Model (5.1) may be equivalently written as

$$\epsilon_{st+v} = \sqrt{h_{st+v}} e_{st+v} \text{ and } h_{st+v} = w(v) + \sum_{i=1}^{q} \alpha_i(v) \epsilon_{st+v-i}^2,\ t \in \mathbb{Z} \tag{5.2}$$

highlighting thus the periodicity in the model which we will make heavy use of (5.2). It is easy to write Model (5.1) in term of the squared process as follows

$$\epsilon_n^2 = w(s_n)\, e_n^2 + \sum_{i=1}^{q} \alpha_i(s_n)\, e_n^2 \epsilon_{n-i}^2, \tag{5.3}$$

which is ready to be cast in a first-order stochastic recurrence equation with random coefficients. Indeed, defining the q-random vectors $\underline{\epsilon}_n = \left(\epsilon_n^2, ..., \epsilon_{n-q+1}^2\right)'$ and $\underline{b}(s_n) = \left(w(s_n)\, e_n^2, \underline{O}'_{(q-1)\times 1}\right)'$ together with the $q \times q$ random matrix $A(s_n)$ given by

$$A(s_n) := \begin{pmatrix} \alpha_1(s_n)\, e_n^2 \;\; \alpha_2(s_n)\, e_n^2 \ldots \alpha_{q-1}(s_n)\, e_n^2 & \alpha_q(s_n)\, e_n^2 \\ I_{(q-1)\times(q-1)} & \underline{O}_{(q-1)\times 1} \end{pmatrix}$$

one can rewrite Model (5.3) in the following generalized AR model

$$\underline{\epsilon}_n = A(s_n)\, \underline{\epsilon}_{n-1} + \underline{b}(s_n)$$

which differs from the standard formulation studied by Bougerol and Picard [16] in that the coefficients $(A(s_n), \underline{b}(s_n))$ are rather independent and periodically distributed ($i.p.d.$). It is well known, that with periodic coefficients, it is possible to embed seasons into a multivariate stationary process (see Bibi and Aknouche [9]). More precisely, $\underline{Y}_t = \left(\underline{\epsilon}'_{st+1}, \underline{\epsilon}'_{st+2}, ..., \underline{\epsilon}'_{st+s}\right)'$ is a $VRCA(1)$ process of the form

$$\underline{Y}_t = C_t \underline{Y}_{t-1} + \underline{B}_t,\ t \in \mathbb{Z} \tag{5.4}$$

where C_t and $\underline{B}_t$ are defined by blocks as

$$C_t := \begin{pmatrix} O_{q\times q} & \cdots & O_{q\times q} & A(st+1) \\ O_{q\times q} & \cdots & O_{q\times q} & A(st+2)A(st+1) \\ \vdots & \vdots & \vdots & \vdots \\ O_{q\times q} & \cdots & O_{q\times q} & \prod_{v=0}^{s-1} A(st+s-v) \end{pmatrix}_{qs\times qs},$$

$$\underline{B}_t := \begin{pmatrix} \underline{b}(st+1) \\ A(st+2)\underline{b}(st+1) + \underline{b}(st+2) \\ \vdots \\ \sum_{k=1}^{s} \left\{ \prod_{v=0}^{s-k-1} A(st+s-v) \right\} \underline{b}(st+k) \end{pmatrix}_{qs\times 1}.$$

However, Equation (5.2) has a periodically strictly stationary (SPS) solution in $\mathbb{L}_1$ if and only if (5.4) has strictly stationary solution in $\mathbb{L}_1$. Bibi and Aknouche

[8] have been analyzed the probabilistic properties of $P-GARCH$ process, such as geometric ergodicity and the strong mixing.These concepts are fundamental in central limit theorem and in the law of large numbers, which can be employed to derive asymptotic normality, consistency of maximum likelihood estimator and inference with the model. The key conditions of interest in determining the geometric ergodicity are summarized in the following assumption

A. $\rho\left(\prod_{v=1}^{s} A_v\right) < 1$ where $\rho(M)$ represents the maximum modulus of the eigenvalues of a squared matrix M.

B. the variable e_1 has a positive density f absolutely continuous with respect to the Lebesgue measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.

The Assumption **A.,** ensures the causality of the process $(\epsilon_t^2, t \in \mathbb{Z})$. Moreover, the solution process is unique, SPS, periodically ergodic (see [9]) and with periodic correlated (PC) structure in the sence that $Cov\left(\epsilon_{l+s}, \epsilon_{k+s}\right) = Cov\left(\epsilon_l, \epsilon_k\right)$ for all integer l, k. Hence, for $P-ARCH(1)$, the above condition reduce to $\prod_{v=1}^{s} \alpha_1(v) < 1$. It is worth noting that the existence of explosive regimes (i.e., $\alpha_1(v) > 1$) does not preclude the periodic second order stationarity of $(\epsilon_t, t \in \mathbb{Z})$. When associated with **B.**, Bibi and Aknouche [8] have showed that the process $(\underline{Y}_t, t \in \mathbb{Z})$ defined by (5.4) is geometrically ergodic, and if initialized from its invariant measure, $(\underline{Y}_t, t \in \mathbb{Z})$ is strictly stationary and β-mixing with exponential decay.

5.2 Conditional least squares estimator and efficiency

Let $\underline{\theta} = (\underline{\theta}_1', \underline{\theta}_2', ..., \underline{\theta}_s')'$ where $\underline{\theta}_i = (w(i), \alpha_1(i), ..., \alpha_q(i))'$, $i = 1, ..., s$, be the parameter vector which is supposed to belong to a compact space $\Theta \subset \left(]0, +\infty[^{2q}\right)^s$. The true parameter value is unknown and is denoted by $\underline{\theta}^0 = \left(\underline{\theta}_1^{0\prime}, \underline{\theta}_2^{0\prime}, ..., \underline{\theta}_s^{0\prime}\right)'$ with $\underline{\theta}_i^0 = \left(w^0(i), \alpha_1^0(i), ..., \alpha_q^0(i)\right)'$. Let $\{\epsilon_1, \epsilon_2, ..., \epsilon_n\}$ be a realization of length $n = Ns$ of the unique, causal and periodically strictly stationary solution $(\epsilon_t, t \in \mathbb{Z})$ to Model (5.2) with true parameter $\underline{\theta}^0 \in \Theta$, i.e., $\epsilon_{st+v} = \sqrt{h_{st+v}} e_{st+v}$ and

$$h_{st+v} = w^0(v) + \sum_{i=1}^{q} \alpha_i^0(v) \epsilon_{st+v-i}^2 = \underline{Z}_t'(v) \underline{\theta}_v^0 \tag{5.5}$$

where $\underline{Z}_t(v) = \left(1, \epsilon^2_{st+v-1}, ..., \epsilon^2_{st+v-q}\right)', v = 1, ..., s$. Beginning this section with a weak vectorial $ARCH$ representation of the square process $(\epsilon^2_t, t \in \mathbb{Z})$ which will be used frequently in the sequel. Set $\underline{\epsilon}^2_t = \left(\epsilon^2_{st+1}, ..., \epsilon^2_{st+s}\right)'$, $\underline{e}^2_t = \left(e^2_{st+1}, ..., e^2_{st+s}\right)'$ and $\underline{h}_t = (h_{st+1}, ..., h_{st+s})'$ then we have $\underline{\epsilon}^2_t = diag\left\{\underline{h}_t\right\}\underline{e}^2_t$. Defining $\mathcal{F}_t$ as the $\sigma-$field generated by $\left\{\underline{\epsilon}_{t-i}, i \geq 0\right\}$ we note $E\left\{\underline{\epsilon}^2_t|\mathcal{F}_{t-1}\right\} = diag\left\{\underline{h}_t\right\}\underline{1} = Z'_t\underline{\theta}$ where $\underline{1} := (1, ..., 1)' \in \mathbb{R}^s$, $Z'_t := diag\left\{\underline{Z}'_t(1), ..., \underline{Z}'_t(s)\right\}$. Conditionally on some initial values properly chosen, the CLS estimator $\widehat{\underline{\theta}}^{CLS}_n$ of $\underline{\theta}^0$ based on the square-transformed variables $\epsilon^2_{1-q}, ..., \epsilon^2_0, \epsilon^2_1, ..., \epsilon^2_n$ is any measurable solution of

$$\widehat{\underline{\theta}}^{CLS}_n := Arg\min_{\underline{\theta}\in\Theta}\widehat{Q}_n\left(\underline{\theta}\right)$$

where

$$\begin{aligned}\widehat{Q}_n\left(\underline{\theta}\right) &= \sum_{t=1}^{n}\left(\underline{\epsilon}^2_t - E\left\{\underline{\epsilon}^2_t|\mathcal{F}_{t-1}\right\}\right)'\left(\underline{\epsilon}^2_t - E\left\{\underline{\epsilon}^2_t|\mathcal{F}_{t-1}\right\}\right) \\ &= \sum_{t=1}^{n}\left(\underline{\epsilon}^2_t - Z'_t\underline{\theta}\right)'\left(\underline{\epsilon}^2_t - Z'_t\underline{\theta}\right).\end{aligned}$$

From the linear regression theory it is easily verified that

$$\widehat{\underline{\theta}}^{CLS}_n = \left(\sum_{t=1}^{n} Z_tZ'_t\right)^{-1}\sum_{t=1}^{n} Z_t\underline{\epsilon}^2_t. \tag{5.6}$$

Remark 1 *It will be shown that the choice of the initial values does not affect the asymtotic results regarding* $\widehat{\underline{\theta}}^{CLS}_n$. *Hence, in practice, the initial values* $\left\{\epsilon_{1-q}, ..., \epsilon_0\right\}$ *can be chosen as* $\epsilon_{1-q} = ... = \epsilon_0 = 0$.

In order to derive the asymptotic behavior of $\widehat{\underline{\theta}}^{CLS}_n$, we need the following assumption.

C. $E\left\{\|\underline{\epsilon}_t\|^4\right\} < \infty$.

The Assumption **C.,** ensures the existence of the finiteness of the fourth-order moment for the solution of Equations (5.2). Noting here that in $P - GARCH$ model, a necessary and sufficient condition for the existence of the fourth-order moment has been established by Bibi and Aknouche [9]. In particular for $P - ARCH\left(q\right)$, the condition $\rho\left(\left\{\prod_{v=1}^{s} E\left\{A^{\otimes 2}(st+v)\right\}\right\}\right) < 1$ implies that $E\left\{\epsilon^4_t\right\} < +\infty$. The following lemma gives the strong consistency and the limit distribution of $\widehat{\underline{\theta}}^{CLS}_n$.

Lemme 2 *Under Assumption **C.**, we have*

1. *almost surely* $\underline{\widehat{\theta}}_n^{CLS} \rightarrow \underline{\theta}$ *as* $n \rightarrow +\infty$.

2. $\sqrt{n}\left(\underline{\widehat{\theta}}_n^{CLS} - \underline{\theta}^0\right) \rightsquigarrow \mathcal{N}\left(\underline{0}, \vartheta^{-1}\left(\underline{\theta}^0\right)\mathcal{I}\left(\underline{\theta}^0\right)\vartheta^{-1}\left(\underline{\theta}^0\right)\right)$ *where* $\vartheta\left(\underline{\theta}^0\right)$ *and* $\mathcal{I}\left(\underline{\theta}^0\right)$ *are block matrices given by*

$$\begin{aligned}
\vartheta\left(\underline{\theta}^0\right) &= \frac{1}{s}\sum_{v=1}^{s} E_{\underline{\theta}^0}\left\{\underline{Z}_t(v)\underline{Z}_t'(v)\right\} \\
&= \frac{1}{s}\sum_{v=1}^{s} E_{\underline{\theta}^0}\left\{\frac{\partial h_{st+v}}{\partial \underline{\theta}_v}\frac{\partial h_{st+v}}{\partial \underline{\theta}_v'}\left(\underline{\theta}^0\right)\right\} \\
\mathcal{I}\left(\underline{\theta}^0\right) &= \frac{1}{s}\sum_{v=1}^{s} E_{\underline{\theta}^0}\left\{\left(e_0^4 - 1\right)\left(\underline{Z}_t'(v)\underline{\theta}_v^0\right)^2 \underline{Z}_t(v)\underline{Z}_t'(v)\right\} \\
&= \frac{1}{s}\sum_{v=1}^{s} E_{\underline{\theta}^0}\left\{\left(e_0^4 - 1\right)h_{st+v}^2\frac{\partial h_{st+v}}{\partial \underline{\theta}_v}\frac{\partial h_{st+v}}{\partial \underline{\theta}_v'}\left(\underline{\theta}^0\right)\right\}
\end{aligned}$$

Proof.

1. The result follows from Equation (5.6) and the ergodic theorem.

2. Note that

$$\begin{aligned}
\underline{\widehat{\theta}}_n^{CLS} - \underline{\theta}^0 &= \left(\sum_{t=1}^{n} Z_t Z_t'\right)^{-1}\left(\sum_{t=1}^{n} Z_t \underline{\epsilon}_t^2 - \sum_{t=1}^{n} Z_t Z_t'\underline{\theta}\right) \\
&= \left(\sum_{t=1}^{n} Z_t Z_t'\right)^{-1}\sum_{t=1}^{n} Z_t\left(\underline{\epsilon}_t^2 - Z_t'\underline{\theta}\right)
\end{aligned}$$

Consider $\frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}\underline{G}_t$ where $\underline{G}_t := \frac{1}{\sqrt{s}}\sum_{v=1}^{s}\underline{Z}_t(v)\left(\epsilon_{st+v}^2 - Z_t'(v)\underline{\theta}_v\right)$ is the $v-th$ component of $Z_t\left(\underline{\epsilon}_t^2 - Z_t'\underline{\theta}\right)$ which is a stationary ergodic zero mean martingale difference with

$$\begin{aligned}
Var_{\underline{\theta}^0}\left(\underline{G}_t\right) &= \frac{1}{s}\sum_{v=1}^{s} E_{\underline{\theta}^0}\left\{\underline{Z}_t(v)\,\underline{Z}_t'(v)\left(\epsilon_{st+v}^2 - Z_t'(v)\underline{\theta}_v\right)^2\right\} \\
&= \frac{1}{s}\sum_{v=1}^{s} E_{\underline{\theta}^0}\left\{\underline{Z}_t(v)\,\underline{Z}_t'(v)\left(Z_t'(v)\underline{\theta}_v\right)^2\left(e_{st+v}^2 - 1\right)^2\right\} \\
&= \frac{1}{s}\sum_{v=1}^{s} E_{\underline{\theta}^0}\left\{\underline{Z}_t(v)\,\underline{Z}_t'(v)\left(Z_t'(v)\underline{\theta}_v\right)^2\left(e_{st+v}^4 - 1\right)\right\}
\end{aligned}$$

Applying the central limit theorem for stationary ergodic martingale difference (see [12]) and the Gramèr-Wold device we find that $\frac{1}{\sqrt{n}}\sum_{t=1}^{n} Z_t\left(\underline{\epsilon}_t^2 - Z_t'\underline{\theta}\right)$

$\rightsquigarrow N\left(\underline{O}, \mathcal{I}\left(\underline{\theta}^0\right)\right)$. Furthermore, by ergodic theorem, we have almost surely $\frac{1}{n}\sum_{t=1}^{n} Z_t Z_t' \rightarrow \vartheta$ as $n \rightarrow +\infty$. The result follows from Slutsky's theorem.

■

Note that the conditional least squares estimator $\widehat{\underline{\theta}}_n^{CLS}$ has the following advantages: (i) it is simple and has an explicit form (5.6), (ii) its construction does not need the knowledge of innovation density $f(.)$. However, it is not asymptotically efficient in general due to the heteroscedasticity. In this paper, based on the LeCam [33] approach we establish the locally asymptotic normality (LAN) theorem. This property implies the asymptotic optimality of the CLS estimator and the related statistics (see [47] and [3] for further discussions). For this purpose, we set down the following assumption.

D. $\kappa_4 = E\left\{e_t^4\right\} < \infty$.

E. The innovation density f is symmetric, twice continuously differentiable and satisfies

$$\begin{aligned}(i) &: 0 < \int \left\{\frac{f'(x)}{f(x)}\right\}^2 f(x)\,dx < \infty, (ii) : \int \left\{\frac{f'(x)}{f(x)}\right\}^4 f(x)\,dx < \infty,\\ (iii) &: \lim_{|x|\rightarrow\infty} x^3 f(x) = 0, (iv) : \lim_{|x|\rightarrow\infty} x^2 f'(x) = 0.\end{aligned}$$

Conditionally to $\mathcal{F}_{t-1}$, the density of ϵ_t is $\frac{1}{\sqrt{h_t}} f\left(\frac{\epsilon_t}{\sqrt{h_t}}\right)$ and thus the distribution of $(\epsilon_1, ..., \epsilon_n)$ denoted by $P_{n,\underline{\theta}}$ with density is $dP_{n,\underline{\theta}} := \prod_{t=1}^{n} \frac{1}{\sqrt{h_t(\underline{\theta})}} f\left(\frac{\epsilon_t}{\sqrt{h_t(\underline{\theta})}}\right)$. Thus for two hypothetical values $\underline{\theta}$ and $\underline{\theta}^0 \in \Theta$, the log likelihood ratio is written as $\Lambda_n(\underline{\theta}, \underline{\theta}') := \log \frac{dP_{n,\underline{\theta}}}{dP_{n,\underline{\theta}^0}} = \sum_{t=1}^{n} \log \Phi_t\left(\underline{\theta}, \underline{\theta}^0\right)$ where

$$\Phi_t\left(\underline{\theta}, \underline{\theta}^0\right) = \frac{\sqrt{h_t\left(\underline{\theta}^0\right)} f\left(\frac{\epsilon_t}{\sqrt{h_t(\underline{\theta})}}\right)}{\sqrt{h_t(\underline{\theta})} f\left(\frac{\epsilon_t}{\sqrt{h_t\left(\underline{\theta}^0\right)}}\right)}.$$

Let $\widehat{\underline{\theta}}_n^{CLS} = \underline{\theta} + \frac{\underline{r}}{\sqrt{n}}$ where $\underline{r} := (\underline{r}_1', ..., \underline{r}_s')'$ with $\underline{r}_v = (r_{0v}, ..., r_{qv})' \in \mathbb{R}^{(q+1)}$ a sequence of parameters such that $\widehat{\underline{\theta}}_n^{CLS} \in \overset{\circ}{\Theta}$. We are now in position to state the LAN theorem for the $P-ARCH$ Model (5.2).

Theorem 3 [*Local Asymptotic Normality*] *Suppose that Assumptions* ***A.***, ***B.-D.***, ***E.*** *holds. Then we have under* $P_{n,\underline{\theta}}$

1. *For all* $\underline{\theta} \in \Theta$*, the log-likelihood ratio* $\Lambda_n\left(\underline{\theta}\right) := \Lambda_n\left(\underline{\theta}, \underline{\theta} + \frac{\underline{r}}{\sqrt{n}}\right)$ *admits the following asymptotic representation*

$$\Lambda_n\left(\underline{\theta}\right) = \frac{\underline{r}'}{\sqrt{n}} \frac{\partial}{\partial \underline{\theta}} \log dP_{n,\underline{\theta}} - \frac{1}{2n} \underline{r}' \frac{\partial^2}{\partial \underline{\theta} \partial \underline{\theta}'} \log dP_{n,\underline{\theta}} \underline{r} + o_p(1)$$

where

$$\begin{aligned} \frac{\partial}{\partial \underline{\theta}} \log dP_{n,\underline{\theta}} &= -\frac{1}{2s} \sum_{t=1}^{n} \frac{1}{h_t\left(\underline{\theta}^0\right)} \left\{\left(1 + \frac{f'\left(e_t\right)}{f\left(e_t\right)} e_t\right) \frac{\partial h_t}{\partial \underline{\theta}}\left(\underline{\theta}^0\right)\right\} \\ &= \sum_{t=1}^{n} \underline{w}_t \\ \frac{\partial^2}{\partial \underline{\theta} \partial \underline{\theta}'} \log dP_{n,\underline{\theta}} &= \frac{1}{4s^2} \sum_{t=1}^{n} \frac{1}{h_t^2\left(\underline{\theta}^0\right)} \left\{\left(1 + \frac{f'\left(e_t\right)}{f\left(e_t\right)} e_t\right)^2 \frac{\partial h_t}{\partial \underline{\theta}}\left(\underline{\theta}^0\right) \frac{\partial h_t}{\partial \underline{\theta}'}\left(\underline{\theta}^0\right)\right\} \end{aligned}$$

2. $\Lambda_n\left(\underline{\theta}, \widehat{\underline{\theta}}_n^{CLS}\right) \rightsquigarrow N\left(-\frac{1}{2}\tau^2\left(\underline{\theta}\right), \tau^2\left(\underline{\theta}\right)\right)$.

Proof. The proof rests classically on a Taylor-series expansion of the function $g(\underline{r}) = \Lambda_n\left(\underline{\theta}, \underline{\theta} + \underline{r}/\sqrt{n}\right)$ around $\underline{0}$. We have

$$\Lambda_n\left(\underline{\theta}\right) = \frac{\underline{r}'}{\sqrt{n}} \frac{\partial}{\partial \underline{\theta}} \log dP_{\underline{\theta},n} - \frac{1}{2n} \underline{r}' \frac{\partial^2}{\partial \underline{\theta} \partial \underline{\theta}'} \log dP_{\underline{\theta},n} \underline{r} + o_p(1).$$

where

$$\begin{aligned} \frac{\partial}{\partial \underline{\theta}} \log P_{\underline{\theta}^0,n} &= -\frac{1}{2s} \sum_{t=1}^{n} \frac{1}{h_t\left(\underline{\theta}^0\right)} \left\{\left(1 + \frac{f'\left(e_t\right)}{f\left(e_t\right)} e_t\right) \frac{\partial h_t}{\partial \underline{\theta}}\left(\underline{\theta}^0\right)\right\} = \sum_{t=1}^{n} \underline{w}_t \\ \frac{\partial^2}{\partial \underline{\theta} \partial \underline{\theta}'} \log P_{\underline{\theta}^0,n} &= \frac{1}{4s^2} \sum_{t=1}^{n} \frac{1}{h_t^2\left(\underline{\theta}^0\right)} \left\{\left(1 + \frac{f'\left(e_t\right)}{f\left(e_t\right)} e_t\right)^2 \frac{\partial h_t}{\partial \underline{\theta}}\left(\underline{\theta}^0\right) \frac{\partial h_t}{\partial \underline{\theta}'}\left(\underline{\theta}^0\right)\right\} \end{aligned}$$

It is easy to see that $\left(\underline{w}_t, \mathcal{F}_t\right)_{t \in \mathbb{Z}}$ constitutes a martingale difference sequence. Indeed, firstly we have

$$\begin{aligned} & -\frac{1}{2} \sum_{t=1}^{n} \int_{-\infty}^{+\infty} \left(1 + \frac{f'\left(e_t\right)}{f\left(e_t\right)} e_t\right) \frac{f\left(e_t\right)}{\sqrt{h_t\left(\underline{\theta}^0\right)}} de_t \\ = & -\frac{1}{2\sqrt{h_t\left(\underline{\theta}^0\right)}} \sum_{t=1}^{n} \int_{-\infty}^{+\infty} \left(f\left(e_t\right) + f'\left(e_t\right) e_t\right) de_t = 0, \end{aligned}$$

thence $E_{\underline{\theta}^0}\left\{\underline{w}_t|\mathcal{F}_{t-1}\right\} = \underline{0}$. Applying the central limit theorem for martingale difference (see [12]), Gramèr-Wold device and Slutsky's theorem, it follows that $\Lambda_n\left(\underline{\theta}, \widehat{\underline{\theta}}_n^{CLS}\right) \rightsquigarrow \mathcal{N}\left(-\frac{1}{2}\tau^2\left(\underline{\theta}^0\right), \tau^2\left(\underline{\theta}^0\right)\right)$ where $\tau^2\left(\underline{\theta}^0\right) := \underline{r}'\Gamma\left(\underline{\theta}^0\right)\underline{r}$ with $\Gamma\left(\underline{\theta}\right)$ is a block diagonal matrix given by

$$\Gamma\left(\underline{\theta}\right) = E_{\underline{\theta}^0}\left\{\frac{1}{4s^2}\sum_{v=1}^{s}\frac{1}{h^2_{st+v}\left(\underline{\theta}\right)}\left(\frac{f'\left(e_{st+v}\right)}{f\left(e_{st+v}\right)}e_{st+v}+1\right)^2\frac{\partial h_{st+v}\left(\underline{\theta}\right)}{\partial\underline{\theta}_v}\frac{\partial h_{st+v}\left(\underline{\theta}\right)}{\partial\underline{\theta}'_v}\right\}$$

is block -diagonal. ■

Recalling here, that an estimator $\widehat{\underline{\theta}}_n$ is called asymptotically efficient if its asymptotic variance equal to $\Gamma^{-1}\left(\underline{\theta}^0\right)$. Hence if $\vartheta^{-1}\left(\underline{\theta}^0\right)\mathcal{I}\left(\underline{\theta}^0\right)\vartheta^{-1}\left(\underline{\theta}^0\right) = \Gamma^{-1}\left(\underline{\theta}^0\right)$, then $\widehat{\underline{\theta}}_n^{CLS}$ is said to be asymptotically efficient. Now we state the following theorem.

Theorem 4 *Suppose that Assumptions **C.** and **D.**, **E.** holds. Then the following assertions hold true.*

1. *The asymptotic variance of* $\widehat{\underline{\theta}}_n^{CLS}$ *satisfies the inequality*

$$\mathcal{V}^{-1}\left(\underline{\theta}^0\right)\mathcal{I}\left(\underline{\theta}^0\right)\mathcal{V}^{-1}\left(\underline{\theta}^0\right) \geq \Gamma^{-1}\left(\underline{\theta}^0\right) \tag{5.7}$$

2. $\widehat{\underline{\theta}}_n^{CLS}$ *is asymptotically efficient if and only if*

   a. $h_{st+v} = w_v + \sum_{i=1}^{q}\alpha_i(v)\epsilon^2_{st+v-i} = k$ *(constant) almost surely (a.s) for all* $v = 1, ..., s$.

   b. $e_t \sim \mathcal{N}(0,1)$.

To prove Theorem 4, we use the following matrix inequality.

Lemme 5 *Let A and B be $r \times m$ and $t \times m$ random matrices, respectively, and h is a positive everywhere random variable. If $E\left\{BB'/h\right\}^{-1}$ exists, then*

$$E\left\{AA'h\right\} \geq E\left\{AB'\right\}\left(E\left\{\frac{BB'}{h}\right\}\right)^{-1}E\left\{AB'\right\}'.$$

The equality holds if and only if there exists a constant $r \times t$ matrix C such that almost surely $hA + CB = O$.

Now we proceed to prove theorem 4.

Proof. In lemma 5, let $A_v = \dfrac{\partial h_{st+v}(\underline{\theta})}{\partial \underline{\theta}_v}, B_v = \dfrac{\partial h_{st+v}(\underline{\theta})}{\partial \underline{\theta}_v}$ and $h = h^2_{st+v} = \left(w(v) + \sum\limits_{i=1}^{q} \alpha_i(v)\epsilon^2_{st+v-i}\right)^2$ then we have

$$\widetilde{\mathcal{I}}_v\left(\underline{\theta}^0\right) \geq \mathcal{V}_v\left(\underline{\theta}^0\right)\widetilde{\Gamma}_v^{-1}\left(\underline{\theta}^0\right)\mathcal{V}_v\left(\underline{\theta}^0\right) \tag{5.8}$$

where

$$\begin{aligned}
\widetilde{\mathcal{I}}\left(\underline{\theta}^0\right) &= E_{\underline{\theta}^0}\left\{h^2_{st+v}\left(\underline{\theta}\right)\frac{\partial h_{st+v}(\underline{\theta})}{\partial \underline{\theta}_v}\frac{\partial h_{st+v}(\underline{\theta})}{\partial \underline{\theta}'_v}\right\}, \\
\mathcal{V}\left(\underline{\theta}^0\right) &= E_{\underline{\theta}^0}\left\{\frac{\partial h_{st+v}(\underline{\theta})}{\partial \underline{\theta}_v}\frac{\partial h_{st+v}(\underline{\theta})}{\partial \underline{\theta}'_v}\right\}, \\
\widetilde{\Gamma}\left(\underline{\theta}^0\right) &= E_{\underline{\theta}^0}\left\{\frac{1}{h^2_{st+v}\left(\underline{\theta}\right)}\frac{\partial h_{st+v}(\underline{\theta})}{\partial \underline{\theta}_v}\frac{\partial h_{st+v}(\underline{\theta})}{\partial \underline{\theta}'_v}\right\}
\end{aligned}$$

From lemma 5 we can see that equality in (5.8) holds if and only if almost surely $w(v) + \sum\limits_{i=1}^{q} \alpha_i(v)\epsilon^2_{st+v-i} = k'$ for all $v = 1, ..., s$. Now we are going to prove the inequality (5.7). For this we use the following inequality

$$\begin{aligned}
& E\left\{e^4_{st+v} - 1\right\} E\left\{\left(1 + \frac{f'\left(e_{st+v}\right)}{f\left(e_{st+v}\right)}e_{st+v}\right)^2\right\} \\
= \;& E\left\{\left(e^2_{st+v} - 1\right)^2\right\} E\left\{\left(1 + \frac{f'\left(e_{st+v}\right)}{f\left(e_{st+v}\right)}e_{st+v}\right)^2\right\} \\
\geq \;& \left|E\left\{\left(e^2_{st+v} - 1\right)\left(1 + \frac{f'\left(e_{st+v}\right)}{f\left(e_{st+v}\right)}e_{st+v}\right)\right\}\right|^2 \text{ (by Schwarz inequality)} \\
= \;& \left\{\int_{-\infty}^{+\infty}\left(e^2_{st+v} - 1\right)\left(1 + \frac{f'\left(e_{st+v}\right)}{f\left(e_{st+v}\right)}e_{st+v}\right)f\left(e_{st+v}\right)de_{st+v}\right\}^2 \\
= \;& \left\{\int_{-\infty}^{+\infty}\left(f'\left(e_{st+v}\right)e^3_{st+v} - f'\left(e_{st+v}\right)e_{st+v}\right)de_{st+v}\right\}^2 \\
= \;& \left\{\left(\lim_{a\to\infty}\left[f\left(e_{st+v}\right)\left(e^3_{st+v} - e_{st+v}\right)\right]_{-a}^{a} - \int_{-\infty}^{+\infty} f\left(e_{st+v}\right)\left(3e^2_{st+v} - 1\right)de_{st+v}\right\}^2 \\
= \;& 4.
\end{aligned}$$

The equality holds if and only if there exists constant $c \neq 0$ such that $c\left(e^2_{st+v} - 1\right) = 1 + \dfrac{f'\left(e_{st+v}\right)}{f\left(e_{st+v}\right)}e_{st+v}$, for all $v = 1, ..., s$. Recalling that e_t is independent of $\{\epsilon_{t-1}, \epsilon_{t-2}, ...\}$ so we obtain $\mathcal{I}_v\left(\underline{\theta}^0\right) = E\left\{e^4_t - 1\right\}\widetilde{\mathcal{I}}_v\left(\underline{\theta}^0\right) \geq \mathcal{V}_v\left(\underline{\theta}^0\right)\Gamma_v^{-1}\left(\underline{\theta}^0\right)\mathcal{V}_v\left(\underline{\theta}^0\right)$

which implies

$$\mathcal{I}\left(\underline{\theta}^0\right) = E\left\{e_t^4 - 1\right\}\widetilde{\mathcal{I}}\left(\underline{\theta}^0\right) \geq \mathcal{V}\left(\underline{\theta}^0\right)\Gamma_v^{-1}\left(\underline{\theta}^0\right)\mathcal{V}\left(\underline{\theta}^0\right).$$

The equality holds if and only if there exist constants $k' \neq 0$ and $c \neq 0$ such that almost surely for all $v = 1, ..., s$.

$$w_v + \sum_{i=1}^{q} \alpha_{v,i}\epsilon_{st+v-i}^2 = k' \text{ and } c\left(e_{st+v}^2 - 1\right) = 1 + \frac{f'\left(e_{st+v}\right)}{f\left(e_{st+v}\right)} e_{st+v} \qquad (5.9)$$

From the second equation in (5.9) the solution becomes

$$f\left(e_t\right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{e_t^2}{2}\right) \qquad (5.10)$$

Then the assertions of theorem 4 follow from (5.9) and (5.10). ■

Chapter 6

Conclusion générale: Remaques et quelques perspectives

Dans cette thèse, nous avons essayé de "dégager le voile" sur un domaine très actuel dans les mathématiques empiriques et dans l'économétrie, en envisageant une étude fidèle des modèles PGARCH proposés par Bollerslev et Ghysels (1996) et Franses et Paap (1999) puis popularisés récemment par Bibi et Aknouche à travers leurs travaux sur le sujet (notons notamment Aknouche et Bibi (2009) [2], Bibi et Aknouche (2008) [9]).

Il est utile de bien rappeler quels étaient les buts que nous nous sommes fixés au début de cette étude. L'idée principale est de proposer une approche pour les $PGARCH$ qui peut être utilisée pour estimer, modéliser, espérons-le, de manière plus explicite et adéquate autre que la quasi-maximum de vraisemblance (QMV) proposée par Aknouche et Bibi [2] tout en tenant compte de l'aspect mathématique de notre étude. Nous avons donc pensé dans un premier temps aux moindres carrés (non standards) et plus tard aux moindres carrés conditionnels puis aux équations de Yule-Walker pour les $PGARCH$.

Nous devons néanmoins rester prudents quant à l'interprétation de la fonction $\widehat{Q}_n(\underline{\theta})$ à minimiser dans l'équation 2.7 du chapitre 2, il est claire que cette dernière dépend de l'innovation non observable qui nous la considérons comme un processus de nuisance. Malgré certaines réserves observées par les référés lors de la révision du papier correspondant à ce chapitre ([10]), les résultats obtenues sont encourageants et, semblent indiquer que l'approche que nous proposons peut s'avérer utile (Excellent discussion sur un sujet voisin peut être consulté dans Francq et Zakoîan [23]). Grace aux critiques fructueux des référés sur ce sujet, nous avons pensé plus tard aux m-estimateurs conditionnels, cette approche à fait l'objet d'une Note $CRASS$.

Notre contribution à l'étude de l'identification et de l'estimation dans les modèles $PGARCH$ continue, cependant nous avons envisagé une étude plus détaillée sur le modèle le plus populaire: $PGARCH(1,1)$. Dans cette classe de modèles, nous avons considéré les équations de Yule-Walker, afin d'obtenir une forme explicite des estimateurs. Contrairement au QMV, la stationnarité au second ordre (au sens périodique) jeu un rôle fondamental, et par conséquent, le développement d'une théorie (ou méthodes) de l'estimation dans les modèles $IPGARCH$ s'impose. La généralisation des équations de Yule-Walker dans les $PGARCH$ avec des ordres plus élevés mérite aussi une étude particulière.

Enfin, de nombreux problèmes sont envisageables et certainement désirables. D'un point de vue économétrique, l'extension des modèles $PGARCH$ aux modèles $GARCH$ à coefficients quasi-périodiques est assez directe et mériterait notre attention. De même, une version multivariée périodique permettrait de prendre en compte les inter-relations dynamiques complexes serai parmi nos occupations primordiales.

Bibliography

[1] Adams, G. J. and Goodwin, G. C. (1995). Parameter estimation for periodic ARMA models. Journal of Time Series Analysis, 16, 127-145.

[2] Aknouche, A. and Bibi, A. (2009). Quasi-maximum likelihood estimation of periodic GARCH and periodic ARMA-GARCH processes. J Time Ser Anal, 29(1), 19-45.

[3] Amano, T. and Taniguchi, M. (2008). Asymptotic efficiency of conditional least squares estimators for ARCH models. Stat, Prob. Letters, 78, 179-185.

[4] Basawa, I.V., and Lund R.B. (2001). Large sample properties of parameter estimates for periodic ARMA models. J. Time Ser. Anal. Vol. 22, No. 6, 651-663.

[5] Berkes, I., Horvath, L. and Kokoskza, P. (2003). GARCH processes: Structure and estimation. Bernoulli, 9, 201-227.

[6] Bessembinder, H. et M., Hertzel (1993). Return autocorrelations around nontrading days. Review of financial studies. 6(1) 155-89.

[7] Bezandry, P.H., and T. Diagana (2011). Almost periodic stochastic processes. Springer.

[8] Bibi, A. and Aknouche, A. (2008). On periodic GARCH processes: Stationarity, existence of moments and geometric ergodicity. Math. Methods of Statist, 17(4), 305-316.

[9] Bibi, A. and Aknouche, A. (2009). Propriétés probabilistes des processus GARCH périodiques. C. R. Acad. Sci. Paris, Ser. I 347, 299–303.

[10] Bibi, A. and Lescheb, I. (2010). Strong consistency and asymptotic normality of least squares estimators for PGARCH and PARMA-PGARCH. Stat. Prob. Letters 80, 1532-1542.

[11] Bibi, A., and Lescheb, I. (2010). A conditional least squares approach to PGARCH and PARMA-PGARCH time series estimation. C.R. Acad. Sci. Paris, Ser. I, Vol. 21-22, 1211-1216.

[12] Billingsley, P. (1995). Probability and measure. (3rd Edition) Wiley-Interscience.

[13] Bittanti, S. and De Nicolao, G. (1993). Spectral factorization of linear periodic systems with application to the optimal prediction of periodic ARMA models. Automatica, 29, 517-522.

[14] Bloomfield, P., Hurd, H. L. and Lund, R. B. (1994). Periodic correlation in stratospheric ozone data. J. Time Ser. Anal. 15:127,150.

[15] Bollerslev, T., and Ghysels, E. (1996). Periodic autoregressive conditional heteroskedasticity. J. of Business & Economic Statistics, 14, 139-151.

[16] Bougerol, P., and N., Picard (1992). Stationarity of GARCH processes and some nonnegative time series. Journal of Econometrics, 52, 115-127.

[17] Brandt, A. (1986). The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. Adv. in Appl. Probab. 18, 211-220.

[18] Cleveland, W. P and Tiao, G. C. (1979). Modeling seasonal time series. Revue Economique Appliquée, 32, 107-129.

[19] Cramér, H. (1961). On some classes of nonstationary stochastic processes. Proceedings of the fourth Berkeley sympposium on mathematical statistics and probability, Vol.II, 57-78.

[20] Francq, C. and Zakoïan, J. M. (1998). Estimation linear representations of nonlinear processes. J. Statist. Planning and inference, Vol. 68(1), 145-165.

[21] Francq, C. and Zakoïan, J. M. (2009). Modèles GARCH: Structure, Inférence statistique et Applications finançières. Ed. Economica.

[22] Francq, C. and Zakoïan, J. M. (2010). Strict stationarity testing and estimation of explosive ARCH models. MPRA.

[23] Francq, C. and Zakoïan, J. M. (2010). Optimal predictions of powers of conditionally heteroskedastic processes. MPRA.

[24] Franses, P. H. (1996). Periodicity and Stochastic Trends in Economic Times Series. Oxford University Press.

[25] Franses, P. H. and Paap, R. (2004). Periodic time series models. Oxford University Press.

[26] Gardner, W. A. and Franks, L. E. (1975). Characterization of cyclostationary, random signal prcesses. IEEE Trans. Inform. Theory, 21, 4-14.

[27] Hallin, M. (1989). Modèles non stationnaires: séries univariées et multivariées. Dans Séries chronologiques: Théorie et pratique des modèles ARIMA. ed., J.J. Droesbeke et all. Chez Economica.

[28] Hurd, H. L. and Miamee, A. G. (2007). Periodically correlated random sequences: Spectral Theory and Practice. Wiley&Sons.

[29] Jiming, J. (2010). Large sample techniques for statistics. Springer.

[30] Kesten, H. and Spitzer, F. (1984). Convergence in distribution for products of random matrices. Z. Wahrsch. Verw. Gebiete, 67, 363-386.

[31] Kim, W., and O., Linton (2011). Estimation of a semiparametric IGARCH(1.1) model. Econometric Theory 27, 1-23.

[32] Kristensen, D., and O., Linton (2006). A closed-form estimator for the GARCH(1.1) model. Econometric Theory 22, 223-337.

[33] LeCam, L., Asymptotic Methods in statistical decision theory. Springer-Verlag. 1986.

[34] Lee, O., and Shin, D.W. (2010). Geometric Ergodicity and moment conditions for seasonal GARCH model with periodic coefficients. Commun. Stat.: Theory and Methods. 39, 38-51.

[35] Lund, R. B., and Basawa I.V. (2000). Recursive prediction and likelihood evaluation for periodic ARMA models. Journal of time series analysis, 21, 75-93.

[36] Maercker, G., and M., Moser (2000). Yule-Walker type estimators in GARCH(1,1) models: Asymtotic normality and Bootstrap. Preprint.

[37] Meyer, R. A. and Burrus, C. S. (1975). A unified analysis of multirate and periodically time varying digital filters. IEEE Trans. Circuits and Systems, 22, 162-168.

[38] Meitz, M., and P., Saikonen (2008). Ergodicity, Mixing, and existence of moments of a class of markov models with applications to GARCH and ACD models. Econometric Theory 24, 1291-1320.

[39] Mikosch,T., and C. Stărică (2000). Limit theory for the sample autocorrelations and extremes of a GARCH(1.1) process. Annals of Statistics 28, 1427-1451.

[40] Nicholls, D. F. and Quinn, B. G. (1982). Random coefficient autoregressive model: An introduction. Springer Verlag, New York.

[41] Pagano, M. (1978). On periodic and multiple autoregression. The Annals of Statistics, 6, 1310-1317.

[42] Parzen, E. and Pagano M. (1979). An approach to modeling seasonally stationary time series. J. Econometrics, 9, 137-153.

[43] Rubin, H. (1950). Note on random coefficients. In T.C. Koopman, ed., Statistical inference in dynamic economic models. J. Wiley, New-York.

[44] Sébastien, L., et J-P., Urbain (2002). L'apport des modèles périodiques à longue mémoire pour la modélisation de l'effet jour sur la volatilité des séries financières, pp. 221-246. Dans Finances publiques, Finances privées, ed. Jurion, B. et Pestieu, P. Chez les Editions.

[45] Shimizu, K. (2010). Bootstrapping stationary ARMA-GARCH models. Vieweg+Teubner Research.

[46] Straumann, D., and T. Mikosch (2006). Quasi-maximum likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equations approach. The Annals of Statistics. Vol. 34, No. 5, pp. 2449-2495.

[47] Taniguchi, M., and Yoshihide, K. (2000). Asymptotic theorey of statistical inference of time series. Springer.

[48] Vecchia, A. V. (1985b). Periodic autoregressive-moving average (PARMA) modeling with application to water resources. Water Resources Bulletin, 21, 721-730.