

Abdelhakim Aknouche<sup>1,2</sup> / Nacer Demmouche<sup>3</sup> / Stefanos Dimitrakopoulos<sup>4</sup> / Nassim Touche<sup>5</sup>

# Bayesian analysis of periodic asymmetric power GARCH models

<sup>1</sup> Faculty of Mathematics, University of Science and Technology Houari, Boumediene, Algeria, E-mail: aknouche\_ab@yahoo.com

<sup>2</sup> College of Science, Qassim University, Buraidah, Saudi Arabia, E-mail: aknouche\_ab@yahoo.com

<sup>3</sup> Department of Mathematics, University Akli Mohand Oulhadj, Bouira, Algeria

<sup>4</sup> Economics Division, Leeds University, Leeds, UK

<sup>5</sup> Faculty of Exact Sciences, University of Bejaia, Bejaia, Algeria

## Abstract:

In this paper, we set up a generalized periodic asymmetric power *GARCH* (*PAP-GARCH*) model whose coefficients, power, and innovation distribution are periodic over time. We first study its properties, such as periodic ergodicity, finiteness of moments and tail behavior of the marginal distributions. Then, we develop an *MCMC* algorithm, based on the Griddy-Gibbs sampler, under various distributions of the innovation term (Gaussian, Student-t, mixed Gaussian-Student-t). To assess our estimation method we conduct volatility and Value-at-Risk forecasting. Our model is compared against other competing models via the Deviance Information Criterion (*DIC*). The proposed methodology is applied to simulated and real data.

**Keywords:** Bayesian forecasting, Deviance Information Criterion, Griddy-Gibbs, periodic asymmetric power *GARCH* model, probability properties, Value at Risk

**MSC 2010:** AMS 2000 Primary 62M10, Secondary 60F99

**DOI:** 10.1515/snde-2018-0112

## 1 Introduction

Nowadays, the generalized autoregressive conditional heteroskedastic (*GARCH*) model, introduced by Engle (1982) and Bollerslev (1986), is a well-established model in the financial econometrics literature. Its popularity comes from the fact that it can capture some stylized facts of asset returns series, such as volatility clustering and excess kurtosis in a very parsimonious way. Since its advent, numerous extensions of the standard *GARCH* model have been proposed to account for additional empirical characteristics of financial time series (Francq and Zakoian, 2019).

In this paper, our contribution to the *GARCH* literature is three-fold. First of all, building upon the work of Aknouche, Al-Eid, and Demouche (2018), we develop an estimation algorithm for a generalized periodic asymmetric power *GARCH* (*PAP-GARCH*( $p, q$ )) model, whose coefficients, power, and innovation distribution are periodic over time. The algorithm accounts for three different cases of innovations. The first case refers to the standard Gaussian distribution, making the innovation independent and identically distributed (*iid*). In the second case, the innovation is rather independent and *S*-periodically distributed (*ipd<sub>S</sub>*), having a standardized Student-t distribution with an *S*-periodic degree of freedom. The third case is a combination of the first two cases.

Aknouche, Al-Eid, and Demouche (2018) set up the *PAP-GARCH*(1, 1) model and established the strong consistency and asymptotic normality of the generalized quasi-maximum likelihood estimator (*GQMLE*) under general weak and tractable assumptions. However, since the *PAP-GARCH* model contains a large number of parameters (as is the case for any periodic model), any estimation method utilizing optimization techniques would be cumbersome and subject to errors.

To overcome these estimation difficulties we resort to Bayesian methods that do not rely on optimization routines. By augmenting the parameter space to include all the parameters, standard Bayesian techniques can be applied. Therefore, we design a Markov Chain Monte Carlo (*MCMC*) algorithmic scheme for the general order *PAP-GARCH*( $p, q$ ) model, using the Griddy-Gibbs sampler (Ritter and Tanner, 1992). This posterior sampler is an extension of the Gibbs sampler to the case where the posterior density of some parameters has a complex or non-standard form. The Griddy-Gibbs has been employed to *GARCH*-related model by Bauwens and Lubrano

Abdelhakim Aknouche is the corresponding author.

© 2019 Walter de Gruyter GmbH, Berlin/Boston.

(1998) and Xia et al. (2017), among others. Also, for the class of models we study, the Bayesian approach has additional advantages over the classical approach, which are explicitly explained in Ardia (2008).

To evaluate the performance of the proposed algorithm we conduct a simulation study, where we monitor convergence and mixing based on nonparametric diagnostic tools, such as the Relative Numerical Inefficiency (RNI), the Numerical Standard Error (NSE) as well as other MCMC correlation measurements. To assess the predictive performance of the proposed model we conduct in-sample and out-of-sample volatility forecasting and Value at Risk (*VaR*) forecasting; see, for example, Akrouche (2017) and Xia et al. (2017).

Second, we carry out a model comparison exercise, using the Deviance Information Criterion (*DIC*, Spiegelhalter et al. 2002). In particular, our model is compared against some of its nested versions; the periodic *GARCH* (*P-GARCH*, Bollerslev and Ghysels, 1996), the periodic threshold *GARCH* (*PT-GARCH*) and the non-periodic asymmetric power *GARCH* (*AP-GARCH*, Ding, Granger, and Engle 1993). The *DIC* is also used to select the order and the period of the *PAP-GARCH*( $p, q$ ) model. We illustrate the methodology of the paper with an empirical application to *S&P500* returns.

Third, for the selection of the priors of some of the parameters of our model, we need to examine their stationary regions. Therefore, prior to designing the MCMC algorithm, we study some stability/probabilistic properties of the *PAP-GARCH*( $p, q$ ) model, such as periodic ergodicity, existence of moments and tail behavior of the marginal distributions.

The motivation of this paper comes from the flexibility and usefulness of the *PAP-GARCH* model in analyzing financial data. To be more specific, in an attempt to control for additional characteristics of financial data, including leverage and Taylor effects, Ding, Granger, and Engle (1993) put forward the asymmetric power *GARCH* model (*AP-GARCH*); see also Hwang and Basawa (2004), Pan, Wang, and Tong (2008), Hamadeh and Zakoian (2011), Francq and Zakoian (2013), Akrouche and Touche (2015), and Xia et al. (2017).

Compared to other competitive volatility models such as Markov Switching *GARCH* models (Haas, Mittnik and Paolella, 2004; Francq and Zakoian, 2008; Bauwens, Dufays and Rombouts, 2014), the *AP-GARCH* specification has a simpler probability structure and it is easier to estimate by maximum likelihood-type methods, a fact that makes it quite popular. However, a serious deficiency of the *AP-GARCH* model is that it assumes time-invariant parameters.

Hence, this model fails to represent time series for which the volatility distribution varies over time according to a seasonal or periodic pattern. Notable examples are the day-of-the-week effect and the month-of-the-year effect, both observed in return series (Bollerslev, Cai and Song, 2000; Franses and Paap, 2000; Tsiakas, 2006; Osborn, Savva and Gill, 2008; Smith, 2010; Rossi and Fantazani, 2015; Akrouche, 2017). Other non-financial applications of periodic volatility models include hourly electricity demand, intraday wind power and wind speed series (Ambach and Croonenbroeck, 2015; Ambach and Schmid, 2015; Ziel, Croonenbroeck and Ambach, 2016).

Bollerslev and Ghysels (1996) introduced a periodic *GARCH* (*P-GARCH*) model, where the volatility coefficients are periodic over time with periodicity  $S$ . The *P-GARCH* model is sufficiently flexible and rich enough to control for periodicity in volatility and other useful characteristics of return series (Franses and Paap, 2000; Osborn, Savva and Gill, 2008; Rossi and Fantazani, 2015). However, it can not capture leverage and Taylor effects as well as heavy tailedness of the marginal distribution.

Akrouche, Al-Eid, and Demouche (2018) generalized the *AP-GARCH*(1, 1) taking also into account periodicity, thus solving the problems encountered both in the *AP-GARCH* and the *P-GARCH* models. In addition, the existing literature on periodic *GARCH* models generally assumes stationarity of the innovation term, so the periodicity of the model is driven solely by the volatility coefficients (Bollerslev and Ghysels, 1996; Franses and Paap, 2000; Osborn, Savva and Gill, 2008; Akrouche and Bibi, 2009; Akrouche and Al-Eid, 2012; Rossi and Fantazani, 2015; Ziel, Steinert and Husmann, 2015; Ziel, Croonenbroeck and Ambach, 2016). In many applications, this might be a restrictive assumption, when there are seasonal return series that are characterized by time-varying shape marginal distributions. This restriction is relaxed in the *PAP-GARCH*(1, 1) model of Akrouche, Al-Eid, and Demouche (2018), as the periodicity is manifested through both the volatility parameters and the distribution of the innovation sequence. This makes the model not only more flexible in representing periodic volatility, at just a minor cost of a few additional parameters, but also important in analyzing financial data.

The paper is organized as follows. In Section 2 we derive the probabilistic properties of the *PAP-GARCH*( $p, q$ ) model, while in Section 3 we develop the Bayesian Griddy-Gibbs algorithm. Section 4 assesses the performance of the algorithm in finite samples through simulation experiments. In Section 5 we conduct our empirical analysis. Section 6 concludes. Two appendices accompany this paper.

## 2 Structure of the *PAP-GARCH* $_S(p, q)$ model

Let  $\{\epsilon_t, t \in \mathbb{Z}\}$  be a *PAP-GARCH* $_S(p, q)$  process with period  $S$  and orders  $p$  and  $q$ , given by the following equation

$$\begin{cases} \epsilon_t = \sigma_t \eta_t \\ \sigma_t^{\delta_t} = \omega_t + \sum_{i=1}^q \alpha_{ti+} (\epsilon_{t-i}^+)^{\delta_{t-i}} + \alpha_{ti-} (\epsilon_{t-i}^-)^{\delta_{t-i}} + \sum_{j=1}^p \beta_{tj} \sigma_{t-j}^{\delta_{t-j}} \end{cases}, \quad t \in \mathbb{Z}, \tag{1}$$

where  $x^+ = \max(x, 0)$ ,  $x^- = -\min(x, 0)$  and  $\{\eta_t, t \in \mathbb{Z}\}$ , called the model innovation, is a sequence of independent and  $S$ -periodically distributed ( $ipd_S$ ) unobservable random variables ( $S \geq 1$ ) such that  $\eta_t$  is independent of  $\{\epsilon_i, i < t\}$ . The volatility parameters  $\omega_t > 0$ ,  $\alpha_{ti+} \geq 0$ ,  $\alpha_{ti-} \geq 0$ ,  $\beta_{tj} \geq 0$ , and  $\delta_i > 0$  ( $1 \leq i \leq q$ ,  $1 \leq j \leq p$ ) are  $S$ -periodic over  $t$ . To emphasize the periodicity of model (1) we rewrite it in the following form

$$\begin{cases} \epsilon_{Sn+v} = \sigma_{Sn+v} \eta_{Sn+v} \\ \sigma_{Sn+v}^{\delta_v} = \omega_v + \sum_{i=1}^q \alpha_{vi+} (\epsilon_{Sn+v-i}^+)^{\delta_{v-i}} + \alpha_{vi-} (\epsilon_{Sn+v-i}^-)^{\delta_{v-i}} + \sum_{j=1}^p \beta_{vj} \sigma_{Sn+v-j}^{\delta_{v-j}} \end{cases} \quad \begin{matrix} 1 \leq v \leq S, \\ n \in \mathbb{Z}, \end{matrix} \tag{2}$$

where for all  $1 \leq v \leq S$ , the  $v$ th season (or channel) stands for the set  $\{\dots, v - S, v, v + S, \dots\}$ .

Model (1), proposed by Aknouche, Al-Eid, and Demouche (2018) for the case  $p = q = 1$ , is quite general and covers a wide range of well-known GARCH-type models. For  $S = 1$ , it is just the asymmetric power GARCH (AP-GARCH( $p, q$ )) model proposed by Ding, Granger, and Engle (1993) and revisited by Pan, Wang, and Tong (2008); see also Francq and Zakoian (2013). It reduces to the periodic GARCH( $p, q$ ) when  $\delta_v = 2$  and  $\alpha_{v+} = \alpha_{v-}$  ( $1 \leq v \leq S$ ), to the periodic power GARCH( $p, q$ ) when  $\alpha_{v+} = \alpha_{v-}$  ( $1 \leq v \leq S$ ) and to the periodic threshold GARCH( $p, q$ ) when  $\delta_v = 1$  for all  $1 \leq v \leq S$ . Besides the stylized facts captured by the AP-GARCH( $p, q$ ) model such as the leverage effect and the Taylor property (Granger, 2005; Haas, 2009; Aknouche and Touche, 2015), model (1) might also account for periodicity in volatility.

To study the probabilistic structure of the PAP-GARCH $_S(p, q)$  model, we express model (1) as a stochastic recurrence equation with  $ipd_S$  coefficients. Let  $r = p + 2q - 2$ ,

$$\begin{aligned} \zeta_t &= \beta_{t+1,1} + \alpha_{t+1,1+} (\eta_t^+)^{\delta_t} + \alpha_{t+1,1-} (\eta_t^-)^{\delta_t}, \\ Y_t &= \left( \sigma_{t+1}^{\delta_{t+1}}, \dots, \sigma_{t-p+2}^{\delta_{t-p+2}}, (\epsilon_t^+)^{\delta_t}, (\epsilon_t^-)^{\delta_t}, \dots, (\epsilon_{t-q+2}^+)^{\delta_{t-q+2}}, (\epsilon_{t-q+2}^-)^{\delta_{t-q+2}} \right)' \in \mathbb{R}^r, \\ B_t &= (\omega_{t+1}, 0, \dots, 0) \in \mathbb{R}^r, \end{aligned}$$

and

$$A_{t-1} = \begin{pmatrix} \zeta_{t-1} & \beta_{t,2} & \dots & \beta_{t,p-1} & \beta_{t,p} & \alpha_{t,2+} & \alpha_{t,2-} & \dots & \alpha_{t,q-1+} & \alpha_{t,q-1-} & \alpha_{t,q+} & \alpha_{t,q-} \\ 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ (\eta_{t-1}^+)^{\delta_{t-1}} & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ (\eta_{t-1}^-)^{\delta_{t-1}} & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 \end{pmatrix}$$

Equation (1) may be cast in the following stochastic recurrence equation

$$Y_t = A_t Y_{t-1} + B_t, \quad t \in \mathbb{Z}, \tag{3}$$

where  $\{(A_t, B_t), t \in \mathbb{Z}\}$  is an  $ipd_S$  sequence valued in  $\mathcal{M}_r(\mathbb{R}) \times \mathbb{R}^r$ , with  $\mathcal{M}_r(\mathbb{R})$  being the set of square matrices of dimension  $r$ . Let  $\gamma^{(S)}$  be the top Lyapunov exponent associated with the recurrence equation (3) which is given by (Aknouche and Bibi, 2009)

$$\gamma^{(S)} = \inf \left\{ \frac{1}{n} E \log \|A_{nS} \dots A_2 A_1\|, \quad n \geq 1 \right\}. \tag{4}$$

The following result gives necessary and/or sufficient conditions for equation (1) to have a unique strictly periodically stationary and periodically ergodic solution (see also Aknouche, Al-Eid, and Demouche 2018 for the definition of periodic ergodicity).

**Theorem 1**

- a. Assume that  $E \log |\eta_v|^{\delta_v} < \infty$  for all  $1 \leq v \leq S$ . A necessary and sufficient condition for model (3) to have a unique nonanticipative strictly periodically stationary and periodically ergodic solution is that

$$\gamma^{(S)} < 0. \tag{5}$$

This solution is given by

$$X_t = \sum_{j=0}^{\infty} \prod_{i=0}^{j-1} A_{t-i} B_{t-j}, \quad t \in \mathbb{Z}, \tag{6}$$

where the series in the right hand side of (6) converges absolutely almost surely.

- b. If (3) admits a strictly periodically stationary solution then

$$\rho \left( \prod_{v=0}^{S-1} \beta_{S-v} \right) < 1, \tag{7}$$

where  $\beta_t$  is the submatrix of  $A_t$  defined by

$$\beta_t = \begin{pmatrix} \beta_{t,1} \beta_{t,2} \dots \beta_{t,p-1} & \beta_{t,p} \\ I_{(p-1) \times (p-1)} & 0_{(p-1) \times 1} \end{pmatrix},$$

and  $\rho(A)$  denotes the spectral radius of the squared matrix  $A$ , i.e. the maximum modulus of the eigenvalues of  $A$ .

In the special case  $p = q = 1$ , condition (5) reduces to

$$\frac{1}{S} \sum_{v=1}^S E \left( \log (\alpha_{0v+} (\eta_{v-1}^+)^{\delta_{v-1}} + \alpha_{0v-} (\eta_{v-1}^-)^{\delta_{v-1}} + \beta_{0v}) \right) < 0,$$

(Aknouche, Al-Eid, and Demouche 2018) while (7) is just  $\prod_{v=0}^{S-1} \beta_{0v} < 1$ . Now conditions for the existence of moments of the PAP-GARCH<sub>S</sub>( $p, q$ ) process are given as follows.

**Theorem 2**

- a. If  $\gamma^S(A) < 0$  then there is  $\kappa > 0$  such that for all  $t$

$$E(\sigma_t^\kappa) < \infty \quad \text{and} \quad E(|\epsilon_t|^\kappa) < \infty. \tag{8}$$

- b. Let  $\{\epsilon_t, t \in \mathbb{Z}\}$  be a strictly periodically stationary solution of (1). A necessary and sufficient condition for  $E(\epsilon_t^{m\delta_t})$  ( $m \in \mathbb{N}^*, 1 \leq t \leq S$ ) to be finite is that

$$\rho \left( \prod_{v=0}^{S-1} E(A_{S-v}^{\otimes m}) \right) < 1, \tag{9}$$

where  $A^{\otimes m}$  is the Kronecker product:  $A \otimes A \otimes \dots \otimes A$  with  $m$  factors.

The following result shows that the  $S$  marginal distributions of the PAP-GARCH<sub>S</sub>( $p, q$ ) model are regularly varying provided that some limiting moment conditions are satisfied. Thus, heavy-tailed marginals can be obtained for the PAP-GARCH<sub>S</sub>( $p, q$ ) model even when its input innovation sequence  $\{\eta_t, t \in \mathbb{Z}\}$  has light-tailed distributions.

Automatically generated rough PDF by ProofCheck from River Valley Technologies Ltd

**Theorem 3**

Assume model (1) satisfies the following three conditions: (i)  $\gamma^{(S)} < 0$  (ii)  $\eta_t$  has  $S$  positive densities on  $\mathbb{R}$ , such that  $E|\eta_v|^{\tau_v} < \infty$  for some  $\tau_v > 0$  ( $1 \leq v \leq S$ ), and (iii)  $\min_{1 \leq v \leq S} (\omega_v) > 0$ . Then for all  $1 \leq v \leq S$ ,

$$P(\epsilon_v > x) \sim c_v x^{-\alpha \delta_v},$$

where  $c_v > 0$  and  $\alpha$  is the unique solution of

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log E \|A_{nS+v} A_{nS+v-1} \dots A_{v+2} A_{v+1}\|^{\frac{\alpha}{2}} = 0.$$

The proofs of the theorems are given in Appendix A.

**3 Prior-Posterior analysis**

**3.1 Likelihood and priors**

Let  $\epsilon^T = (\epsilon_1, \dots, \epsilon_T)'$  be a series of observations generated from the  $PAP-GARCH_S(p, q)$  model (1) with sample-size  $T = NS$  ( $N \geq 1$ ). We first need to specify the distribution of the  $ipd_S$  innovation  $\{\eta_t, t \in \mathbb{Z}\}$ .

Consider the following three cases:

Case (i) *The pure Gaussian case:*  $\eta_1, \dots, \eta_S$  are normally distributed with mean zero and unit variance ( $\eta_v \sim N(0, 1)$ ), i.e.

$$f(\eta_v) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\eta_v^2} \quad (1 \leq v \leq S).$$

Case (ii) *The pure Student-t case:*  $\eta_1, \dots, \eta_S$  are (standardized) Student-t distributed with  $\tau_1, \dots, \tau_S > 2$  degrees of freedom, respectively ( $\frac{\tau_v}{\tau_v-2}\eta_v \sim t(\tau_v)$ ), i.e.

$$f(\eta_v) = \frac{1}{\sqrt{\pi(\tau_v-2)}} \frac{\Gamma(\frac{\tau_v+1}{2})}{\Gamma(\frac{\tau_v}{2})} \left(1 + \frac{\eta_v^2}{(\tau_v-2)}\right)^{-\frac{\tau_v+1}{2}} \quad (1 \leq v \leq S).$$

Case (iii) *The mixed Gaussian-Student-t case:* It is assumed that for certain seasons  $\{v_1, \dots, v_k\} \subset \{1, \dots, S\}$  ( $1 \leq k \leq S-1$ ) the innovations  $\eta_{v_1}, \dots, \eta_{v_k}$  are normally distributed with mean zero and unit variance. For the remaining seasons  $\{v_{k+1}, \dots, v_S\}$ ,  $\eta_{v_j}$  has a standardized Student-t distribution with  $\tau_{v_j} > 2$  ( $j = k+1, \dots, S$ ) degrees of freedom. Cases (i) and (ii) are particular instances of Case (iii) and they correspond to  $k = S$  and  $k = 0$ , respectively.

Observe that the innovation  $\{\eta_t, t \in \mathbb{Z}\}$  is *iid* in the first case and *ipd<sub>S</sub>* in the second and third cases. The choices of  $k$  and  $\{v_1, \dots, v_k\}$  are motivated by practical considerations, such as the empirical Kurtosis of each season and the shapes of the seasonal empirical distributions. Effective model selection criteria, such as the *DIC*, allow us to select the best model. The MCMC method of Section 3.2 refers to the general Case (iii).

Assuming that the power  $\delta_v$  ( $1 \leq v \leq S$ ) and  $\tau_{v_j}$  ( $k+1 \leq j \leq S$ ) are both unknown, the parameter vector to be estimated is denoted by  $\theta = (\tau', \theta'_1, \dots, \theta'_S)' \in (0, \infty)^{r^*}$ , where  $r^* = (p+2q+2)S + (S-k)$ ,  $\tau' = (\tau_{v_{k+1}}, \dots, \tau_{v_S})'$ ,  $\theta'_v = (\omega_v, \alpha'_{v+}, \alpha'_{v-}, \beta'_v, \delta_v)' \in (0, \infty)^{(p+2q+2)}$ ,  $\alpha_{v+} = (\alpha_{v_1+}, \dots, \alpha_{v_q+})'$ ,  $\alpha_{v-} = (\alpha_{v_1-}, \dots, \alpha_{v_q-})'$  and  $\beta_v = (\beta_{v_1}, \dots, \beta_{v_p})'$  ( $1 \leq v \leq S$ ).

Adopting the Bayesian approach, the parameter vector  $\theta$  is assigned a prior distribution  $f(\theta)$ . The goal is, thus, to make inference about the posterior distribution  $f(\theta/\epsilon^T)$ , which satisfies the following proportionality

$$f(\theta/\epsilon^T) \propto f(\theta)f(\epsilon^T/\theta), \tag{10}$$

where  $f(\epsilon^T/\theta)$  is the likelihood function given by

$$f(\epsilon^T/\theta) = \prod_{t=1}^T f(\epsilon_t/\epsilon^{t-1}, \theta) = \prod_{n=0}^{N-1} \prod_{v=1}^S l_{v+nS}(\theta) \tag{11}$$

$$l_{v+nS}(\theta) = \begin{cases} \frac{1}{\sigma_{v+nS}(\theta)} e^{-\frac{\epsilon_{v+nS}^2}{\sigma_{v+nS}^2(\theta)}} & \text{if } v \in \{v_1, \dots, v_k\}, \quad 0 \leq k \leq S \\ \frac{1}{\sigma_{v+nS}(\theta) \sqrt{(\tau_v - 2)}} \frac{\Gamma\left(\frac{\tau_v + 1}{2}\right)}{\Gamma\left(\frac{\tau_v}{2}\right)} \left(1 + \frac{\epsilon_{v+nS}^2}{(\tau_v - 2) \sigma_{v+nS}^2(\theta)}\right)^{-\frac{\tau_v + 1}{2}} & \text{if } v \in \{v_{k+1}, \dots, v_S\} \end{cases} \quad (12)$$

$$\sigma_{v+nS}(\theta) = \left( \omega_v + \sum_{i=1}^q \alpha_{vi+} (\epsilon_{S_{n+v-i}}^+)^{\delta_{v-i}} + \alpha_{vi-} (\epsilon_{S_{n+v-i}}^-)^{\delta_{v-i}} + \sum_{j=1}^p \beta_{vj} \sigma_{S_{n+v-j}}^{\delta_{v-j}}(\theta) \right)^{\frac{1}{\delta_v}}. \quad (13)$$

Because of the periodic structure of the  $PAP-GARCH_S(p, q)$  model, it is natural to assume that the parameters  $\theta_1, \dots, \theta_S$  are independent of each other. The joint posterior distribution  $f(\theta/\epsilon^T)$  can, then, be estimated using the Gibbs sampler, provided we can draw samples from any of the  $r^*$  conditional posterior distributions:  $f(\tau_{v_j}/\theta_{-\{\tau_{v_j}\}}, \epsilon^T)$  ( $k + 1 \leq j \leq S$ ),  $f(\delta_v/\theta_{-\{\delta_v\}}, \epsilon^T)$ ,  $f(\omega_v/\theta_{-\{\omega_v\}}, \epsilon^T)$ ,  $f(\alpha_{vi+}/\theta_{-\{\alpha_{vi+}\}}, \epsilon^T)$ ,  $f(\alpha_{vi-}/\theta_{-\{\alpha_{vi-}\}}, \epsilon^T)$  ( $i = 1, \dots, q$ ) and  $f(\beta_{vj}/\theta_{-\{\beta_{vj}\}}, \epsilon^T)$  ( $j = 1, \dots, p$ ) ( $1 \leq v \leq S$ ), where  $x_{-\{\theta_j\}}$  denotes the vector obtained from  $x$  after removing the parameter component  $\theta_j$ .

To get the latter conditional posterior distributions, prior distributions of the parameters have to be determined. In general, the choice of conjugate priors is appealing, as it simplifies the posterior analysis. However, since the volatility  $\sigma_t(\theta)$  is a nonlinear function of  $\theta$  with a feedback mechanism, it is difficult to find conjugate priors. That's why noninformative priors are used. They are given, for all  $1 \leq v \leq S$  and  $k + 1 \leq j \leq S$ , by

$$\begin{aligned} \omega_v &\sim U_{(0, A)}, \quad \alpha_{v+} \sim U_{(0, A)^q}, \quad \alpha_{v-} \sim U_{(0, A)^q}, \\ \beta_v &\sim U_{(0, A)^p}, \quad \delta_v \sim U_{(a, A)}, \quad \tau_{v_j} \sim U_{(2, A)}, \end{aligned}$$

where  $U_D$  denotes the uniform distribution on the set  $D$ ,  $A$  is a fairly large positive number and  $a < A$  is a quite small positive number.

To simplify the computations, we reparametrize the standardized Student-t distribution with  $\psi_{v_j} = \frac{1}{\tau_{v_j}}$ . The prior distribution for  $\psi_v$  is then

$$\psi_{v_j} \sim U_{\left(\frac{1}{A}, 0.5\right)}, \quad (1 \leq v \leq S).$$

Based on the above priors, the conditional posterior distributions of  $\psi_{v_j}$ ,  $\delta_v$ ,  $\omega_v$ ,  $\alpha_{v+}$ ,  $\alpha_{v-}$  and  $\beta_v$  can easily be derived from (10), except for a scale factor. For example, the kernel of the conditional posterior of  $\delta_v$  is written as follows

$$f(\delta_v/\theta_{-\{\delta_v\}}, \epsilon^T) \propto \prod_{n=0}^{N-1} \prod_{v=1}^S l_{v+nS}(\theta), \quad 1 \leq v \leq S, \quad (14)$$

where  $l_{v+nS}(\theta)$  is given by (12)–(13). The kernels of the remaining distributions  $f(\psi_{v_j}/\theta_{-\{\tau_{v_j}\}}, \epsilon^T)$ , ( $k + 1 \leq j \leq S$ ),  $f(\omega_v/\epsilon^T, \theta_{-\{\omega_v\}})$ ,  $f(\alpha_{vi+}/\epsilon^T, \theta_{-\{\alpha_{vi+}\}})$ ,  $f(\alpha_{vi-}/\epsilon^T, \theta_{-\{\alpha_{vi-}\}})$  ( $i = 1, \dots, q$ ) and  $f(\beta_{vj}/\epsilon^T, \theta_{-\{\beta_{vj}\}})$  ( $1 \leq v \leq S$ ) are obtained as in (14). However, the parameters  $\omega_v$ ,  $\alpha_{v+}$ ,  $\alpha_{v-}$  and  $\beta_v$  are restricted to lie in the periodic stationarity domain described by (5).

### 3.2 MCMC algorithm

Once determining the kernel of  $f(\theta_j/\epsilon^T, \theta_{-\{\theta_j\}})$  ( $j = 1, \dots, r^*$ ), we can use some indirect sampling algorithms in order to draw each component of  $\theta$ . We choose the Griddy-Gibbs sampler (Ritter and Tanner, 1992), whose implementation seems simple in our context. We illustrate its principle on the power parameter  $\delta_v$  ( $1 \leq v \leq S$ ). The same scheme can be followed for the remaining parameters  $\psi_{v_j} = \frac{1}{\tau_{v_j}}$  ( $k + 1 \leq j \leq S$ ) and  $(\omega_v, \alpha'_{v+}, \alpha'_{v-}, \beta'_v)'$  ( $1 \leq v \leq S$ ).

#### Griddy-Gibbs scheme

1. Select a grid of  $g$  points  $\delta_{v1} \leq \delta_{v2} \leq \dots \leq \delta_{vg}$  from a given interval  $[\delta_{v1}, \delta_{vg}]$ ; then, evaluate the conditional posterior  $f(\delta_v/\theta_{-\{\delta_v\}}, \epsilon^T)$  at each one of these points, to get  $f_{vs} = f(\delta_{vs}/\epsilon^T, \theta_{-\{\delta_v\}})$ , ( $1 \leq s \leq g$ ).

2. From the values  $f_{v1}, f_{v2}, \dots, f_{vg}$ , build the discrete distribution  $p_v(\cdot)$  defined at  $\delta_{vs}$  ( $1 \leq s \leq g$ ) by  $p(\delta_{vs}) = \frac{f_{vs}}{\sum_{j=1}^g f_{vj}}$ .  
This can be seen as an approximation to the inverse cumulative distribution of  $f(\delta_v/\theta_{-\{\delta_v\}}, \epsilon^T)$ .
3. Generate a number from the uniform distribution on  $(0, 1)$  and transform it using the discrete distribution  $p(\cdot)$  obtained in step (2) to get a random draw for  $\delta_v$ .

The following algorithm summarizes the Griddy-Gibbs sampler for drawing from the conditional posterior distribution  $f(\theta/\epsilon^T)$ .

**Algorithm 1 (Griddy-Gibbs sampler for the PAP-GARCH<sub>S</sub>( $p, q$ ) model).**

For  $l = 0, 1, \dots, L$ , let  $\theta^{(l)}$  be the Griddy-Gibbs draw of  $\theta$  at the  $l$ -th Gibbs iteration.

**Step 0** Specify starting values  $\psi_{vj}^{(0)}, \theta_v^{(0)} = (\omega_v^{(0)}, \alpha_{v+}^{(0)}, \alpha_{v-}^{(0)}, \beta_v^{(0)}, \delta_v^{(0)})'$  ( $k+1 \leq j \leq S, 1 \leq v \leq S$ ).

**Step 1** Repeat for  $l = 0, 1, \dots, L-1$ .

(a) For  $1 \leq v \leq S$ , sample  $\delta_v$  from  $f(\delta_v/\theta_{-\{\delta_v\}}, \epsilon^T)$  using the following Griddy scheme:

(a1) Select a grid  $\delta_{v1}^{(l+1)} \leq \delta_{v2}^{(l+1)} \leq \dots \leq \delta_{vg}^{(l+1)}$ .

(a2) For  $1 \leq s \leq g$  calculate  $f_{vs}^{(l+1)} = f(\delta_{vs}^{(l+1)}/\epsilon^T, \theta_{-\{\delta_v^{(l)}\}}^{(l)})$  from (14) and define the inverse distribution

$$p(\delta_{vs}^{(l+1)}) = \frac{f_{vs}^{(l+1)}}{\sum_{j=1}^g f_{vj}^{(l+1)}}, \quad 1 \leq s \leq g.$$

(a3) Generate a number  $u$  from the uniform  $(0, 1)$  distribution and transform it using the inverse distribution  $p(\cdot)$  to get  $\delta_v^{(l+1)}$ .

(b) Using a Griddy step similarly to (a), sample from  $f(\omega_v/\epsilon^T, \theta_{-\{\omega_v^{(l)}\}}^{(l)})$  to get  $\omega_v^{(l+1)}$ .

(c) Using a Griddy step similarly to (a), sample from  $f(\alpha_{vi+}/\epsilon^T, \theta_{-\{\alpha_{vi+}^{(l)}\}}^{(l)})$  to get  $\alpha_{vi+}^{(l+1)}, 1 \leq i \leq q$ .

(d) Using a Griddy step similarly to (a), sample from  $f(\alpha_{vi-}/\epsilon^T, \theta_{-\{\alpha_{vi-}^{(l)}\}}^{(l)})$  to get  $\alpha_{vi-}^{(l+1)}, 1 \leq i \leq q$ .

(e) Using a Griddy step similarly to (a), sample from  $f(\beta_{vj}/\epsilon^T, \theta_{-\{\beta_{vj}^{(l)}\}}^{(l)})$  to get  $\beta_{vj}^{(l+1)}, 1 \leq j \leq p$ .

(f) Using a Griddy step similarly to (a), sample from  $f(\psi_{vj}/\epsilon^T, \theta_{-\{\psi_{vj}^{(l)}\}}^{(l)})$  to get  $\psi_{vj}^{(l+1)}, k+1 \leq j \leq S$ .

**Step 3** Return  $\theta^{(l)}, l = 1, \dots, L$ .

The Griddy-Gibbs estimate  $\hat{\theta}_G$  of  $\theta$  is obtained by averaging the posterior draws of  $\theta$ , giving

$$\hat{\theta}_G = \frac{1}{L} \sum_{l=1}^L \theta^{(l)}. \quad (15)$$

It is important to note that the efficiency of the Griddy-Gibbs scheme is very sensitive to the choice of the grid  $\{\delta_{v1}, \dots, \delta_{vg}\}$ . Here, we follow a similar approach to Tsay (2010), according to which, at the  $l$ -th Gibbs iteration, the range of  $\delta_v$  is taken to be  $[\delta_{vl}^1, \delta_{vl}^2]$ , where

$$\delta_{vl}^1 = 0.6 \max(\delta_v^{(0)}, \delta_v^{(l-1)}), \quad \delta_{vl}^2 = 1.4 \min(\delta_v^{(0)}, \delta_v^{(l-1)}), \quad (16)$$

with  $\delta_v^{(l-1)}$  and  $\delta_v^{(0)}$  being the estimate of  $\delta_v$  at the  $(l-1)$ -th Gibbs iteration and the initial value, respectively.

Despite the fact that this choice may depend on the initial parameter draws, it gives quite satisfactory results that are not case-sensitive, as indicated by the used MCMC diagnostic tools (see A). On the other hand, it is well-known that the Griddy algorithm could be enhanced by considering a trapezoidal integration or the Simpson rule (Bauwens and Lubrano, 1998; Bauwens, Dufays and Rombouts, 2014). However, the use of the Griddy-Gibbs Step a2) in Algorithm 1 and the variable grid in (16) both seem suitable in our context, as we get fairly good estimates (see Section 4).

### 3.3 The Deviance Information Criterion

Selecting the best order and period of the  $PAP-GARCH_S(p, q)$  model is carried out using the  $DIC$  (Spiegelhalter et al. 2002). This criterion, which may be seen as a Bayesian generalization of the  $AIC$  (Akaike Information Criterion), is easily obtained from  $MCMC$  draws, without extra calculations.

In the context of the  $PAP-GARCH_S(p, q)$ ,  $DIC := DIC(p, q, S)$  is defined to be

$$DIC = -4E_{\theta/\epsilon^T} (\log (f (\epsilon^T / \theta))) + 2 \log (f (\epsilon^T / \bar{\theta})),$$

where  $f (\epsilon^T / \theta)$  is the likelihood given by (11)–(13) and  $\bar{\theta} = E (\theta / \epsilon^T)$  is the posterior mean of  $\theta$ . From the Griddy-Gibbs draws, the expectation  $E_{\theta/\epsilon^T} (\log (f (\epsilon^T / \theta)))$  can be estimated by averaging the conditional log-likelihood,  $\log f (\epsilon^T / \theta)$ , over the posterior draws of  $\theta$ . Moreover, the joint posterior mean estimate of  $\bar{\theta}$  can be approximated by the mean of the posterior draws of  $(\theta^{(l)})_{1 \leq l \leq L}$ .

The same model comparison criterion is used, when the proposed model is compared against several nested versions of it.

### 3.4 Forecast evaluation

To evaluate our  $PAP-GARCH_S(p, q)$  model, in terms of forecasting, we calculate in-sample and out-of-sample predictions of the volatility and the Value at Risk ( $VaR$ ), following a simulation-based approach; see, for example, Chen and So, 2006; Hoogerheide and van Dijk, 2010; Xia et al., 2017, among others. Once we generate the posterior draws  $\theta^{(l)}$  ( $l = 1, \dots, L$ ) from Algorithm 1, we can use them to readily generate in-sample volatilities  $(\sigma_t^{2(l)}, t = 1, \dots, T)$  according to (13), while replacing  $\theta$  by  $\theta^{(l)}$ , i.e.

$$\begin{aligned} \sigma_{nS+v}^{2(l)} &= \sigma_{nS+v}^2 (\theta^{(l)}), \quad l = 1, \dots, L, \\ &= \left( \omega_v^{(l)} + \sum_{i=1}^q \alpha_{vi+}^{(l)} (\epsilon_{Sn+v-i}^+)^{\delta_{v-i}^{(l)}} + \alpha_{vi-}^{(l)} (\epsilon_{Sn+v-i}^-)^{\delta_{v-i}^{(l)}} + \sum_{j=1}^p \beta_{vj}^{(l)} \sigma_{Sn+v-j}^{\delta_{v-j}^{(l)}} (\theta^{(l)}) \right)^{\frac{2}{\delta_v}}. \end{aligned} \quad (17)$$

Thus,  $(\sigma_t^{2(l)})_{1 \leq l \leq L}$  can be seen as a posterior sample from  $f (\sigma_t^2 / \epsilon^T)$ , and a Bayesian in-sample estimate of  $\sigma_t^2$  is given by

$$\hat{\sigma}_t^2 = \frac{1}{L} \sum_{l=1}^L \sigma_t^{2(l)}, \quad t = 1, \dots, T. \quad (18)$$

To forecast future volatilities  $\sigma_{T+h}^2$  ( $h = 1, 2, \dots$ ), we use a sequential method on  $h$  as follows:

- a. For  $h = 1$ , since  $\sigma_{T+1}^2$  depends on  $\epsilon_T, \epsilon_{T-1}, \dots$  which are available in the sample, we can easily compute  $\sigma_{T+1}^{2(l)}$  ( $1 \leq l \leq L$ ) from (17). This may be seen as a posterior sample from the predictive distribution  $f (\sigma_{T+1}^2 / \epsilon^T)$ . Therefore, the volatility forecast  $\hat{\sigma}_{T+1}^2$  is given, as in (18), by  $\hat{\sigma}_{T+1}^2 = \frac{1}{L} \sum_{l=1}^L \sigma_{T+1}^{2(l)}$ . Then,  $\epsilon_{T+1}^{(l)}$  ( $1 \leq l \leq L$ ) can be generated from  $f (\epsilon_{T+1} / \epsilon^T, \theta^{(l)})$  using (11) and the one-step ahead predicted return is given by  $\hat{\epsilon}_{T+1} = \frac{1}{L} \sum_{l=1}^L \epsilon_{T+1}^{(l)}$ .
- b. For  $h = 2$ , with  $\epsilon_{T+1}^{(l)}$  available, we can generate  $\sigma_{T+2}^{2(l)}$  ( $1 \leq l \leq L$ ) using (17). Then  $\epsilon_{T+2}^{(l)}$  ( $1 \leq l \leq L$ ) can be generated from  $f (\epsilon_{T+2} / \epsilon^{T+1}, \theta^{(l)})$  using again (11). This can be utilized to sample  $\sigma_{T+3}^{2(l)}$  ( $1 \leq l \leq L$ ) in the following step.
- c. For  $h \geq 3$  we can sequentially repeat the steps (i) and (ii).

As a by-product of the above volatility prediction scheme, we can also forecast the  $VaR$  (Francq and Zakoian 2019). Under the  $PAP-GARCH_S(p, q)$  model, the one-step  $VaR_{T+1}$  at the significance level  $\phi$  is the quantile of  $\epsilon_{T+1}$  (to within a sign “-”) at level  $\phi$ , i.e.,  $\phi = P (\epsilon_{T+1} \leq -VaR_{T+1})$ . More explicitly, it is given by

$$VaR_{T+1} = -F_{T+1}^{-1} (\phi) \sqrt{\sigma_{T+1}^2 (\theta)},$$

where  $F_{T+1}^{-1}(\cdot)$  is the inverse of the probability cumulative function of  $\eta_{T+1}$  and is  $S$ -periodic over time due to the periodic stationarity of the innovation  $\{\eta_t, t \in \mathbb{Z}\}$ . The  $l$ -th draw of  $VaR$  is given by  $VaR_{T+1}^{(l)} = -F_{T+1}^{-1}(\phi) \sqrt{\sigma_{T+1}^2(\theta^{(l)})}$  ( $1 \leq l \leq L$ ) and is readily obtained from the volatility forecast sample  $(\sigma_{T+1}^2(\theta^{(l)}))_{1 \leq l \leq L}$ . Hence, the estimated one-period ahead  $VaR$  for  $T + 1$  is given by

$$\widehat{VaR}_{T+1} = \frac{1}{L} \sum_{l=1}^L VaR_{T+1}^{(l)}.$$

More generally, the  $h$ -step  $VaR_{T+h}$  ( $h \geq 1$ ) is defined to be the  $\phi$ -quantile of  $\epsilon_{T+1} + \dots + \epsilon_{T+h}$ , i.e. a solution of the equation

$$\phi = P(\epsilon_{T+1} + \dots + \epsilon_{T+h} \leq -VaR_{T+h}).$$

Evaluating the  $h$ -step ahead  $VaR_{T+h}$  requires the estimation of the  $\phi$ -quantile of  $\epsilon_{T+1} + \dots + \epsilon_{T+h}$ , which can be easily obtained from the  $MCMC$  sample  $(\epsilon_{T+1}^{(l)} + \dots + \epsilon_{T+h}^{(l)})_{1 \leq l \leq M}$  as above.

### 4 Simulation study

Using simulated series generated from the 5-periodic  $PAP-GARCH_5(1, 1)$  model, this section examines the finite-sample performance of the Griddy-Gibbs estimate,  $\hat{\theta}_G$ , based on volatility and  $VaR$  forecasting as well as on other measures. The choice of  $S = 5$  is motivated by computational as well as practical considerations.

Two cases are considered for the distribution of the innovation. The first case refers to the standard Gaussian distribution, where the innovation sequence is *iid* with  $\eta_v \sim N(0, 1)$ , ( $1 \leq v \leq S$ ). The second case refers to the standardized Student-t distribution for which the innovation is *ipd*<sub>S</sub> with  $\eta_v \sim t(\frac{1}{\psi_v})$  ( $1 \leq v \leq S$ ).

The parameter  $\theta$  is fixed for each case so as to be in accordance with the empirical evidence, satisfying also the strict periodic stationarity condition (5). Moreover, at each season, the  $\alpha_{v+}$  is significantly different from  $\alpha_{v-}$ , to ensure asymmetry in the two models. In addition, different power values across seasons are allowed.

For each case, we consider 1000 replications of the  $PAP-GARCH_5(1, 1)$  series of sample size  $T = 1000$  for which we calculate the Griddy-Gibbs estimate  $\hat{\theta}_G$ . In evaluating  $\hat{\theta}_G$ , we use  $L = 1000$  iterations from which we discard the first 400 iterations. The initial parameter draw  $\theta^{(0)}$  is taken to be the true value of  $\theta$ . In the Griddy Gibbs iteration, the range of the grid is given by (16) and each component of  $\theta$  is generated using  $g = 300$  grid points.

The posterior means and standard deviations (*Std*) of the parameters are reported in Table 1 (Gaussian case) and Table 2 (Student-t case). From these tables, it can be seen that the parameters are well estimated with quite small bias and small standard deviations.

**Table 1:** Mean and standard deviation (Std) of  $\hat{\theta}_G$  for the 5-periodic Gaussian  $PAP-GARCH_5(1,1)$  series with  $T = 1000, L = 1000, g = 300$  and 1000 repetitions.

Season $v$		$\omega_v$	$\alpha_{v+}$	$\alpha_{v-}$	$\beta_v$	$\delta_v$
1	True	0.2	0.25	0.35	0.4	1.3
	Mean	0.2120	0.2786	0.3827	0.4108	1.3071
	Std	0.0501	0.0815	0.1062	0.1155	0.2894
2	True	0.1	0.15	0.3	0.2	1.2
	Mean	0.1050	0.1654	0.3166	0.1952	1.1970
	Std	0.0253	0.0479	0.0897	0.0508	0.1712
3	True	0.15	0.2	0.1	0.25	0.8
	Mean	0.1525	0.2167	0.1068	0.2407	0.8046
	Std	0.0265	0.0625	0.0298	0.0523	0.0868
4	True	0.4	0.3	0.2	0.15	1
	Mean	0.4115	0.3369	0.2241	0.1627	1.0294
	Std	0.0661	0.0982	0.0618	0.0467	0.2238
5	True	0.15	0.1	0.18	0.1	1.6
	Mean	0.1586	0.1125	0.1987	0.1088	1.5983
	Std	0.0381	0.0324	0.0563	0.0316	0.2562

Automatically generated rough PDF by ProofCheck from River Valley Technologies Ltd

**Table 2:** Mean and standard deviation (Std) of  $\hat{\theta}_G$  for the 5-periodic Student-t PAP0-GARCH<sub>5</sub>(1,1) series with  $\psi_v = \frac{1}{\tau_v}$ ,  $T = 1000$ ,  $L = 1000$ ,  $g = 300$  and 1000 repetitions.

Season $v$		$\omega_v$	$\alpha_{v+}$	$\alpha_{v-}$	$\beta_v$	$\delta_v$	$\psi_v$
1	True	0.2	0.25	0.35	0.4	1.3	0.2
	Mean	0.2295	0.2875	0.3722	0.4210	1.3420	0.1884
	Std	0.0423	0.0732	0.0608	0.0802	0.2952	0.0420
2	True	0.1	0.15	0.3	0.2	1.2	0.25
	Mean	0.1144	0.1701	0.3156	0.1951	1.2202	0.2355
	Std	0.0241	0.0354	0.0713	0.0407	0.2208	0.0410
3	True	0.15	0.2	0.1	0.25	0.8	0.3333
	Mean	0.1602	0.2135	0.1202	0.2281	0.8135	0.3140
	Std	0.0280	0.0534	0.0206	0.0312	0.0954	0.0720
4	True	0.4	0.3	0.2	0.15	1	0.25
	Mean	0.3859	0.3409	0.2388	0.1649	1.0584	0.2323
	Std	0.0570	0.0841	0.0530	0.0291	0.2014	0.0489
5	True	0.15	0.1	0.18	0.1	1.6	0.2
	Mean	0.1640	0.1233	0.2021	0.1203	1.6424	0.1896
	Std	0.0288	0.0287	0.0437	0.0258	0.3620	0.0411

To further evaluate the performance of the algorithm, we conduct volatility and VaR forecasting. We first generate a 5-periodic PAP-GARCH<sub>5</sub>(1, 1) series, using the parameters of Table 1 for the Gaussian case and of Table 2 for the Student-t case. Then, we compute the true volatility  $\sigma_t^2$  for  $t = 1, \dots, 1000 + h$ , where the horizon of prediction  $h$  is taken to be in the set  $\{1, \dots, 8\}$ . Finally, we compute the Griddy-Gibbs estimate  $\hat{\theta}_G$  from which, using (18), we obtain the prediction  $\hat{\sigma}_t^2$  and its standard deviation over the  $L$  Gibbs draws, for  $t = 1, \dots, 1000 + h$ .

For the Gaussian and Student-t cases, Table 3 and Table 4 show, respectively, the true volatility  $\sigma_{1000+h}^2$  for all  $h \in \{1, \dots, 8\}$ , and the posterior means and standard deviations (Std) of the estimated volatilities  $(\sigma_{1000+h}^{2(l)})$ . We conclude that the predicted volatilities are close to their corresponding true values, indicating that the proposed MCMC algorithm performs satisfactory.

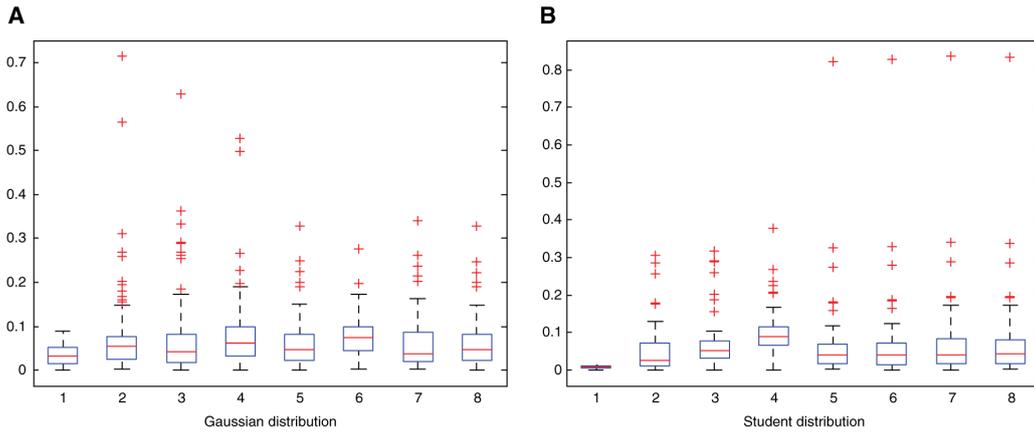
**Table 3:** Volatility forecasts from the 5-periodic Gaussian PAP-GARCH<sub>5</sub>(1,1) in Table 4 for the horizon  $h$  with  $h = 1, \dots, 8$ .

	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	$h = 8$
True value $\sigma_{1000+h}^2$	0.2487	0.2576	0.2556	0.2481	0.2777	0.2273	0.2370	0.2493
Mean: $\hat{\sigma}_{1000+h}^2$	0.2321	0.2541	0.2550	0.2571	0.2480	0.2526	0.2521	0.2468
Std	0.0282	0.0695	0.0672	0.0801	0.0618	0.0887	0.0696	0.0642

**Table 4:** Volatility forecasts from the 5-periodic Student-t PAP-GARCH<sub>5</sub>(1, 1) in Table 2 for the horizon  $h$  with  $h = 1, \dots, 8$ .

	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	$h = 8$
True value $\sigma_{1000+h}^2$	0.3372	0.3269	0.4194	0.3442	0.2907	0.3570	0.3171	0.2771
Mean: $\hat{\sigma}_{1000+h}^2$	0.3308	0.3741	0.3911	0.3726	0.3729	0.3958	0.3701	0.3583
Std	0.035	0.1255	0.1245	0.1178	0.1145	0.1499	0.1021	0.0849

On the basis of 100 replications of this generated series, we calculate the mean absolute error (MAE)  $|\hat{\sigma}_{1000+h}^2 - \sigma_{1000+h}^2|$  for  $h = 1, \dots, 8$  for both cases and obtain the corresponding boxplots (Figure 1). The MAEs are small enough for all time horizons. Of course, the MAE of volatilities prediction becomes large as the time horizon  $h$  increases.



**Figure 1:** Mean absolute error of the  $h$ -step ahead absolute prediction error of the volatility for  $h = 1, \dots, 8$ . Boxplot of the MAE of  $|\hat{\sigma}_{1000+h}^2 - \sigma_{1000+h}^2|$  ( $h = 1, \dots, 8$ ) for the Gaussian (panel (A)) and Student-t (panel (B)) cases.

Next, we turn our attention to  $VaR$  forecasting. The posterior mean and standard deviation of the  $h$ -step  $VaR_{1000+h}$  estimate ( $h = 1, \dots, 8$ ) at the probability levels  $\phi = 0.01$  and  $\phi = 0.05$  are both presented in Table 5 (for the Gaussian case) and Table 6 (for the Student-t case). We observe that for both cases, the standard deviations of the  $VaR$  forecasts are quite small across all time horizons  $h$  and naturally increase with respect to  $h$ . For the Student-t case, the  $VaR$  estimates are slightly more precise than those obtained from the Gaussian case.

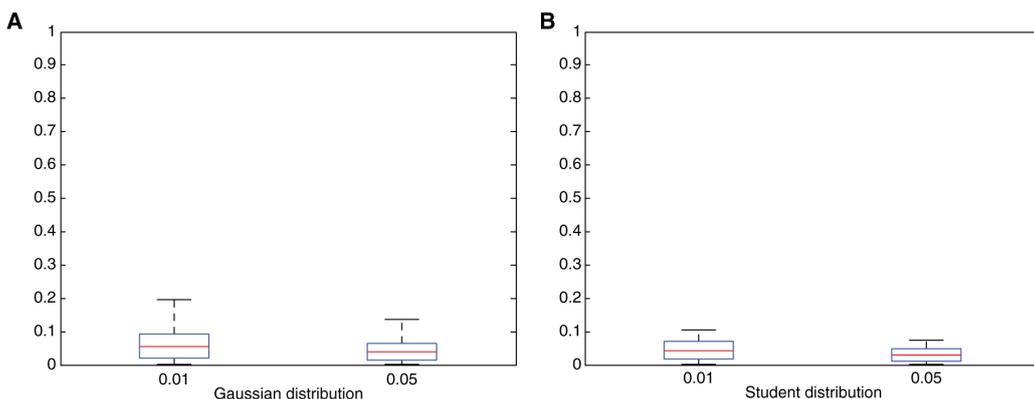
**Table 5:** Mean and standard deviations of  $VaR_{1000+h}$  forecast at the levels  $\phi = 0.01$  and  $\phi = 0.05$  for  $h = 1, \dots, 8$  using the 5-periodic Gaussian PAP-GARCH<sub>5</sub>(1, 1) in Table 1.

$(VaR_{1000+h}^{(\phi)})_t$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	$h = 8$
Mean ( $\phi = 0.01$ )	1.1654	1.6949	2.0750	2.3448	2.6313	2.8606	3.0818	3.2881
Std ( $\phi = 0.01$ )	0.0104	0.0202	0.0607	0.0494	0.0426	0.0477	0.0841	0.0755
Mean ( $\phi = 0.05$ )	0.8083	1.1663	1.4353	1.6113	1.8355	1.9619	2.1566	2.2774
Std ( $\phi = 0.05$ )	0.0118	0.0286	0.0340	0.0183	0.0304	0.0300	0.0298	0.0507

**Table 6:** Mean and standard deviations of  $VaR_{1000+h}$  forecast at the levels  $\phi = 0.01$  and  $\phi = 0.05$  for  $h = 1, \dots, 8$  using the 5-periodic Student-t PAP-GARCH<sub>5</sub>(1, 1) in Table 2.

$(VaR_{1000+h}^{(\phi)})_t$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	$h = 8$
Mean ( $\phi = 0.01$ )	1.6320	2.5928	3.0001	3.5134	3.7597	4.2026	4.3246	4.7986
Std ( $\phi = 0.01$ )	0.4251	0.2396	0.2364	0.1014	0.0861	0.1581	0.0655	0.0978
Mean ( $\phi = 0.05$ )	0.9040	1.4366	1.8599	2.1883	2.4368	2.6574	2.8440	3.0553
Std ( $\phi = 0.05$ )	0.0285	0.0324	0.0515	0.0608	0.0864	0.1102	0.0602	0.1439

For the one-step ahead prediction corresponding to  $h = 1$ , the boxplots of the MAE of  $|\widehat{VaR}_{1001} - VaR_{1001}|$  for the two cases in question are given in Figure 2.



**Figure 2:** Mean absolute error of the one-step ahead absolute Value-at-Risk error. Boxplot of the MAE of  $|\widehat{VaR}_{1001} - VaR_{1001}|$  for the Gaussian (panel (A)) and Student-t (panel (B)) cases at the levels  $\phi = 0.01$  and  $\phi = 0.05$ .

The boxplots in Figure 2 also show that the MAEs are quite small. Therefore, the estimates have a good predictive performance.

## 5 An application to daily S&P 500 returns

To show the usefulness of the proposed modelling framework, we fit a modified version of the  $PAP-GARCH_S(1, 1)$  model with periodicity  $S = 5$  to daily returns on the S&P500 index (closing value). The relevant literature has documented that daily financial asset returns are characterized by the day-of-the-week effect, suggesting the existence of periodicity of magnitude  $S = 5$  in volatility (Bollerslev and Ghysels, 1996; Franses and Paap, 2000; Tsiakas, 2006; Osborn, Savva and Gill, 2008; Regnard and Zakoian, 2011).

Because of the presence of holidays in our sample, model (1) in which  $\theta_v = \theta_{nS+v}$  ( $1 \leq v \leq S, n \in \mathbb{Z}$ ) does not seem a suitable choice; each day of a week may correspond to a different specification as compared to the same day of the week before. So, under  $S = 5$  we consider the following variant of model (1), denoted by  $PAP-GARCH_5^*(1, 1)$ :

$$\begin{cases} \epsilon_t = \sigma_t^{\delta_{d(t)}} \eta_t \\ \sigma_t^{\delta_{d(t)}} = \omega_{d(t)} + \alpha_{d(t)+} (\epsilon_{t-1}^+)^{\delta_{d(t-1)}} + \alpha_{d(t)-} (\epsilon_{t-1}^-)^{\delta_{d(t-1)}} + \beta_{d(t)} \sigma_{t-1}^{\delta_{d(t-1)}} \end{cases} \quad (19)$$

in which  $d(t)$  is defined to be

$$d(t) = \begin{cases} 1 & \text{if the day corresponding to } t \text{ is a Monday} \\ 2 & \text{if the day corresponding to } t \text{ is a Tuesday} \\ \vdots & \\ 5 & \text{if the day corresponding to } t \text{ is a Friday.} \end{cases}$$

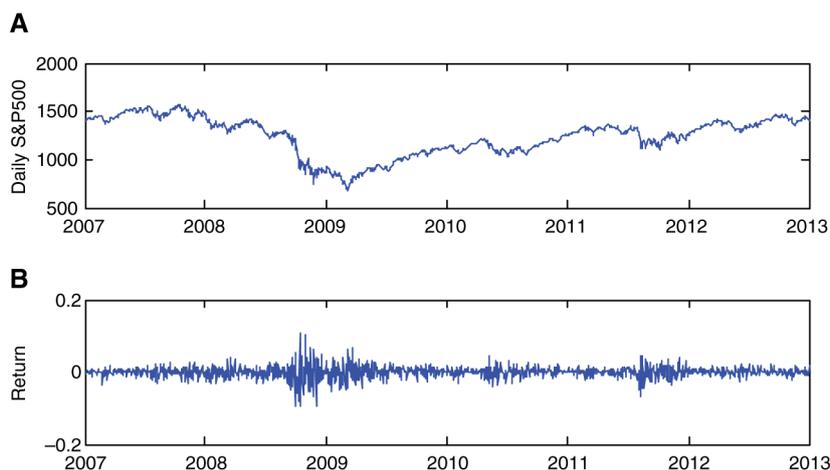
The  $PAP-GARCH_5^*(1, 1)$  specification (19) with missing values (Franses and Paap, 2000; Regnard and Zakoian, 2011; Aknouche, 2017) is capable of accommodating the day-of-the-week effect.

As in our simulation study, two cases of the distribution of innovation  $\{\eta_t, t \in \mathbb{Z}\}$  in (19) are examined. In the first case,  $\eta_1, \dots, \eta_S$  are normally distributed with mean zero and unit variance ( $\eta_v \sim N(0, 1)$ ), whereas in the second case  $\eta_1, \dots, \eta_S$  are (standardized) Student-t distributed with  $\tau_1, \dots, \tau_S > 2$  degrees of freedom, respectively.

We estimate the two models using our proposed algorithmic scheme, which we run for  $L = 1000$  iterations with a burn-in period of 400 draws. The initial parameter vector  $\theta^{(0)}$  is chosen as follows. The initial power parameter  $\delta_v$  at a day  $v$  is taken to be inversely proportional to the Kurtosis of that day. For the remaining parameters, we take the values obtained from the estimation of a  $GARCH(1, 1)$  model, applied to the series of each day. In the Griddy-Gibbs step, 500 grid points are used and the range of parameters at the  $l$ -th Gibbs iteration is given in (16).

### 5.1 The data and the day-of-the-week effect

Our dataset consists of  $T = 1509$  observations, spanning the period from January 01, 2007 to December 31, 2012. The time series plots of the index (panel (A)) and its return (panel (B)) are presented in Figure 3. The same dataset was considered by Chan and Grant (2016) and Aknouche (2017).



**Figure 3:** Daily levels and returns of the S&P 500 from January 2007 to December 2012. (A) level, (B) return.

Some descriptive statistics concerning the S&P500 returns, such as the absolute returns, the squared returns and the log-absolute returns can be found in Aknouche (2017). To highlight the day-of-the week effect, Table 7 shows for each (trading) day the average return, the volatility (approximated by the absolute value), the kurtosis, and skewness, where it may be seen that these measures are quite different from one day to another.

**Table 7:** Day of the week effect in daily S&P 500 returns.

	Sample size	Mean of ( $\epsilon_t$ )	Mean of ( $ \epsilon_t $ )	Mean of ( $\epsilon_t^2$ )	Kurtosis	Skewness
Full series	1509	$4.4711e^{-06}$	0.0102	$2.4646e^{-04}$	10.4975	-0.2643
1 Monday	284	0.0012	0.0108	$2.6627e^{-04}$	9.4713	1.0406
2 Tuesday	308	-0.0003	0.0099	$2.2774e^{-04}$	10.2425	-1.5226
3 Wednesday	311	0.0001	0.0109	$2.5983e^{-04}$	7.0415	-0.7051
4 Thursday	305	-0.0002	0.0088	$1.5129e^{-04}$	5.5615	0.0655
5 Friday	301	-0.0007	0.0108	$3.2866e^{-04}$	12.7415	-0.1795

This is also confirmed by Figure 4 which exhibits the kernel estimate of the distribution of return for each trading day together with the full series. These distributions seem to have different shapes. The same finding may be observed in the boxplot of each day (Figure 5). This reinforces the intuition that a periodic model with a periodic innovation might be better in explaining the day specificities than a non-periodic model.

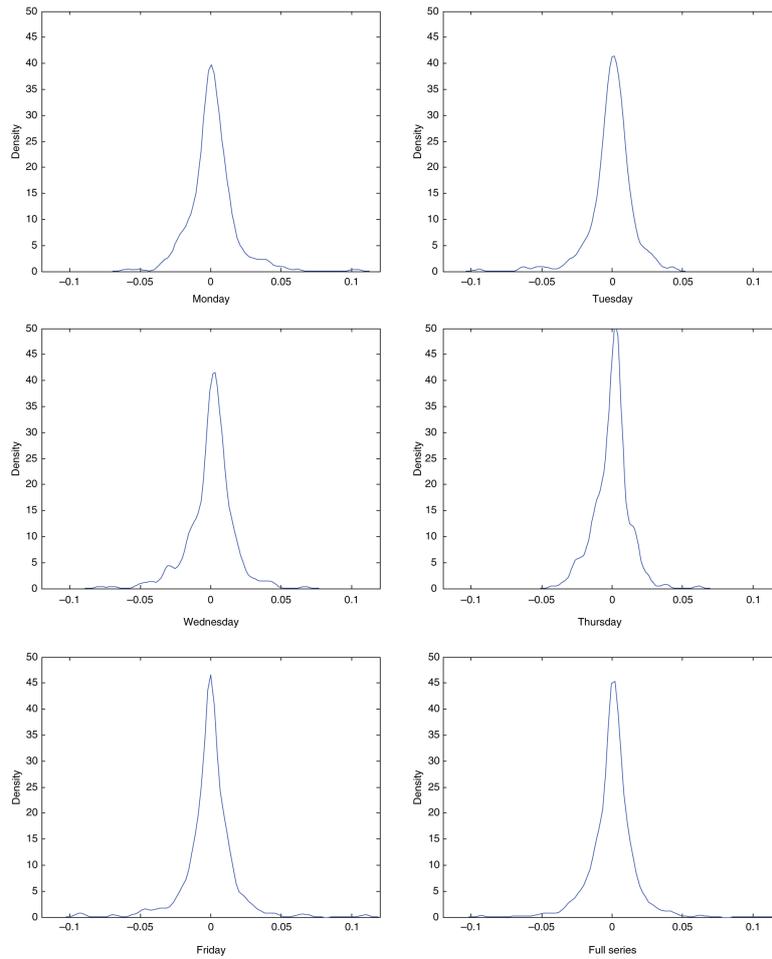


Figure 4: Kernel density estimate of the distribution of returns for the full series and each trading day.

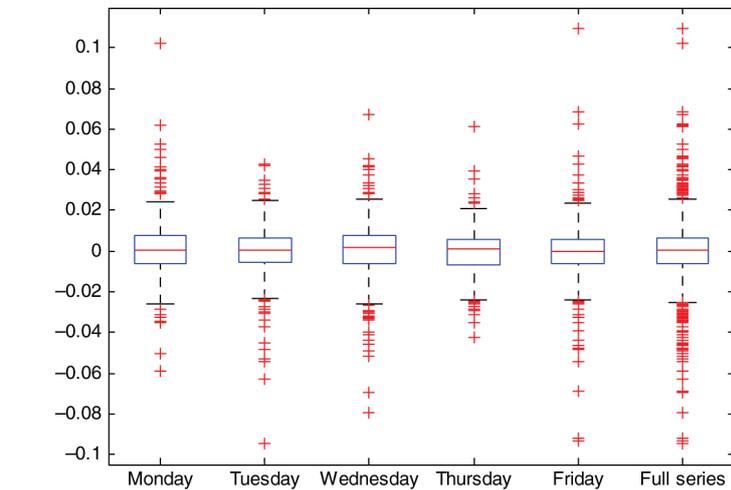


Figure 5: Boxplots of the distribution of returns for the full series and each trading day.

### 5.2 The $PAP-GARCH_5^*(1, 1)$ model with Gaussian errors

As a first step in our empirical analysis, we compare the  $PAP-GARCH_5^*(1, 1)$  given by (19) with other alternative competing models, some of which have already been used in the relevant literature. The set of these models includes the  $PAP-GARCH_5(1, 1)$ , given by (1), for  $S \in \{2, \dots, 4\}$ , the standard  $GARCH(1, 1)$ , the  $P-GARCH_5^*(1, 1)$  of Bollerslev and Ghysels (1996), the  $AP-GARCH(1, 1)$  of Ding, Granger, and Engle (1993) and the 5-periodic threshold  $GARCH_5^*(1, 1)$ , which is named  $PT-GARCH_5^*(1, 1)$ . For all models in question, the innovations are *iid* having a standard Gaussian distribution.

Table 8 shows the estimated *DIC* values along with the rank of each one of the competing models. It can be observed that the *DIC*s are significantly different from each other. In particular, the best model is the  $PAP-GARCH_5^*(1, 1)$ , as it produced the smallest *DIC* ( $-7034.3004$ ). The second best position is occupied by the  $AP-GARCH(1, 1)$  model, followed by the standard  $GARCH(1, 1)$  model, which is preferred to the threshold and periodic  $GARCH$  models. Thus, we retain the  $PAP-GARCH_5^*(1, 1)$  model in our analysis. Note that the ranking of the models is related to this particular data set and therefore we can not extract safe conclusions as to why, for example the second and third best models are the  $AP-GARCH(1, 1)$  and the  $GARCH(1, 1)$ , respectively. In any case, the proposed  $PAP-GARCH_5^*(1, 1)$  is the dominant one.

**Table 8:** *DIC* values for various conditionally Gaussian periodic GARCH models.

Model	$AP-GARCH_1(1, 1)$	$PAP-GARCH_2(1, 1)$	$PAP-GARCH_3(1, 1)$	$PAP-GARCH_4(1, 1)$
<i>DIC</i>	$-7030.6480$	$-6942.8207$	$-7019.6105$	$-7000.7523$
Rank	2	7	4	5
Model	$PAP-GARCH_5^*(1, 1)$	$PT-GARCH_5^*(1, 1)$	$P-GARCH_5^*(1, 1)$	$GARCH(1, 1)$
<i>DIC</i>	$-7034.3004$	$-6930.7031$	$-6971.5962$	$-7025.0591$
Rank	1	8	6	3

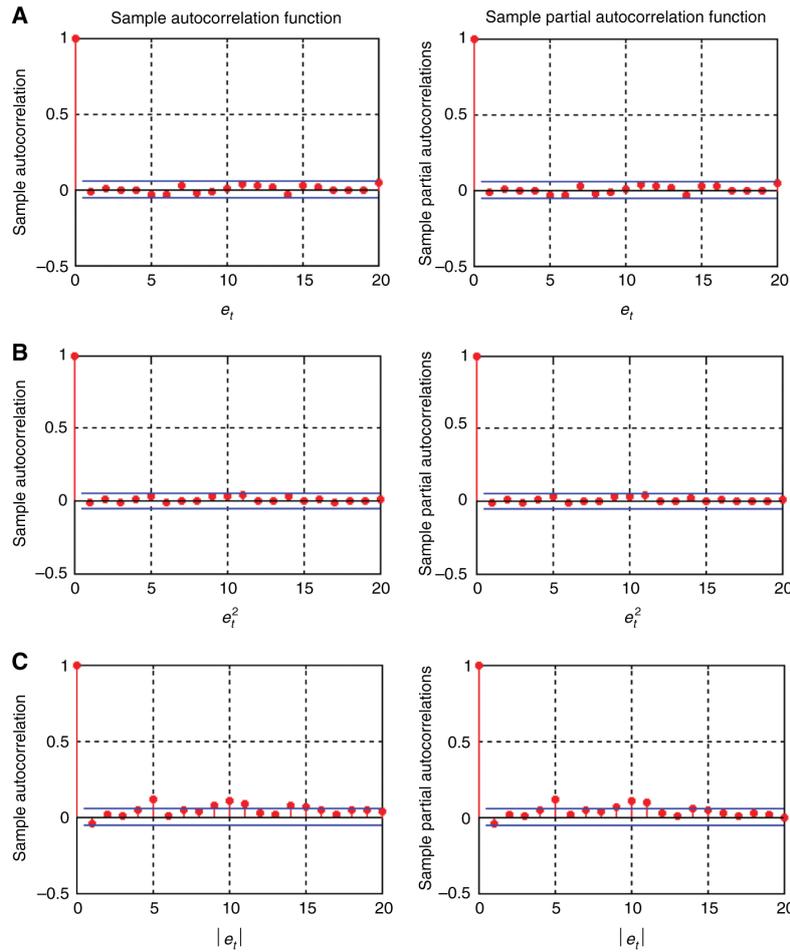
Table 9 displays the posterior means of the parameters for the  $PAP-GARCH_5^*(1, 1)$  model, along with their standard deviations (*Std*), their numerical standard errors (*NSE*) and their relative numerical inefficiency (*RNI*); see appendix B for a description of the *NSE* and *RNI*.

**Table 9:** Bayesian Griddy-Gibbs estimates of the Gaussian  $PAP-GARCH_5^*(1, 1)$  model for the S&P500 returns.

Day	Parameters	Mean	Std	RNI	NSE
Monday	$\delta_1$	0.6183	0.0258	3.8105	0.0016
	$\omega_1$	$0.2830e^{-04}$	$0.0207e^{-04}$	3.0642	$0.1146e^{-06}$
	$\alpha_{1+}$	0.3056	0.0630	3.2201	0.0037
	$\alpha_{1-}$	0.5115	0.0391	3.7839	0.0020
	$\beta_1$	0.7466	0.0181	2.9539	0.0012
Tuesday	$\delta_2$	0.6239	0.0245	4.2003	0.0016
	$\omega_2$	$0.2023e^{-04}$	$0.0125e^{-04}$	3.0120	$0.0687e^{-06}$
	$\alpha_{2+}$	0.2028	0.0379	3.1586	0.0022
	$\alpha_{2-}$	0.3265	0.0468	3.2612	0.0024
	$\beta_2$	0.8536	0.0206	2.4796	0.0015
Wednesday	$\delta_3$	0.9222	0.0371	3.5016	0.0022
	$\omega_3$	$0.4265e^{-04}$	$0.0709e^{-04}$	2.9078	$0.3821e^{-06}$
	$\alpha_{3+}$	0.2688	0.0646	2.9837	0.0034
	$\alpha_{3-}$	0.4649	0.0423	3.4362	0.0021
	$\beta_3$	0.6673	0.0191	2.4585	0.0013
Thursday	$\delta_4$	0.8139	0.0595	3.7871	$0.0037e^{-06}$
	$\omega_4$	$0.1062e^{-04}$	0.0343	3.1505	0.1923
	$\alpha_{4+}$	0.1244	0.0241	2.7845	0.0013
	$\alpha_{4-}$	0.2627	0.0415	3.1916	0.0023
	$\beta_4$	0.7912	0.0276	2.6221	0.0021
Friday	$\delta_5$	0.4589	0.0224	4.0096	$0.0014e^{-06}$
	$\omega_5$	0.2352	0.0462	2.8572	0.2470
	$\alpha_{5+}$	0.3220	0.0686	2.1526	0.0041
	$\alpha_{5-}$	0.3935	0.0647	2.9243	0.0035
	$\beta_5$	0.6785	0.0203	2.4539	0.0015

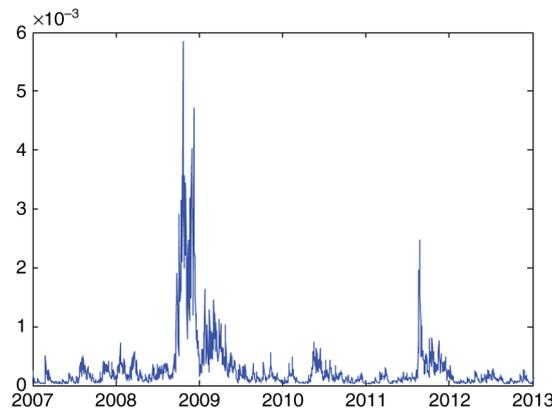
The performance of our algorithm is satisfactory, judging from the small standard deviations, *RNI* and *NSE* values. Moreover, the parameters are quite different from 1 day to another. In particular,  $\alpha_{v-}$  is quite different from  $\alpha_{v+}$  for all  $v \in \{1, \dots, S\}$ , which implies the presence of asymmetric volatility in the data. On the other hand, the estimated model is characterized by high persistence and overall the estimates are comparable with those from similar models in the literature when  $S = 1$  (e.g. Pan, Wang and Tong, 2008; Xia et al., 2017).

In addition, simple and partial (sample) autocorrelations of the residuals given by  $e_t = \frac{\epsilon_t}{\sigma_t(\hat{\theta})}$  ( $t = 1, \dots, T$ ), the squared residuals ( $e_t^2$ ) and the absolute residuals ( $|e_t|$ ) are shown in Figure 6. The residuals look like an independent white noise that validates the estimated model. Note that these results are stable enough under different initial values and different numbers of iterations for the Griddy-Gibbs sampler.



**Figure 6:** Sample autocorrelations of the residuals (panel (A)), the squared residuals (panel (B)) and absolute residuals (panel (C)) under the Gaussian innovation.

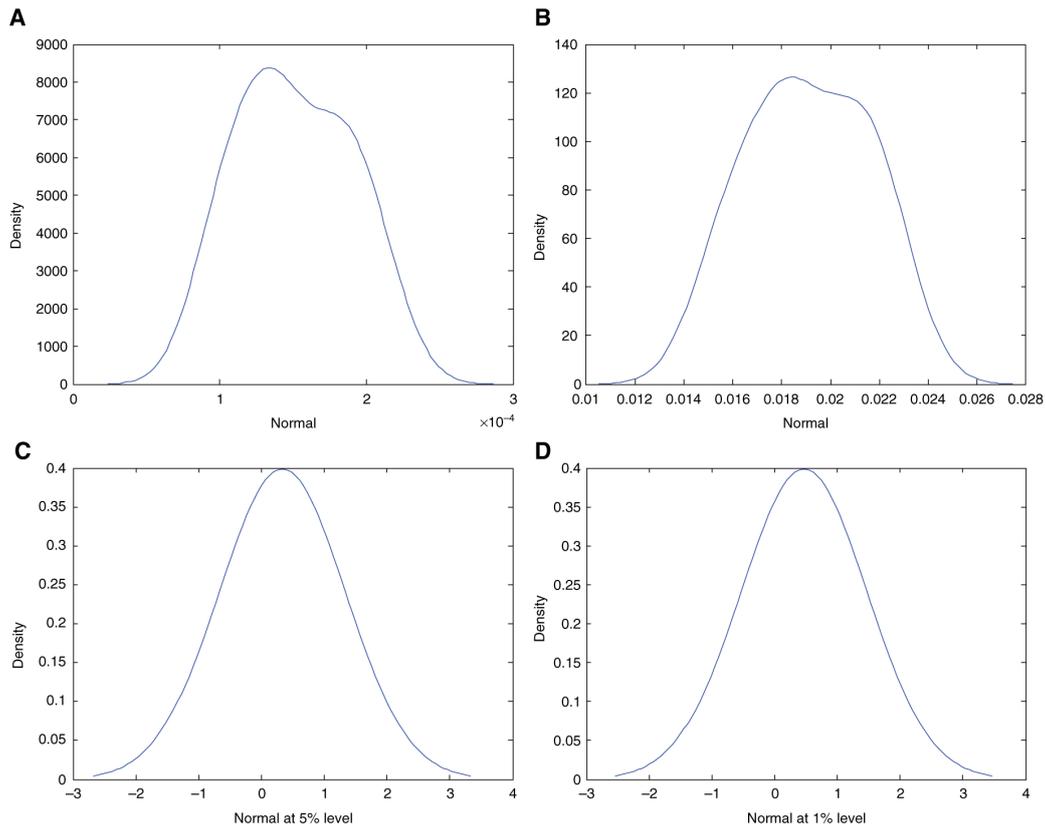
The estimated volatility induced by the Gaussian  $PAP-GARCH_5^*(1, 1)$  is displayed in Figure 7, showing a pattern similar to the actual variability of the S&P500 returns.



**Figure 7:** Volatility induced by the Gaussian  $PAP-GARCH_5^*(1, 1)$  model.

We finally use the estimated  $PAP-GARCH_5^*(1, 1)$  model to get one-step ahead predictive distribution (cf. Figure 8) of the volatility (panel (A)), the return (panel (B)) and the value at risk at levels 5% (panel (C)) and 1% (panel (D)). The posterior means of these predictive distributions together with their MCMC standard deviations are reported in Table 10. In summary, the estimated  $PAP-GARCH_5^*(1, 1)$  model seems to give good results.

Automatically generated rough PDF by ProofCheck from River Valley Technologies Ltd



**Figure 8:** The one-step-ahead predictive distribution of the volatility (panel (A)), the return (panel (B)) and the value at risk at 5% (panel (C)) and 1% (panel (D)) for the Gaussian case.

(A) Volatility  $\hat{\sigma}_{T+1}$ , (B) return  $\hat{e}_{T+1}$ , (C) 5%  $\widehat{VaR}_{T+1}$ . (D) 1%  $\widehat{VaR}_{T+1}$ .

**Table 10:** One day-ahead prediction of the volatility, the return and the value at risk at levels 5% and 1% for the S&P500 under the Gaussian assumption.

	$\hat{e}_{T+1}$	$\hat{\sigma}_{T+1}^2$	$\alpha = 0.01 \widehat{VaR}_{T+1}$	$\alpha = 0.05 \widehat{VaR}_{T+1}$
Mean	0.0191	$1.5111e^{-04}$	0.4343	0.3071
Std	0.0024	$3.6861e^{-05}$	$0.8348e^{-4}$	$0.5009e^{-4}$

### 5.3 The $PAP-GARCH_5^*(1, 1)$ model with Student-t errors

We now estimate a  $PAP-GARCH_5^*(1, 1)$  given by (19) based on the same S&P500 series but under the assumption that the innovation  $\{\eta_t, t \in \mathbb{Z}\}$  is  $ipd_5$  with a standardized Student-t distribution, i.e.  $\frac{\tau_v}{\tau_v - 2} \eta_v \sim t_{(\psi_v^{-1})}$  ( $1 \leq v \leq 5$ ). The means of the posterior estimates, their standard deviations, their RNI and their NSE are reported in Table 11. The estimates are quite good considering their small *std*, RNI and NSE. The periodicity of the model is significant in view of the estimates across seasons which are quite different.

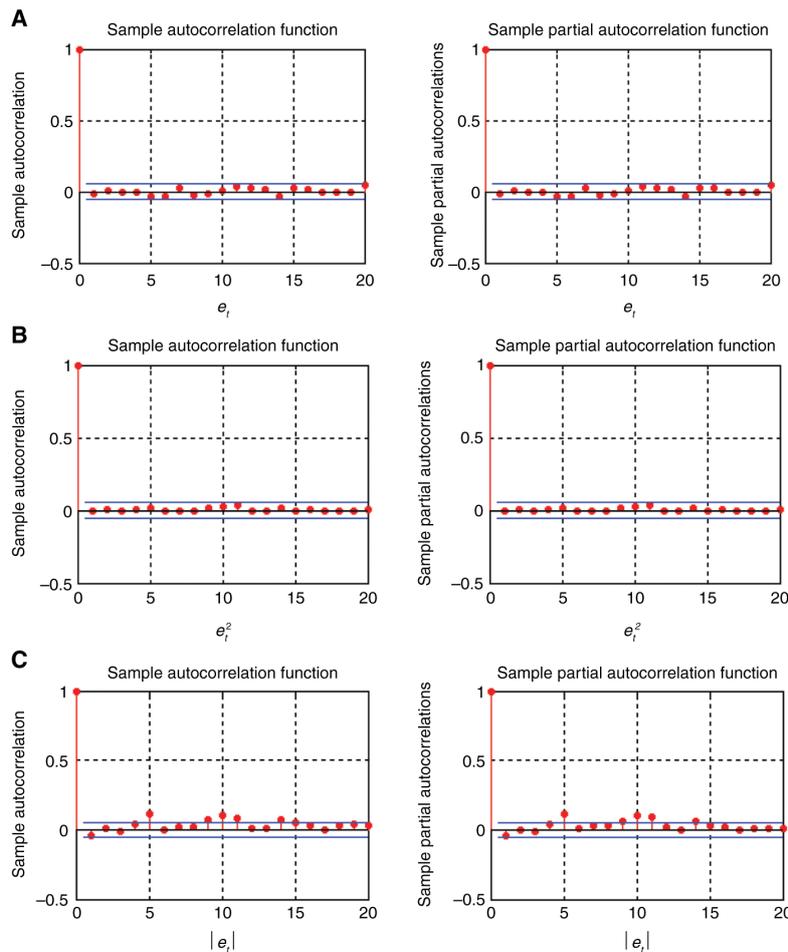
**Table 11:** Bayesian Griddy-Gibbs estimates of the Student-t  $PAP-GARCH_5^*(1, 1)$  model for the S&P500 returns.

Day	Parameters	Mean	Std	RNI	NSE
Monday	$\delta_1$	0.5680	0.0146	4.2532	0.0078
	$\omega_1$	$0.1441e^{-04}$	$0.0038e^{-04}$	2.1617	$0.7803e^{-07}$
	$\alpha_{1+}$	0.2785	0.0360	4.2542	0.0019
	$\alpha_{1-}$	0.3052	0.0117	4.1500	0.0063
	$\beta_1$	0.7001	0.0900	4.1542	0.0048
Tuesday	$\psi_1$	0.0910	0.0082	4.2376	0.0044
	$\delta_2$	0.6438	0.0166	4.3200	0.0088
	$\omega_2$	$0.1027e^{-04}$	$0.0037e^{-04}$	2.2259	$0.7852e^{-07}$

Automatically generated rough PDF by ProofCheck from River Valley Technologies Ltd

Wednesday	$\alpha_{2+}$	0.1798	0.0230	4.4200	0.0012
	$\alpha_{2-}$	0.2227	0.0860	4.1011	0.0046
	$\beta_2$	0.8098	0.0104	4.3542	0.0056
	$\psi_2$	0.0993	0.0090	4.3654	0.0048
	$\delta_3$	0.9467	0.0243	4.5312	0.0130
	$\omega_3$	$0.2395e^{-04}$	$0.0018e^{-04}$	2.7221	$0.4291e^{-07}$
	$\alpha_{3+}$	0.2918	0.0037	4.2200	0.0020
Thursday	$\alpha_{3-}$	0.2868	0.0110	4.2104	0.0059
	$\beta_3$	0.6326	0.0810	4.5425	0.0043
	$\psi_3$	0.1241	0.0112	4.1298	0.0060
	$\delta_4$	0.8521	0.0219	4.4465	0.0117
	$\omega_4$	$0.0917e^{-04}$	$0.0010e^{-04}$	2.3329	$0.2066e^{-07}$
	$\alpha_{4+}$	0.1689	0.0220	4.5420	0.0012
	$\alpha_{4-}$	0.2003	0.0770	4.2507	0.0041
Friday	$\beta_4$	0.7988	0.0103	4.2542	0.0055
	$\psi_4$	0.1655	0.0149	4.2145	0.0080
	$\delta_5$	0.4734	0.0122	4.0213	0.0065
	$\omega_5$	$0.1376e^{-04}$	$0.0024e^{-04}$	2.0988	$0.5007e^{-07}$
	$\alpha_{5+}$	0.3592	0.0460	4.2425	0.0025
	$\alpha_{5-}$	0.3471	0.0134	4.2762	0.0071
	$\beta_5$	0.6664	0.0860	4.2114	0.0046
	$\psi_5$	0.0828	0.0075	4.2376	0.0040

Moreover, the model is able to absorb asymmetry since  $\alpha_{v-}$  and  $\alpha_{v+}$  are quite different for all  $v \in \{1, \dots, 5\}$ . Finally, with respect to the sample autocorrelations of  $(e_t)$ ,  $(e_t^2)$  and  $(|e_t|)$  (cf. Figure 9), the residuals seem compatible with the independence assumption.



**Figure 9:** Sample autocorrelations of the residuals (panel (A)), the squared residuals (panel (B)) and the absolute residuals (panel (C)) under Student innovations.

The Student-t  $PAP-GARCH_5^*(1, 1)$  model does worse than the Gaussian  $PAP-GARCH_5^*(1, 1)$  model, in terms of model fit; the  $DIC$  value for the former is 13421.0183. Furthermore, by comparing the Table 9 and Table 11,

we observe that the two models produced quite similar estimation results for most of the parameters ( $\delta$ ,  $\alpha_{v-}$  and  $\alpha_{v+}$ , etc.).

We repeated the same forecast analysis for the Student-t  $PAP-GARCH_5^*(1, 1)$  model, as did for the Gaussian  $PAP-GARCH_5^*(1, 1)$  model. In particular, the volatility estimated by the Student-t version of the model is shown in Figure 10. Under the Student-t assumption, we calculate the one-step ahead predictive distribution (cf. Figure 11) of the volatility (panel (A)), the return (panel (B)) and the value at risk at levels 5% (panel (C)) and 1% (panel (D)). The posterior means of these predictive distributions together with their MCMC standard deviations are reported in Table 12. The forecast results obtained from the Student-t  $PAP-GARCH_5^*(1, 1)$  model are quite similar to those obtained from the Gaussian  $PAP-GARCH_5^*(1, 1)$ . This is justified by the fact that the two model specifications are quite similar.

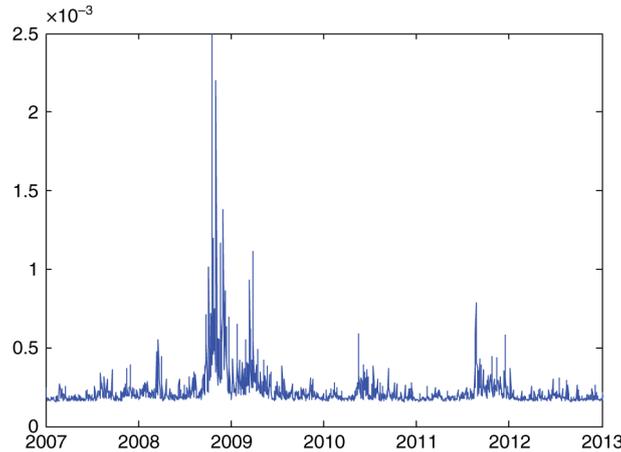


Figure 10: Volatility induced by the Student-t  $PAP-GARCH_5^*(1, 1)$  model.

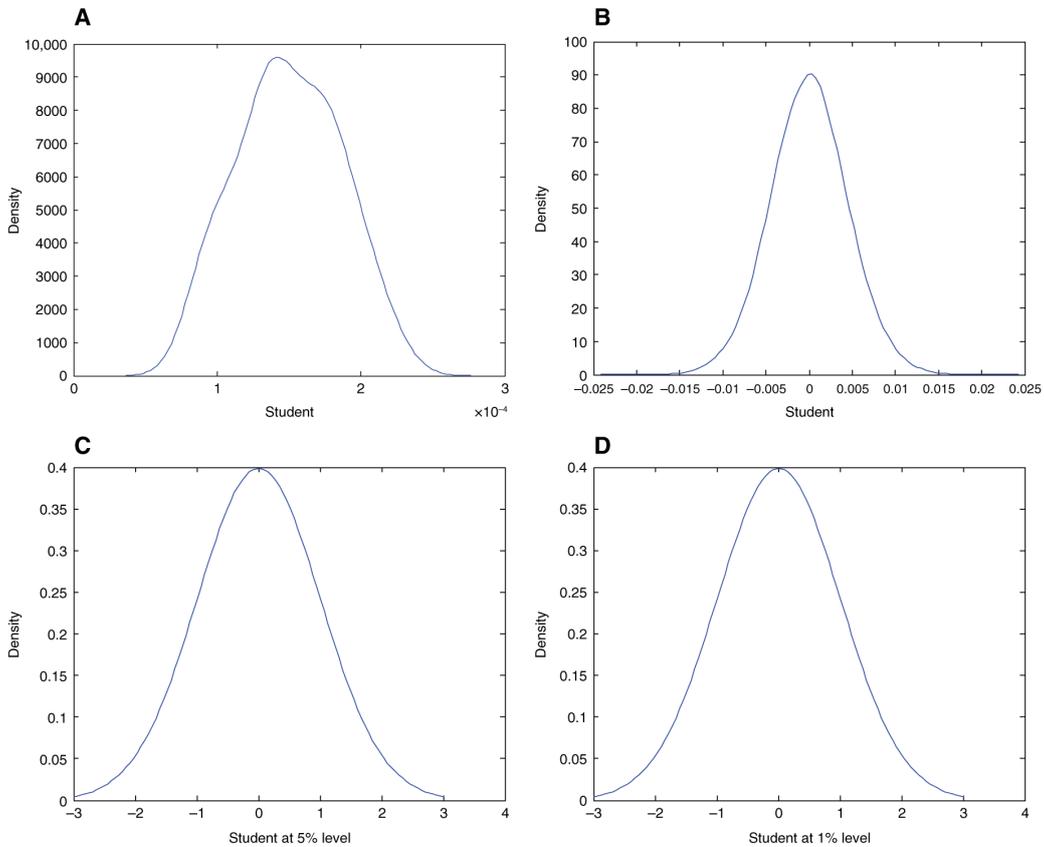


Figure 11: The one-step-ahead predictive distribution of the volatility (panel (A)), the return (panel (B)) and the value at risk at 5% (panel (C)) and 1% (panel (D)) for the Student-t case.

(A) Volatility  $\hat{\sigma}_{T+1}$ . (B) return  $\hat{e}_{T+1}$ . (C) 5%  $\widehat{VaR}_{T+1}$ . (D) 1%  $\widehat{VaR}_{T+1}$ .

**Table 12:** One day-ahead prediction of the volatility, the return and the value at risk at levels 5% and 1% for the S&P500 under the Student-t assumption.

	$\hat{\epsilon}_{T+1}$	$\hat{\sigma}_{T+1}^2$	$\alpha = 0.01 \overline{\text{VaR}}_{T+1}$	$\alpha = 0.05 \overline{\text{VaR}}_{T+1}$
Mean	-0.0020	1.1425e <sup>-04</sup>	0.4410	0.3118
Std	1.5515e <sup>-04</sup>	1.7553e <sup>-05</sup>	0.3278e <sup>-04</sup>	0.1945e <sup>-04</sup>

## 6 Conclusions

We propose an easily adaptable Markov Chain Monte Carlo algorithm for a flexible periodic asymmetric power GARCH model under various distributions for the periodic innovation sequence. This model is designed to account for periodicity in volatility as well as for other important features of volatility, including asymmetry and periodic power dependence. Forecast evaluation of the proposed model is conducted with volatility and Value at Risk forecasting, while model comparison is based on the Deviance Information Criterion. Our simulation studies showed the good performance of the proposed algorithm. We illustrated the usefulness of the model with an application to the daily S&P500 returns. Our model had the best fit to the data and produced good results.

## Acknowledgement

The authors are deeply grateful to the Editor-in-Chief and two referees for their relevant comments and suggestions that were very helpful in improving the first draft.

## Appendix A

### Proofs

*Proof of Theorem 1.*

(i) *Sufficiency:* Equation (3) may be cast in the following system of  $S$  recurrence equations

$$Y_{nS+v} = A_{nS+v} Y_{(n-1)S+v} + B_{nS+v}, \quad n \in \mathbb{Z}, \quad v \in \{0, \dots, S-1\}, \quad (20)$$

where  $A_{nS+v} = \prod_{i=0}^{S-1} A_{nS+v-i}$  and  $B_{nS+v} = \sum_{j=0}^{S-1} \prod_{i=0}^{j-1} A_{nS+v-i} B_{nS+v-j}$ , so  $\{(A_{nS+v}, B_{nS+v}), n \in \mathbb{Z}\}$  is iid for all  $v \in \{0, \dots, S-1\}$ . The top Lyapunov exponent  $\gamma_v^{(S)}$  associated with (20) is given for all  $v \in \{0, \dots, S-1\}$  by (Bougerol and Picard, 1992)

$$\begin{aligned} \gamma_v^{(S)} &= \inf \left\{ \frac{1}{n} E \log \|A_{nS+v} A_{(n-1)S+v} \dots A_{S+v}\|, n \geq 1 \right\} \\ &= \inf \left\{ \frac{1}{n} E \log \|A_{nS+v} A_{nS+v-1} \dots A_{v+1}\|, n \geq 1 \right\}, \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_{nS+v} A_{nS+v-1} \dots A_{v+1}\| \text{ a.s.} \end{aligned} \quad (21)$$

Since  $E \log |\eta_v|^{\delta_v} < \infty$  for all  $0 \leq v \leq S-1$ , it follows that  $E \log^+ \|A_v\| < \infty$  and  $E \log^+ \|B_v\| < \infty$ . Therefore, by Theorem 2.5 of Bougerol and Picard (1992), equation (20) admits a unique nonanticipative strictly stationary and ergodic solution  $\{Y_{nS+v}, n \in \mathbb{Z}\}$  provided that  $\gamma_v^{(S)} < 0$ . The solution is given for all  $v \in \{0, \dots, S-1\}$  by

$$Y_{nS+v} = \sum_{j=0}^{\infty} \prod_{i=0}^{j-1} A_{(n-i)S+v} B_{(n-j)S+v}, \quad n \in \mathbb{Z}, \quad v \in \{0, \dots, S-1\}, \quad (22)$$

where the series in equality (22), which is exactly (6), converges absolutely *a.s.* This shows that  $\{Y_t, t \in \mathbb{Z}\}$  is the unique causal strictly periodically stationary and periodically ergodic solution of (3). Note finally that by a sandwiching argument, it is easy to see that for all  $v \in \{0, \dots, S-1\}$

$$\gamma_v^{(S)} = \lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_{nS+v} A_{nS+v-1} \dots A_{v+1}\| = \lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_{nS} A_{nS-1} \dots A_1\| := \gamma^{(S)}.$$

*Necessity:* Assume that model (3) admits a nonanticipative strictly periodically stationary solution  $\{Y_t, t \in \mathbb{Z}\}$ . From the non-negativity of the coefficients of  $A_t$  in (3) it follows that for all  $k > 1$ ,

$$Y_v \geq \sum_{j=0}^k \prod_{i=0}^{j-1} A_{v-i} B_{v-j}, \quad a.s.,$$

implying that the series  $\sum_{j=0}^{\infty} \prod_{i=0}^{j-1} A_{v-i} B_{v-j}$  converges *a.s.* Therefore,

$$\prod_{i=0}^{j-1} A_{v-i} B_{v-j} \rightarrow 0, \quad a.s. \text{ as } j \rightarrow \infty,$$

from which we have to show

$$\prod_{i=0}^{j-1} A_{v-i} \rightarrow 0, \quad a.s. \text{ as } j \rightarrow \infty. \quad (23)$$

This holds whenever

$$\lim_{j \rightarrow \infty} \prod_{i=0}^{j-1} A_{v-i} e_m = 0, a.s. \text{ for all } 1 \leq m \leq r, \quad (24)$$

where  $r = p + 2q - 2$  and  $(e_m)_{1 \leq m \leq r}$  is the canonical basis of  $\mathbb{R}^r$ . Since  $A_t$  has the same “sparsity” as the matrix  $A_t$  in Pan, Wang, and Tong (2008, p. 373), then (24) follows from their results using similar arguments (see also Akrouche and Bibi (2009) for the particular *P-GARCH* case).

(ii) Since  $\{A_t, t \in \mathbb{Z}\}$  is nonnegative then

$$\gamma^S(A) \geq \gamma^S(\beta) := \log \rho \left( \prod_{v=0}^{S-1} \beta_{S-v} \right). \quad (25)$$

If (3) has a strictly periodically stationary solution, then  $\gamma^S(A) < 0$ . In view of (25), it follows that  $\gamma^S(\beta) < 0$  establishing (7). ■

*Proof of Theorem 2.*

(i) The proof is similar to that of Lemma 2.3 of Berkes, Horv ath, and Kokoszka (2003). First, we have to show that if  $\gamma^S(A) < 0$  then there is  $\delta > 0$  and  $n_0$  such that

$$E \left( \|A_{n_0 S} A_{n_0 S-1} \dots A_1\|^\delta \right) < 1. \quad (26)$$

Since  $\gamma^S(A) = \inf_{n \in \mathbb{N}^*} \left\{ \frac{1}{n} E \left( \log \|A_{nS} A_{nS-1} \dots A_1\| \right) \right\}$  is strictly negative, there is a positive integer  $n_0$  such that

$$E \left( \log \|A_{n_0 S} A_{n_0 S-1} \dots A_1\| \right) < 0.$$

On the other hand, working with a multiplicative norm and by the *ipd<sub>S</sub>* property of the sequence  $\{A_t, t \in \mathbb{Z}\}$  we have

$$\begin{aligned} E \left( \|A_{n_0 S} A_{n_0 S-1} \dots A_1\| \right) &= \|E \left( A_{n_0 S} A_{n_0 S-1} \dots A_1 \right)\| \\ &= \|E \left( A_S A_{S-1} \dots A_1 \right)^{n_0}\| \\ &\leq \|E \left( A_S A_{S-1} \dots A_1 \right)\|^{n_0} < \infty. \end{aligned}$$

Let  $f(x) = E\left(\|A_{n_0 S} A_{n_0 S-1} \dots A_1\|^x\right)$ . Since  $f'(0) = E\left(\log \|A_{n_0 S} A_{n_0 S-1} \dots A_1\|\right) < 0$ ,  $f(x)$  decrease in a neighborhood of 0 and since  $f(0) = 1$ , it follows that there exists  $0 < \delta < 1$  such that (26) holds. Now from (6) we have for some  $v \in \{1, \dots, S\}$

$$\|Y_v\| \leq \sum_{k=1}^{\infty} \left\| \prod_{j=0}^{k-1} A_{v-j} \right\| \|B_{v-k}\| + \|B_v\|.$$

Since  $0 < \kappa < 1$ , then

$$\|Y_v\|^\kappa \leq \sum_{k=1}^{\infty} \left\| \prod_{j=0}^{k-1} A_{v-j} \right\|^\kappa \|B_{v-k}\|^\kappa + \|B_v\|^\kappa,$$

which, by the independence of  $A_{v-j}$  and  $B_{v-k}$  for  $j < k$ , implies that

$$\begin{aligned} E \|Y_v\|^\kappa &\leq \sum_{k=1}^{\infty} E \left( \left\| \prod_{j=0}^{k-1} A_{v-j} \right\|^\kappa \right) E (\|B_{v-k}\|^\kappa) + E (\|B_v\|^\kappa) \\ &\leq B(\kappa) \sum_{k=1}^{\infty} E \left( \left\| \prod_{j=0}^{k-1} A_{v-j} \right\|^\kappa \right) + E (\|B_v\|^\kappa), \end{aligned}$$

where  $B(\kappa) = \max_{0 \leq v \leq S-1} E (\|B_{v-k}\|^\kappa)$ . In view of (26) there exist  $a_v > 0$  and  $0 < b_v < 1$  such that

$$E \left( \left\| \prod_{j=0}^{k-1} A_{v-j} \right\|^\kappa \right) \leq a_v b_v^k \leq ab^k,$$

where  $ab^k = \max_{0 \leq v \leq S-1} \{a_v b_v^k\}$ . This proves that  $E \|Y_v\|^\kappa < \infty$ , showing (8).

(ii) Define  $\{\tilde{Y}_t, t \in \mathbb{Z}\}$  by

$$\begin{cases} \tilde{Y}_t = A_t \tilde{Y}_{t-1} + B_t & t \geq 1 \\ \tilde{Y}_t = 0 & t \leq 0, \end{cases} \tag{27}$$

and let  $Y^{(v)}$  ( $0 \leq v \leq S-1$ ) be a random variable having the same distribution as the term  $Y_{nS+v}$  of the unique periodically stationary solution given by (22). It is clear that  $\tilde{Y}_{nS+v} \xrightarrow{\mathcal{L}} Y^{(v)}$  as  $n \rightarrow \infty$ . Let  $m = 2$ . From the weak convergence theory (Billingsley 1968), to show that  $E(\text{vec}(Y^{(v)} Y^{(v)'}))$  is finite for all  $v$ , it is sufficient to show that  $\liminf_{n \rightarrow \infty} E(\text{vec}(\tilde{Y}'_{nS+v} \tilde{Y}_{nS+v})) < \infty$  for all  $v$ . Set  $V_{nS+v} = E(\text{vec}(\tilde{Y}'_{nS+v} \tilde{Y}_{nS+v}))$ . From (27) we get the following first-order  $S$ -periodic difference equation

$$V_{nS+v} = E(A_v^{\otimes 2}) V_{nS+v-1} + [E(A_v \otimes B_v) + E(B_v \otimes A_v)] E(\tilde{Y}_{nS+v}) + \text{vec}(E(B_v B_v')), \tag{28}$$

where  $E(A_t^{\otimes 2})$ ,  $E(A_t \otimes B_t)$  and  $\text{vec}(E(B_t B_t'))$  are finite  $S$ -periodic matrices in  $t$ . Since, the last two terms of the right-hand side of (27) are bounded, it follows that  $\lim_{n \rightarrow \infty} V_{nS+v}$  exists for every  $1 \leq v \leq S$  whenever (9) holds, which completes the proof for  $m = 2$ . For general  $m$ , the proof is similar. ■

*Proof of Theorem 3.*

The proof is very similar to that of Corollary 3.5 of Basrak, Davis, and Mikosch (2002). ■

### MCMC diagnostic tools

In order to assess the convergence of the proposed algorithm, we use some MCMC diagnostic tools, such as the autocorrelation of posterior draws, the Relative Numerical Inefficiency (RNI, Geweke 1989) and the Numerical Standard Error (NSE, Geweke 1989). The autocorrelations of parameter draws indicate how the posterior draws

mix. The *RNI* measures the degree of the inefficiency due to the serial correlation of the *MCMC* draws. It is given by

$$RNI = 1 + 2 \sum_{h=1}^B K\left(\frac{h}{B}\right) \hat{\rho}_h,$$

where  $B = 500$  is the bandwidth,  $K(\cdot)$  is the Parzen kernel (e.g. Kim, Shephard, and Chib 1998) and  $\hat{\rho}_h$  is the sample autocorrelation for the lag  $h$  of the parameter draws.

The *NSE* is the square-root of the estimated asymptotic variance of the *MCMC* estimator. It is given by

$$NSE = \sqrt{\frac{1}{L} \left( \hat{\gamma}_0 + 2 \sum_{h=1}^B K\left(\frac{h}{B}\right) \hat{\gamma}_h \right)},$$

where  $\hat{\gamma}_h$  is the sample autocovariance at lag  $h$  of the parameter draws.

## References

- Akrouche, A. 2017. "Periodic Autoregressive Stochastic Volatility." *Statistical Inference for Stochastic Processes* 20: 139–177.
- Akrouche, A., and E. Al-Eid. 2012. "Asymptotic Inference of Unstable Periodic ARCH Processes." *Statistical Inference for Stochastic Processes* 15: 61–79.
- Akrouche, A., and A. Bibi. 2009. "Quasi-Maximum Likelihood Estimation of Periodic GARCH and Periodic ARMA-GARCH Processes." *Journal of Time Series Analysis* 30: 19–46.
- Akrouche, A., and N. Touche. 2015. "Weighted Least Squares-Based Inference for Stable and Unstable Threshold Power ARCH Processes." *Statistics & Probability Letters* 97: 108–115.
- Akrouche, A., E. Al-Eid, and N. Demouche. 2018. "Generalized Quasi-Maximum Likelihood Inference for Periodic Conditionally Heteroskedastic Models." *Statistical Inference for Stochastic Processes* 21: 485–511.
- Ambach, D., and C. Croonenbroeck. 2015. "Obtaining Superior Wind Power Predictions from a Periodic and Heteroskedastic Wind Power Prediction Tool." In *Stochastic Models, Statistics and Their Applications*, Springer Proceedings in Mathematics & Statistics Vol. 122, 225–232. Cham: Springer.
- Ambach, D., and W. Schmid. 2015. "Periodic and Long Range Dependent Models for High Frequency Wind Speed Data." *Energy* 82: 277–293.
- Ardia, D. 2008. "Bayesian Estimation of a Markov-Switching Threshold Asymmetric GARCH Model with Student-t Innovations." *The Econometrics Journal* 12: 105–126.
- Basrak, B., R. A. Davis, and T. Mikosch. 2002. "Regular Variation of GARCH Processes." *Stochastic Processes and Their Applications* 99: 95–115.
- Bauwens, L., and M. Lubrano. 1998. "Bayesian Inference on GARCH Models Using Gibbs Sampler." *Journal of Econometrics* 1: 23–46.
- Bauwens, L., A. Dufays, and J. V. K. Rombouts. 2014. "Marginal Likelihood for Markov-Switching and Change-Point GARCH Models." *Journal of Econometrics* 178: 508–522.
- Berkes, I., L. Horváth, and P. Kokoszka. 2003. "GARCH Processes: Structure and Estimation." *Bernoulli* 9: 201–227.
- Billingsley, P. 1968. *Probability and Measure*. New York: Wiley.
- Bollerslev, T. 1986. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics* 31: 307–327.
- Bollerslev, T., and E. Ghysels. 1996. "Periodic Autoregressive Conditional Heteroskedasticity." *Journal of Business & Economic Statistics* 14: 139–152.
- Bollerslev, T., J. Cai, and F. M. Song. 2000. "Intraday Periodicity, Long Memory Volatility, and Macroeconomic Announcement Effects in the US Treasury Bond Market." *Journal of Empirical Finance* 7: 37–55.
- Bougerol, P., and N. Picard. 1992. "Stationarity of GARCH Processes and some Nonnegative Time Series." *Journal of Econometrics* 52: 115–127.
- Chan, J. C. C., and A. L. Grant. 2016. "On the Observed-Data Deviance Information Criterion for Volatility Modeling." *Journal of Financial Econometrics* 14: 772–802.
- Chen, C. W. S., and M. K. P. So. 2006. "On a Threshold Heteroscedastic Model." *International Journal of Forecasting* 22: 73–89.
- Ding, Z., C. W. J. Granger, and R. F. Engle. 1993. "A Long Memory Property of Stock Market Returns and a New Model." *Journal of Empirical Finance* 1: 83–106.
- Engle, R. F. 1982. "Autoregressive Conditional Heteroskedasticity with Estimates of Variance of U.K. Inflation." *Econometrica* 50: 987–1008.
- Francq, C., and J. M. Zakoian. 2008. "Deriving the Autocovariances of Powers of Markov-Switching GARCH Models, with Applications to Statistical Inference." *Computational Statistics & Data Analysis* 52: 3027–3046.
- Francq, C., and J. M. Zakoian. 2013. "Optimal Predictions of Powers of Conditionally Heteroskedastic Processes." *Journal of Royal Statistical Society B75*: 345–367.
- Francq, C., and J. M. Zakoian. 2019. *GARCH Models: Structure, Statistical Inference and Financial Applications*. 2nd ed. Hoboken, NJ: John Wiley.
- Franses, P. H., and R. Paap. 2000. "Modeling Day-of-the-Week Seasonality in the S&P 500 Index." *Applied Financial Economics* 10: 483–488.
- Geweke, J. 1989. "Bayesian Inference in Econometric Models Using Monte Carlo Integration." *Econometrica* 57: 1317–1339.
- Granger, C. W. J. 2005. "The Past and Future of Empirical Finance: Some Personal Comments." *Journal of Econometrics* 129: 35–40.

- Haas, M. 2009. "Persistence in Volatility, Conditional Kurtosis, and the Taylor Property in Absolute Value GARCH Processes." *Statistics & Probability Letters* 79: 1674–1683.
- Haas, M., S. Mittnik, and M. S. Paolella. 2004. "A New Approach to Markov Switching GARCH Models." *Journal of Financial Econometrics* 4: 493–530.
- Hamadeh, T., and J. M. Zakoian. 2011. "Asymptotic Properties of LS and QML Estimators for a Class of Nonlinear GARCH Processes." *Journal of Statistical Planning and Inference* 141: 488–507.
- Hoogerheide, L., and H. K. van Dijk. 2010. "Bayesian Forecasting of Value at Risk and Expected Shortfall Using Adaptive Importance Sampling." *International Journal of Forecasting* 26: 231–247.
- Hwang, S. Y., and I. V. Basawa. 2004. "Stationarity and Moment Structure for Box-Cox Transformed Threshold GARCH(1, 1) Processes." *Statistics and Probability Letters* 68: 209–220.
- Kim, S., N. Shephard, and S. Chib. 1998. "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models." *The Review of Economic Studies* 65: 361–393.
- Osborn, D. R., C. S. Savva, and L. Gill. 2008. "Periodic Dynamic Conditional Correlations between Stock Markets in Europe and the US." *Journal of Financial Econometrics* 6: 307–325.
- Pan, J., H. Wang, and H. Tong. 2008. "Estimation and Tests for Power-Transformed and Threshold GARCH Models." *Journal of Econometrics* 142: 352–378.
- Regnard, N., and J. M. Zakoian. 2011. "A Conditionally Heteroskedastic Model with Time-Varying Coefficients for Daily Gas Spot Prices." *Energy Economics* 33: 1240–1251.
- Ritter, C., and M. A. Tanner. 1992. "Facilitating the Gibbs Sampler: The Gibbs Stopper and the Griddy-Gibbs Sampler." *Journal of the American Statistical Association* 87: 861–870.
- Rossi, E., and D. Fantazani. 2015. "Long Memory and Periodicity in Intraday Volatility." *Journal of Financial Econometrics* 13: 922–961.
- Smith, M. S. 2010. "Bayesian Inference for a Periodic Stochastic Volatility Model of Intraday Electricity Prices." In: Kneib T., Tutz G. (eds) *Statistical Modelling and Regression Structures*. Heidelberg, 353–376. Berlin: Physica-Verlag HD.
- Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde. 2002. "Bayesian Measures of Model Complexity and Fit." *Journal of Royal Statistical Society B* 64: 583–639.
- Tsay, R. S. 2010. *Analysis of Financial Time Series: Financial Econometrics*, 3rd ed. New York: Wiley.
- Tsiakas, I. 2006. "Periodic Stochastic Volatility and Fat Tails." *Journal of Financial Econometrics* 4: 90–135.
- Xia, Q., H. Wong, J. Liu, and R. Liang. 2017. "Bayesian Analysis of Power-Transformed and Threshold GARCH Models: A Griddy-Gibbs Sampler Approach." *Computational Economics* 50: 353–372.
- Ziel, F., R. Steinert, and S. Husmann. 2015. "Efficient Modeling and Forecasting of Electricity Spot Prices." *Energy Economics* 47: 98–111.
- Ziel, F., C. Croonenbroeck, and D. Ambach. 2016. "Forecasting wind Power Modeling Periodic and Non-Linear Effects Under Conditional Heteroscedasticity." *Applied Energy* 177: 285–297.

**Supplementary Material:** The online version of this article offers supplementary material (DOI: <https://doi.org/10.1515/snde-2018-0112>).