Seasonal FIEGARCH Processes

Sílvia R.C. Lopes and Taiane S. Prass[*]

Mathematics Institute
Federal University of Rio Grande do Sul
Porto Alegre - RS - Brazil

May 25, 2013

Abstract

Here we develop the theory of seasonal FIEGARCH processes, denoted by SFIEGARCH, establishing conditions for the existence, the invertibility, the stationarity and the ergodicity of these processes. We analyze their asymptotic dependence structure by means of the autocovariance and autocorrelation functions. We also present some properties regarding their spectral representation. All properties are illustrated through graphical examples and an application of SFIEGARCH models to describe the volatility of the S&P500 US stock index log-return time series in the period from December 13, 2004 to October 02, 2009 is provided.

Keywords. Long-Range Dependence, Volatility, Periodicity, FIEGARCH Process.

Mathematics Subject Classification (2010). 60G10, 62M10, 62M15, 91B84, 97M30.

Introduction

Introduced by Bollerslev and Mikkelsen (1996), FIEGARCH processes are one of the main models used to describe the volatility in financial time series. This class of models has not only the capability of capturing the asymmetry in the log-returns, as in the EGARCH models, but also it takes into account the characteristic of long memory in the volatility, as in the FIGARCH models, with the advantage of been weakly stationary. Lopes and Prass (2013) present a study on the theoretical properties of these processes, including results on the volatility forecast. The authors also analyze the finite sample performance of the quasi-likelihood estimator for four different FIEGARCH models and present the analysis of an observed time series. The simulated study presented by Lopes and Prass (2013) considers the same parameters values as the ones in the models adjusted to the observed time series considered in Prass and Lopes (2012, 2013).

More recently, economists have noticed that FIEGARCH models are not fully satisfactory, specially when modelling volatility of intra-daily financial returns. The main discovery is that volatility of high frequency financial time series shows long-range dependence merged with periodic behavior. According to Bordignon et al. (2007), these patterns, in the case of exchange rate returns, are generally attributed to different openings of European, Asian and North American markets superimposed each other. Similar patterns are found in stock markets, mainly due to the so-called time-of-day phenomena, such as market opening, closing operations, lunch-hour and overlapping effects. Once again, the focus is on the squared, log-squared and absolute returns. Periodic components are represented as marked peaks at some frequencies of the time series periodogram function and it can also

---

[*]Corresponding author. E-mail: taianeprass@gmail.com

be identified through a persistent cyclical behavior on the autocorrelation function with oscillations decaying very slowly. From the theoretical point of view, modelling and prediction of the volatility dynamics may be seriously affected if this empirical evidence is neglected.

Bordignon et al. (2007, 2009) introduced new GARCH-type models characterized by long memory behavior of periodic type. The generalized long memory GARCH (G-GARCH) introduces generalized periodic long-memory filters, based on Gegenbauer polynomials, into the equation describing the time-varying volatility of standard GARCH models. The periodic long-memory GARCH (PLM-GARCH) process represents a natural extension of the FIGARCH model proposed for modelling the volatility long-range persistence. Although periodic long memory versions of EGARCH (PLM-EGARCH) models were also considered in Bordignon et al. (2009) , we feel that there are several theoretical results related to these processes that were not yet explored. For instance, conditions for the existence, stationarity and ergodicity are yet to be established. Moreover, the autocovariance structure and the spectral representation of these processes are of extreme importance in both theoretical and practical point of view and hence, their study is an important matter.

Here we develop the theory of seasonal FIEGARCH processes, denoted by $\text{SFIEGARCH}(p, d, q)_s$, where p, d and q have the same meaning as in the so-called $\text{FIGARCH}(p, d, q)$ process and s is the length of the periodic component. This model is similar to the PLM-EGARCH process introduced by Bordignon et al. (2009) but, in the definition considered here, for any SFIEGARCH process $\{X_t\}_{t \in \mathbb{Z}}$, the process $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$ is a SARFIMA one, where σ_t^2 is the conditional variance of X_t, for all $t \in \mathbb{Z}$. In particular, if $s = 1$, it is an $\text{ARFIMA}(q, d, p)$ process (see Lopes, 2008). This result is useful for establishing whether the process $\{X_t\}_{t \in \mathbb{Z}}$ is well defined.

Results regarding the process $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$ are already known in the literature and can be found in Bisognin and Lopes (2009) and references therein. Moreover, for an SFIEGARCH process the sequence of random variables $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$ is not directly observable and we study its characteristics only to obtain the properties of the processes $\{X_t\}_{t \in \mathbb{Z}}$ and $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$, which are the observable ones. In this work we extend the range of the parameter d for the invertibility and we present an alternative asymptotic expression for the autocovariance function $\gamma_{\ln(\sigma_t^2)}(\cdot)$. These results are useful to derive the exact and the asymptotic expressions for the autocovariance and spectral density functions of the process $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$.

The paper is organized as follows: in Section 1 we present the $\text{SFIEGARCH}(p, d, q)_s$ processes and we discuss the existence of a power series representation for the function $\lambda(z) = \frac{\alpha(z)}{\beta(z)}(1 - z^s)^{-d}$ and the asymptotic behavior of the coefficients in this representation. A recurrence formula to calculate those coefficients is also provided. In Section 1 we also analyze the existence of the process $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$ and its invertibility property. This analysis is important to guarantee the existence of the process $\{X_t\}_{t \in \mathbb{Z}}$ itself. Section 2 is devoted to study the asymptotic dependence structure of both $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$ and $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$ processes, where $\{X_t\}_{t \in \mathbb{Z}}$ is an SFIEGARCH process. Section 3 presents the spectral representation of both processes $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$ and $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$. Section 5 shows an application of SFIEGARCH models to describe the volatility of the S&P500 US stock index log-return time series in the period from December 13, 2004 to October 02, 2009. Section 6 presents the final conclusions. All proofs are presented in Appendix A.

1 SFIEGARCH Process

In this section we define the *Seasonal* FIEGARCH (SFIEGARCH) process which describes the volatility varying in time, volatility clusters (known as ARCH/GARCH effects), volatility periodic long-memory and asymmetry. Since the existence of a solution $\{X_t\}_{t \in \mathbb{Z}}$ for expression (1.1) depends on the existence of the stochastic process $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$ satisfying expression (1.2), we show that the random variable $\ln(\sigma_t^2) - \omega$ is finite with probability one, for all $t \in \mathbb{Z}$, if and only if $d < 0.5$. We show that $\{\ln(\sigma_t^2) - \omega\}_{t \in \mathbb{Z}}$ is an invertible process, with respect to $\{g(Z_t)\}_{t \in \mathbb{Z}}$, if and only if, $d \in (-1, 0.5)$, extending the range given in Bisognin and Lopes (2009). We also discuss the similarities between this model and the PLM-EGARCH model, introduced by Bordignon et al. (2009).

Hereafter, $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denote, respectively, the floor and ceiling functions and $\mathbb{I}_A(\cdot)$ is the indicator function defined as $\mathbb{I}_A(z) = 1$, if $z \in A$, and 0, otherwise. Whenever $T = \mathbb{N}$ or $T = \mathbb{Z}$, we define $T^* := T \backslash \{0\}$. Throughout the paper, given two real/complex valued functions $f(\cdot)$ and $g(\cdot)$, $f(x) = O(g(x))$, means that $|f(x)| \leq c|g(x)|$, for some $c > 0$, as $x \to \infty$; $f(x) = o(g(x))$ means that $f(x)/g(x) \to 0$, as $x \to \infty$; $f(x) \sim g(x)$ means that $f(x)/g(x) \to 1$, as $x \to \infty$. We also say that $f(x) \approx g(x)$, as $x \to \infty$, if for any $\varepsilon > 0$, there exists $x_0 \in \mathbb{R}$ such that $|f(x) - g(x)| < \varepsilon$, for all $x \geq x_0$. Similar definitions can be obtained upon replacing the functions $f(\cdot)$ and $g(\cdot)$ by sequences of real numbers $\{a_k\}_{k \in \mathbb{N}}$ and $\{b_k\}_{k \in \mathbb{N}}$ or if one considers any constant a or $-\infty$ instead of ∞.

Definition 1.1. Let $\{X_t\}_{t \in \mathbb{Z}}$ be the stochastic process defined by the expressions

$$X_t = \sigma_t Z_t, \tag{1.1}$$

$$\ln(\sigma_t^2) = \omega + \frac{\alpha(\mathcal{B})}{\beta(\mathcal{B})}(1 - \mathcal{B}^s)^{-d} g(Z_{t-1}), \quad \text{for all } t \in \mathbb{Z}, \tag{1.2}$$

where $\{Z_t\}_{t \in \mathbb{Z}}$ is a sequence of i.i.d. random variables, with zero mean and variance equal to one, $g(\cdot)$ is defined by

$$g(Z_t) = \theta Z_t + \gamma \big[|Z_t| - \mathbb{E}\big(|Z_t|\big)\big], \quad \text{with } \theta, \gamma \in \mathbb{R}, \quad \text{for all } t \in \mathbb{Z}, \tag{1.3}$$

$\omega \in \mathbb{R}$, $\mathcal{B}$ is the backward shift operator defined by $\mathcal{B}^{sk}(X_t) = X_{t-sk}$, for all $s, k \in \mathbb{N}$, $\alpha(\cdot)$ and $\beta(\cdot)$ are, respectively, polynomials of order p and q, with no common roots defined by

$$\alpha(z) = \sum_{i=0}^{p} (-\alpha_i) z^i \quad \text{and} \quad \beta(z) = \sum_{j=0}^{q} (-\beta_j) z^j, \tag{1.4}$$

with $\alpha_0 = -1 = \beta_0$, and $\beta(z) \neq 0$ in the closed disk $\{z : |z| \leq 1\}$, $d \in \mathbb{R}$ is the differencing parameter, $s \in \mathbb{N}^*$ is the length of the periodic component, $(1 - \mathcal{B}^s)^{-d}$ is the seasonal difference operator, defined by its Maclaurin series expansion, namely,

$$(1 - \mathcal{B}^s)^{-d} = \sum_{k=0}^{\infty} \frac{\Gamma(k+d)}{\Gamma(k+1)\Gamma(d)} (\mathcal{B}^s)^k := \sum_{k=0}^{\infty} \delta_{-d,k} \mathcal{B}^{sk} := \sum_{k=0}^{\infty} \pi_{d,k} \mathcal{B}^k, \tag{1.5}$$

where $\Gamma(\cdot)$ is the Gamma function, $\pi_{d,k} := 0$, if $k/s \notin \mathbb{N}$, and $\pi_{d,sj} = \delta_{-d,j} := \frac{\Gamma(k+d)}{\Gamma(k+1)\Gamma(d)}$, for all $j \in \mathbb{N}$. Then, $\{X_t\}_{t \in \mathbb{Z}}$ is a seasonal FIEGARCH process, with seasonal period s and differencing parameter d, denoted by SFIEGARCH$(p, d, q)_s$.

Remark 1.1. The assumptions that $\beta(z) \neq 0$ in the closed disk $\{z : |z| \leq 1\}$ and that $\alpha(\cdot)$ and $\beta(\cdot)$ have no common roots guarantee that the operator $\frac{\alpha(\mathcal{B})}{\beta(\mathcal{B})}$ is well defined.

Example 1.1. Figure 1.1 presents a simulated SFIEGARCH$(0, d, 0)_s$ time series $\{X_t\}_{t=1}^n$ and its conditional standard deviation $\{\sigma_t\}_{t=1}^n$, defined by expressions (1.1) and (1.2). For these graphs, $Z_0 \sim \mathcal{N}(0, 1)$, $\omega = 5.0$, $\theta = -0.25$, $\gamma = 0.24$, $d = 0.35$ and $s = 6$.

Remark 1.2. In this work we consider the case where the conditional variance σ_t^2 is defined through expression (1.2), with $\mathbb{E}(Z_0) = 0$, $\text{Var}(Z_0) = \mathbb{E}(Z_0^2) = 1$ and the function $g(\cdot)$ defined by expression (1.3). However the results presented here can be easily extended if one considers $\text{Var}(Z_0) = \sigma^2 \neq 1$ and replaces $g(\cdot)$ by any measurable function satisfying $\text{Var}(g(Z_0)) < \infty$.

Observe that the series expansion of the operator $(1 - \mathcal{B}^s)^d$ is obtained upon replacing $-d$ by d in expression (1.5). Moreover, when $d \in \mathbb{N}$, $(1 - \mathcal{B}^s)^d$ is merely the seasonal difference operator $1 - B^s$ iterated d times. Thus, one can easily see that an equivalent definition for SFIEGARCH process is given if one replaces expression (1.2) by

$$\beta(\mathcal{B})(1 - \mathcal{B}^s)^d(\ln(\sigma_t^2) - \omega) = \alpha(\mathcal{B})g(Z_{t-1}), \quad \text{for all } t \in \mathbb{Z}. \tag{1.6}$$
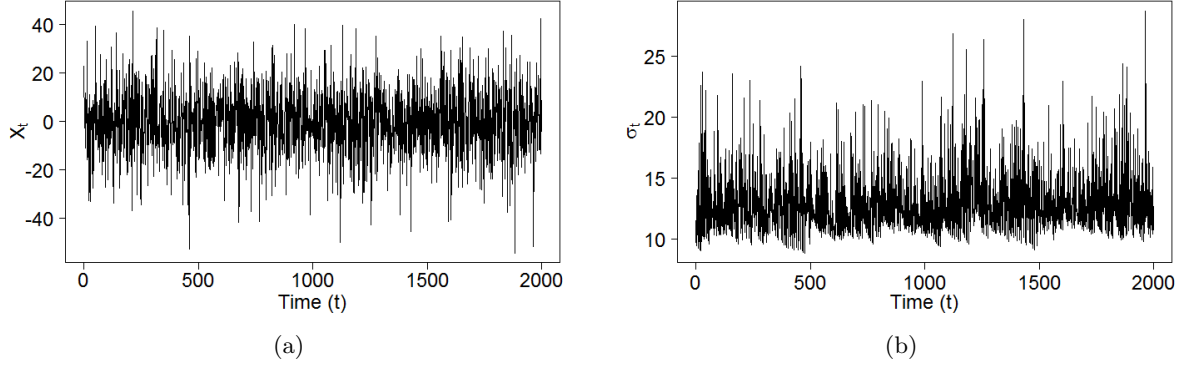
Figure 1.1: Samples from an SFIEGARCH$(0, d, 0)_s$ processes, with $n = 2000$ observations, considering $Z_0 \sim \mathcal{N}(0, 1)$, $\omega = 5.0$, $\theta = -0.25$, $\gamma = 0.24$, $d = 0.35$ and $s = 6$. Panel (a) shows the time series $\{X_t\}_{t=1}^n$. Panel (b) presents the time series $\{\sigma_t\}_{t=1}^n$, where σ_t is the conditional standard deviation of X_t, for all $t \in \{1, \cdots, n\}$.

This expression is similar to the one in the definition of the PLM-EGARCH process, presented by Bordignon et al. (2009). For a PLM-EGARCH(p, m, d, q, s), the conditional variance σ_t^2 of X_t is defined through the equation

$$(1 - \mathcal{B}^s)^d \phi(\mathcal{B})\big(\ln(\sigma_t^2) - \omega\big) = a(\mathcal{B})Z_t + c(\mathcal{B})\big(|Z_t| - \mathbb{E}(|Z_t|)\big), \quad \text{for all } t \in \mathbb{Z}, \tag{1.7}$$

where $a(z) = \sum_{k=1}^p a_k z^k$ and $c(z) = \sum_{l=1}^m c_l z^l$ are polynomials of order p and m, respectively, $\phi(z) = \sum_{j=0}^{q-s} \phi_j z^j$ is a polynomial of order $q-s$, which satisfies $(1 - \mathcal{B}^s)^d \phi(\mathcal{B}) = 1 - b(\mathcal{B})$, where $b(z) = \sum_{i=1}^q b_i z^i$ is a polynomial of order q.

Notice that in the PLM-EGARCH, the polynomials $a(\cdot)$ and $c(\cdot)$ do not necessarily have the same order. Also, it is easy to see that, by setting

$$\begin{aligned}
\phi_i &:= -\beta_i, \quad \text{for all } i = 0, \cdots, q, \\
a_{j+1} &:= -\theta\alpha_j \quad \text{and} \quad c_{j+1} := -\gamma\alpha_j, \quad \text{for all } j = 0, \cdots, p,
\end{aligned} \tag{1.8}$$

one can rewrite the right hand side of expression (1.6) as the right hand side of (1.7). Under this point of view, the PLM-EGARCH model seems more general than the SFIEGARCH one. On the other hand, the left hand side of expression (1.6) is more general then the left hand side of (1.7). This is so because in the SFIEGARCH model, no restriction is made in the order of the product $\beta(z)(1 - z^s)^d$, allowing for the parameter d to be fractional.

Remark 1.3. It is immediate that, if $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$ is a stationary process with finite mean, then $\omega = \mathbb{E}(\ln(\sigma_t^2))$, for all $t \in \mathbb{Z}$. Also, if $d = 0$, we have the EGARCH(p, q) model proposed by Nelson (1991) and, if $s = 1$, we have the FIEGARCH(p, d, q) process defined by Bollerslev and Mikkelsen (1996). A study on the theoretical properties of FIEGARCH(p, d, q) process are presented in Lopes and Prass (2013).

From Definition 1.1, one easily concludes that the existence of the stochastic process $\{X_t\}_{t \in \mathbb{Z}}$ depends on the existence of the stochastic process $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$ which satisfies equation (1.2). The existence of a solution for equation (1.2) is discussed in the sequel.

From now on, let $\lambda(\cdot)$ be the polynomial defined as

$$\lambda(z) := \frac{\alpha(z)}{\beta(z)}(1 - z^s)^{-d} = \sum_{k=0}^\infty \lambda_{d,k} z^k, \quad |z| < 1, \tag{1.9}$$

where $\alpha(\cdot)$ and $\beta(\cdot)$ are defined in (1.4). Notice that, by definition, $\beta(z)$ has no roots in the closed disk $\{z : |z| \le 1\}$, and $\alpha(\cdot)$ and $\beta(\cdot)$ have no common roots. Therefore, the function $\lambda(z)$ is analytic in

the open disc $\{z : |z| < 1\}$ and, if $d \leq 0$, in the closed disk $\{z : |z| \leq 1\}$. So, it has a unique power series representation and the operator given in (1.2) can be rewritten as $\frac{\alpha(\mathcal{B})}{\beta(\mathcal{B})}(1-\mathcal{B}^s)^{-d} = \sum_{k=0}^{\infty} \lambda_{d,k}\mathcal{B}^k = \lambda(\mathcal{B})$. This representation is more convenient and will be used from now on. In the following, we analyze the asymptotic behavior of the coefficients $\lambda_{d,k}$, for all $k \in \mathbb{Z}$, defined by expression (1.9). This result is fundamental for proving the results regarding the existence, invertibility, stationarity and ergodicity of SFIEGARCH processes.

It is immediate that, if $p \geq 0$ and $q = 0$, one can rewrite (1.9) as,

$$\lambda(z) = \begin{cases} (1 - z^s)^{-d} = \sum_{k=0}^{\infty} \pi_{d,k} z^k, & \text{if } p = 0 = q; \\ \alpha(z)(1 - z^s)^{-d} = \sum_{k=0}^{\infty} \Big[\sum_{i=0}^{\min\{p,k\}} (-\alpha_i \pi_{d,k-i}) \Big] z^k, & \text{if } p > 0 \text{ and } q = 0. \end{cases}$$

Thus, for all $r \in \{0, \cdots, s-1\}$ and all $k \in \mathbb{N}$,

$$\lambda_{d,sk+r} = \pi_{d,sk+r}, \quad \text{if} \quad p = 0 = q,$$

and, whenever $p > 0$ and $q = 0$,

$$\lambda_{d,sk+r} = -\sum_{j=0}^{\min\{p,sk+r\}} \alpha_j \pi_{d,sk+r-j} = \begin{cases} 0, & \text{if } p < r; \\ -\sum_{j=0}^{\min\{\lfloor \frac{p-r}{s}\rfloor,k\}} \alpha_{sj+r} \pi_{d,sk-sj}, & \text{otherwise.} \end{cases}$$

Consequently, given $r > 0$,

$$\sum_{k=0}^{\infty} |\lambda_{d,k}|^r < \infty \quad \text{if and only if} \quad \sum_{k=0}^{\infty} |\pi_{d,k}|^r < \infty.$$

Theorem 1.1 bellow shows that this result also holds in the general case $p \geq 0$ and $q > 0$. The proofs of all results stated in this work are given in the Appendix.

Remark 1.4. By Stirling's formula and from lemma 3.1 in Kokoszka and Taqqu (1995), one easily concludes that

$$\pi_{d,sk} := \frac{\Gamma(k+d)}{\Gamma(d)\Gamma(k+1)} = \frac{1}{\Gamma(d)k^{1-d}} + O(k^{d-2}) \sim \frac{1}{\Gamma(d)k^{1-d}}, \quad \text{as } k \to \infty. \tag{1.10}$$

Since (integral convergence test)

$$\sum_{k=1}^{m} k^{-r(d-1)} \leq \int_1^m x^{-r(d-1)} \mathrm{d}x = \frac{1}{1-r(d-1)}\Big|_1^m, \quad \text{for any } r > 0,$$

converges to a finite constant as $m \to \infty$ if and only if $1 - (1-d)r < 0$, it follows that $\sum_{k=0}^{\infty} |\pi_{d,k}|^r < \infty$ if and only if $(1-d)r > 1$.

Theorem 1.1. *Let $\lambda_{d,k}$, for $k \in \mathbb{N}$, be the coefficients of the polynomial $\lambda(\cdot)$, given by expression (1.9). Let $f(z) := \frac{\alpha(z)}{\beta(z)} = \sum_{k=0}^{\infty} f_k z^k$. Then, for each $r \in \{0, \cdots, s-1\}$ and any $\nu > 0$, one has*

$$\lambda_{d,sk+r} = \pi_{d,sk} \mathscr{K}(sk+r) + o(k^{-\nu}), \quad \text{as } k \to \infty, \tag{1.11}$$

where $\mathscr{K}(\cdot)$ satisfies $\lim_{k \to \infty} \sum_{r=0}^{s-1} \mathscr{K}(sk+r) = \frac{\alpha(1)}{\beta(1)}$. Thus, $\sum_{r=0}^{s-1} \lambda_{d,sk+r} \sim \pi_{sk}\frac{\alpha(1)}{\beta(1)}$, as $k \to \infty$.

Theorem 1.2 presents an alternative asymptotic representation for the coefficients $\lambda_{d,k}$, as k goes to infinity. While expression (1.11) is more convenient for proving the asymptotic behavior of $\gamma_{\ln(X_t^2)}(\cdot)$ (see Theorem 2.3), expression (1.12) is useful for simulation purpose (see Remark 1.5).

Theorem 1.2. *Let $\lambda_{d,k}$, for $k \in \mathbb{N}$, be the coefficients of the polynomial $\lambda(\cdot)$, given by expression* (1.9), *with $d < 0.5$. Then, for each $r \in \{0, \cdots, s-1\}$, one can write*

$$\lambda_{d,sk+r} = \frac{1}{\Gamma(d)k^{1-d}} \frac{\alpha(1)}{\beta(1)} - \frac{1}{\Gamma(d)k^{1-d}} \sum_{j=0}^{sk+r} f_j \mathbb{I}_{\mathbb{R}\setminus\mathbb{N}} \left(\frac{|j-r|}{s} \right) + O(k^{d-2}) \tag{1.12}$$

$$= O(k^{d-1}) + O(k^{d-1})\mathbb{I}_{\mathbb{N}\setminus\{0,1\}}(s) + O(k^{d-2}), \qquad as \ k \to \infty,$$

where $f(z) = \frac{\alpha(z)}{\beta(z)} = \sum_{k=0}^{\infty} f_k z^k$.

Remark 1.5. From Theorem 1.1 one observes that $\lambda_{d,sk+r}$ behaves asymptotically as the coefficient $\pi_{d,sk}$, as k goes to infinity. This property is very useful to prove the results stated in Section 2.1. On the other hand, from Theorem 1.2,

$$\lambda_{d,k} \approx \frac{s^{1-d}}{\Gamma(d)k^{1-d}} \frac{\alpha(1)}{\beta(1)}, \quad as \quad k \to \infty.$$

This approximation has a closed formula which also takes into account the magnitude of $\frac{\alpha(1)}{\beta(1)}$. Although this is a rough approximation, it can be used to estimate a truncation point m for $\lambda(\cdot)$ in Monte Carlo simulation studies. That is, given $\varepsilon > 0$, if one chooses $m \gg s\left[\frac{1}{|\Gamma(d)|\varepsilon} \frac{\alpha(1)}{\beta(1)}\right]^{\frac{1}{1-d}}$, one gets $|\lambda_{d,m}| < \varepsilon$.

In the following proposition we present a recurrence formula to calculate the coefficients $\lambda_{d,k}$, for all $k \in \mathbb{N}$. This recurrence formula is very useful in Monte Carlo simulation studies.

Proposition 1.1. *Let $\lambda(\cdot)$ be the polynomial defined by* (1.9). *Suppose $\alpha(\cdot)$ and $\beta(\cdot)$ have no common roots and $\beta(z) \neq 0$ in the closed disk $\{z : |z| \leq 1\}$. Then, the coefficients $\lambda_{d,k}$, for all $k \in \mathbb{N}$, are given by*

$$\lambda_{d,k} = \begin{cases} 1, & if \quad k = 0, \\[2mm] -\alpha_k + \sum_{i=0}^{k-1} \lambda_{d,i}\left(\sum_{j=0}^{(k-i)\wedge q} \delta^*_{d,\frac{k-i-j}{s}} \beta_j \right), & if \quad k \leq p; \\[2mm] \sum_{i=0}^{k-1} \lambda_{d,i}\left(\sum_{j=0}^{(k-i)\wedge q} \delta^*_{d,\frac{k-i-j}{s}} \beta_j \right), & if \quad k > p, \end{cases}$$

where $(k-i) \wedge q = \min\{k-i, q\}$ and, by definition,

$$\delta^*_{d,m} = \begin{cases} \delta_{d,m}, & if \quad m \in \mathbb{N}, \\[2mm] 0, & if \quad m \notin \mathbb{N}, \end{cases} \tag{1.13}$$

with $\delta_{d,m}$, for all $m \in \mathbb{N}$, defined in (1.5).

The following proposition presents some properties of the stochastic process $\{g(Z_t)\}_{t\in\mathbb{Z}}$. Although the proof is straightforward and follows immediately from the fact that $\{Z_t\}_{t\in\mathbb{Z}}$ is a sequence of i.i.d. random variables, the proposition is fundamental to establish the result in Lemma 1.1, Corollary 1.1 and Theorem 1.3.

Proposition 1.2. *Let $g(\cdot)$ be the function defined by* (1.3) *and $\{Z_t\}_{t\in\mathbb{Z}}$ be a sequence of i.i.d. random variables, with zero mean and variance equal to one. Then, $\{g(Z_t)\}_{t\in\mathbb{Z}}$ is a white noise process with i.i.d. random variables and its variance σ_g^2 is given by*

$$\sigma_g^2 = \theta^2 + \gamma^2 - [\gamma\mathbb{E}(|Z_0|)]^2 + 2\theta\gamma\mathbb{E}(Z_0|Z_0|). \tag{1.14}$$

Moreover, the stochastic process $\{g(Z_t)\}_{t\in\mathbb{Z}}$ is stationary (weak and strictly) and ergodic.

Remark 1.6. Henceforth, GED shall stands for the so-called Generalized Error Distribution (see Nelson, 1991). Whenever we consider $Z_0 \sim \text{GED}(\nu)$, and ν is the tail-thickness parameter, we assume that the random variable was normalized to have mean zero and variance equal to one.

Remark 1.7. If the random variable Z_0 is symmetric, then $\mathbb{E}(Z_0|Z_0|) = 0$ and (1.14) is replaced by $\sigma_g^2 = \theta^2 + \gamma^2(1 - \mathbb{E}(|Z_0|)^2)$. Besides, if $Z_0 \sim \mathrm{GED}(\nu)$, then

$$\sigma_g^2 = \theta^2 + \gamma^2 \left(1 - \frac{\Gamma(2/\nu)}{\sqrt{\Gamma(1/\nu)\Gamma(3/\nu)}} \right), \quad \text{for any} \quad \nu > 1.$$

If $\nu = 2$, one has the Gaussian case, that is, $Z_0 \sim \mathcal{N}(0,1)$ and $\sigma_g^2 = \theta^2 + \gamma^2(1 - 2/\pi)$.

Example 1.2. Figure 1.2 (a) shows the graphs of σ_g^2 as a function of θ and γ, when $Z_0 \sim \mathcal{N}(0,1)$. Figures 1.2 (b) and (c) consider $Z_0 \sim \mathrm{GED}(\nu)$, with $\nu > 1$ ($\nu = 2$ corresponds to the Gaussian case), and present the graphs of σ_g^2, respectively, as a function of γ and ν, for $\theta = 0.25$, and as a function of θ and ν, for $\gamma = 0.24$. Notice that for these graphs, Z_0 is a symmetric random variable, so we only consider positive values of θ and γ.

From Figure 1.2 one observes that, although σ_g^2 is increasing in both θ and γ, for any $\nu > 1$, it varies faster as θ increases than when γ does (notice the scales for the ordinate axis). Moreover, for each γ and θ fixed, σ_g^2 is decreasing in ν. This is expected since $\mathbb{E}(|Z_0|)$ is increasing for $\nu \in (1, 5.00]$. This fact is illustrated in Figure 1.3 where the graphs of $\mathbb{E}(|Z_0|)$ and σ_g^2 as functions of ν are presented. In Figure 1.3 (b) we fixed $\theta = 0.25$ and $\gamma = 0.24$. From this figure it is easy to see that σ_g^2 is indeed decreasing in ν.
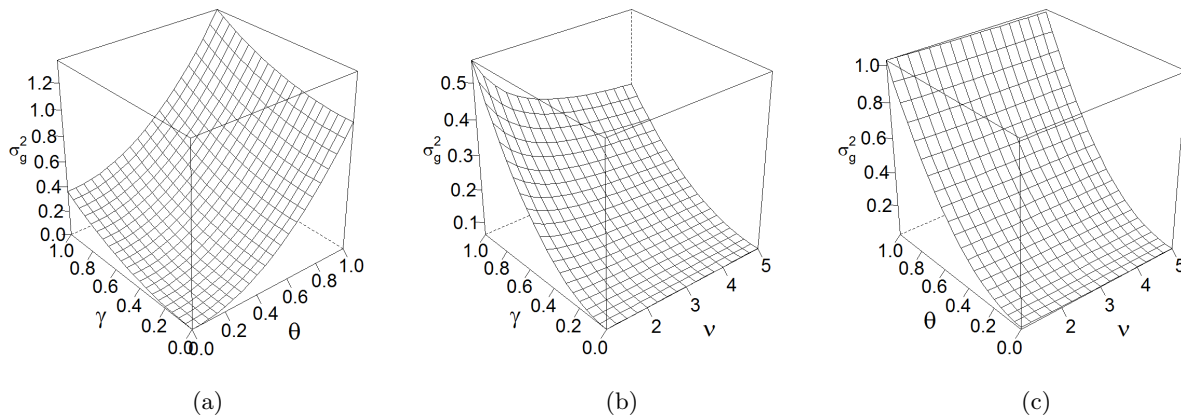

(a)   (b)   (c)

Figure 1.2: This figure presents the behavior of σ_g^2, the variance of the process $\{g(Z_t)\}_{t\in\mathbb{Z}}$, as a function of parameters θ, γ and ν, when $Z_0 \sim \mathrm{GED}(\nu)$. Panel (a) considers σ_g^2 as a function of θ and γ when $\nu = 2$, that is, when $Z_0 \sim \mathcal{N}(0,1)$. Panel (b) shows σ_g^2 as a function of γ and ν when $\theta = 0.25$. Panel (c) considers σ_g^2 as a function of θ and ν when $\gamma = 0.24$.
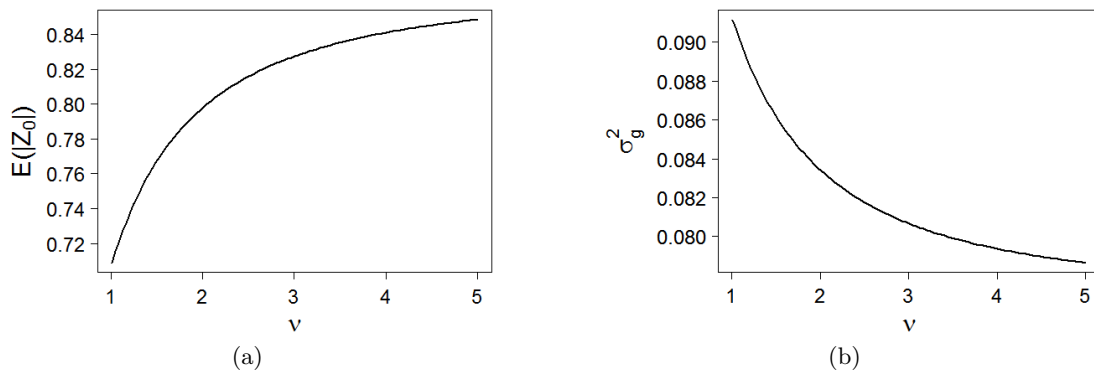

(a)   (b)

Figure 1.3: This figure considers $Z_0 \sim \mathrm{GED}(\nu)$. Panel (a) presents the graph of $\mathbb{E}(|Z_0|)$ as a function of ν. Panel (b) shows the the graph of $\mathbb{E}(g(Z_0)^2) = \sigma_g^2$, the variance of the process $\{g(Z_t)\}_{t\in\mathbb{Z}}$, as a function of ν when $\theta = 0.25$ and $\gamma = 0.24$.

Lemma 1.1 provides the necessary and sufficient conditions for the existence of the process $\{\ln(\sigma_t^2)\}_{t\in\mathbb{Z}}$.

Lemma 1.1. *Suppose that $\{Z_t\}_{t\in\mathbb{Z}}$ is a sequence of i.i.d. random variables with zero mean and variance equal to one. Let $\{g(Z_t)\}_{t\in\mathbb{Z}}$ be the process defined by* (1.3), *ω be a real constant and $\lambda(\cdot)$ be the operator defined by* (1.9). *Define*

$$\ln(\sigma_t^2) - \omega = \sum_{k=0}^{\infty} \lambda_{d,k} g(Z_{t-1-k}), \quad \text{for all } t \in \mathbb{Z}. \tag{1.15}$$

Thus, the series (1.15) *is well defined and converges a.s. if and only if $d < 0.5$. Moreover, the series* (1.15) *converges absolutely a.s. for $d \leq 0$.*

Corollaries 1.1 and 1.2 follow immediately from Lemma 1.1 and show, respectively, that the process $\{\ln(\sigma_t^2)\}_{t\in\mathbb{Z}}$ is a causal SARFIMA process and that X_t is finite with probability one, for all $t \in \mathbb{Z}$. We emphasize that, causality and invertibility are defined in terms of convergence in the linear space L^2 (see Palma, 2007) and not in the linear space L^1 (as in Brockwell and Davis, 1991). The same approach is considered in Bloomfield (1985) and Bondon and Palma (2007).

Corollary 1.1. *Let $\{\ln(\sigma_t^2)\}_{t\in\mathbb{Z}}$ be the stochastic process defined by expression* (1.2), *with $d < 0.5$. Then, $\{\ln(\sigma_t^2)\}_{t\in\mathbb{Z}}$ is a causal $\mathrm{SARFIMA}(p,0,q) \times (0,D,0)_s$ process, with $D = d$.*

Corollary 1.2. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be an $\mathrm{SFIEGARCH}(p,d,q)_s$ process, with $d < 0.5$. Then, the random variable X_t is finite with probability one, for all $t \in \mathbb{Z}$.*

Bisognin and Lopes (2009) show that a $\mathrm{SARFIMA}(p,d,q) \times (P,D,Q)_s$ process is invertible whenever $|d + D| < 0.5$. Moreover, it is usually stated that an $\mathrm{ARFIMA}(p,d,q)$ process is invertible for $|d| < 0.5$ (see for instance Hosking, 1981; Brockwell and Davis, 1991). However, Bloomfield (1985) proves that, for an $\mathrm{ARFIMA}(0,d,0)$, this range can be extended to $d \in (-1, 0.5)$. Bondon and Palma (2007) show that this result actually holds for any $\mathrm{ARFIMA}(p,d,q)$. Although the spectral density function of $\{\ln(\sigma_t^2) - \omega\}_{t\in\mathbb{Z}}$ does not satisfy all conditions imposed in Bondon and Palma (2007), with some modifications in their proof, we show here that the results still holds for a $\mathrm{SARFIMA}(p,0,q) \times (0,D,0)_s$ process (see Theorem 1.3).

Theorem 1.3. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be an $\mathrm{SFIEGARCH}(p,d,q)_s$, defined by* (1.1) *and* (1.2), *with γ and θ, not both equal to zero. Assume that $\alpha(z) \neq 0$, for all $|z| \leq 1$. Let $Y_t := \ln(\sigma_t^2) - \omega$, for all $t \in \mathbb{Z}$. Then,*

$$\lim_{m\to\infty} \mathbb{E}\left(\left|\sum_{k=0}^{m} \tilde{\lambda}_{d,k} Y_{t-k} - g(Z_{t-1})\right|^{\mathfrak{p}}\right) = 0, \quad \text{for all} \quad 0 < \mathfrak{p} \leq 2,$$

if and only if $d \in (-1, 0.5)$, where $\{\tilde{\lambda}_{d,k}\}_{k\in\mathbb{N}}$ is the sequence of coefficients in the series expansion of $\tilde{\lambda}(z) := \lambda^{-1}(z)$, for $|z| < 1$, that is,

$$\sum_{k=0}^{\infty} \tilde{\lambda}_{d,k} z^k = \tilde{\lambda}(z) := \lambda^{-1}(z) = \frac{\beta(z)}{\alpha(z)}(1 - z^s)^d, \quad |z| < 1.$$

2 Stationarity and Ergodicity

Here we show that for any $\mathrm{SFIEGARCH}(p,d,q)_s$, with θ and γ not both equal to zero and $d < 0.5$, the processes $\{X_t\}_{t\in\mathbb{Z}}$ and $\{\sigma_t^2\}_{t\in\mathbb{Z}}$ are strictly stationary and ergodic processes. We also prove that if $\mathbb{E}([\ln(Z_0^2)]^2) < \infty$, the process $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$ is well defined and it is stationary (weakly and strictly) and ergodic. Weakly stationarity of the processes $\{X_t\}_{t\in\mathbb{Z}}$ and $\{\sigma_t^2\}_{t\in\mathbb{Z}}$ is also discussed. For any stationary SFIEGARCH process, we give the expressions for the autocovariance and autocorrelation functions of $\{\ln(\sigma_t^2)\}_{t\in\mathbb{Z}}$ and $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$ and study their relation and asymptotic behavior. We also provide expression for the asymmetry (also known as skewness) and kurtosis measures for any stationary SFIEGARCH process $\{X_t\}_{t\in\mathbb{Z}}$.

Lemma 2.1 presents the conditions for the stationarity of the SARFIMA process $\{\ln(\sigma_t^2) - \omega\}_{t \in \mathbb{Z}}$. This lemma is useful to prove Theorem 2.1 that presents results on the stationarity of the processes $\{X_t\}_{t \in \mathbb{Z}}$ and $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$.

Lemma 2.1. *Let $\{\ln(\sigma_t^2) - \omega\}_{t \in \mathbb{Z}}$ be defined by (1.2), with γ and θ not both equal to zero. If $d < 0.5$, the stochastic process $\{\ln(\sigma_t^2) - \omega\}_{t \in \mathbb{Z}}$ is stationary (strictly and weakly) and ergodic.*

Corollary 2.1. *If $d < 0.5$, the stochastic process $\{\sigma_t^2\}_{t \in \mathbb{Z}}$ is strictly stationary and ergodic.*

Theorem 2.1 shows that both processes $\{X_t\}_{t \in \mathbb{Z}}$ and $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$ are strictly stationary and ergodic, whenever $d < 0.5$ and $\mathbb{E}(|\ln(Z_0)|) < \infty$, regardless the distribution of the random variable Z_0. This theorem also provides the necessary condition for $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$ to be a weakly stationary process.

Theorem 2.1. *Let $\{X_t\}_{t \in \mathbb{Z}}$ be an SFIEGARCH$(p, d, q)_s$, defined by (1.1) and (1.2). Suppose that γ and θ, given in (1.3), are not both equal to zero. If $d < 0.5$*

i) *the stochastic process $\{X_t\}_{t \in \mathbb{Z}}$ is strictly stationary and ergodic.*

ii) *if $\mathbb{E}(|\ln(Z_0^2)|) < \infty$, the process $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$ is well defined and it is strictly stationary and ergodic. Moreover, if $\mathbb{E}([\ln(Z_0^2)]^2) < \infty$ then it is also weakly stationary.*

Although both processes $\{X_t\}_{t \in \mathbb{Z}}$ and $\{\sigma_t^2\}_{t \in \mathbb{Z}}$ are strictly stationary, whenever $d < 0.5$, they are not necessarily weakly stationary. This property depends on the distribution of Z_0 and not only on the existence of the second moment for this random variable. Theorem 2.2 gives the condition for the existence of the r-th moment, for any $\mathsf{r} > 0$, for both processes $\{X_t\}_{t \in \mathbb{Z}}$ and $\{\sigma_t^2\}_{t \in \mathbb{Z}}$.

Theorem 2.2. *Let $\{X_t\}_{t \in \mathbb{Z}}$ be an SFIEGARCH$(p, d, q)_s$ process, with $d < 0.5$. Assume that θ and γ are not both equal to zero. If there exists $\mathsf{r} > 2$ such that*

$$\mathbb{E}(|Z_0|^{\mathsf{r}}) < \infty \quad and \quad \sum_{k=0}^{\infty} \left| \mathbb{E}\left(\exp\left\{ \frac{\mathsf{r}}{2} \lambda_{d,k} g(Z_0) \right\} \right) - 1 \right| < \infty, \tag{2.1}$$

then, $\mathbb{E}(|X_t|^{\mathsf{m}}) < \infty$ and $\mathbb{E}(|\sigma_t|^{\mathsf{m}}) < \infty$, for all $t \in \mathbb{Z}$ and $0 < \mathsf{m} \leq \mathsf{r}$.

Assume that $d < 0.5$ and $Z_0 \sim \mathcal{N}(0, 1)$. Let $\Phi(\cdot)$ and erf$(\cdot)$ be, respectively, the standard Gaussian distribution and the error function, that is,

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \mathrm{d}t \quad \text{and} \quad \mathrm{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} \mathrm{d}t = 2\Phi(z\sqrt{2}) - 1, \quad \text{for all } z \in \mathbb{R}.$$

It follows that,

$$\begin{aligned}
\mathbb{E}\left(\exp\left\{ \mathsf{r}\lambda_{d,k} g(Z_0) \right\} \right) &= \int_{-\infty}^{\infty} \exp\left\{ \mathsf{r}\lambda_{d,k}\left[\theta|z| + \gamma\left(|z| - \sqrt{2/\pi} \right) \right] \right\} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \mathrm{d}z \\
&= e^{-\mathsf{r}\lambda_{d,k}\gamma\sqrt{2/\pi}} \left[\exp\left\{ \frac{(\mathsf{r}\lambda_{d,k})^2(\theta - \gamma)^2}{2} \right\} \Phi\left(\mathsf{r}\lambda_{d,k}(\gamma - \theta) \right) \right. \\
&\qquad\qquad \left. + \exp\left\{ \frac{(\mathsf{r}\lambda_{d,k})^2(\theta + \gamma)^2}{2} \right\} \Phi\left(\mathsf{r}\lambda_{d,k}(\gamma + \theta) \right) \right] \\
&= \frac{1}{2} e^{-\mathsf{r}\lambda_{d,k}\gamma\sqrt{2/\pi}} \left(e^{\frac{(\mathsf{r}\lambda_{d,k})^2(\gamma - \theta)^2}{2}} \left[1 + \mathrm{erf}\left(\frac{\mathsf{r}\lambda_{d,k}(\gamma - \theta)}{\sqrt{2}} \right) \right] \right. \\
&\qquad\qquad \left. + e^{\frac{(\mathsf{r}\lambda_{d,k})^2(\gamma + \theta)^2}{2}} \left[1 + \mathrm{erf}\left(\frac{\mathsf{r}\lambda_{d,k}(\gamma + \theta)}{\sqrt{2}} \right) \right] \right), \tag{2.2}
\end{aligned}$$

for all $k \in \mathbb{N}$ and all $\mathsf{r} > 0$. Since $e^x = 1 + x + O(x^2)$ and $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}x + O(x^3)$, as $x \to 0$, one can rewrite (2.2) as,

$$\mathbb{E}\Big(\exp\Big\{\mathsf{r}\lambda_{d,k}g(Z_0)\Big\}\Big) = \frac{1}{2}\left[1 - \mathsf{r}\lambda_{d,k}\gamma\sqrt{\frac{2}{\pi}} + O(\lambda_{d,k}^2)\right]\left[2 + 2\mathsf{r}\lambda_{d,k}\gamma\sqrt{\frac{2}{\pi}} + O(\lambda_{d,k}^2)\right] = 1 + O(\lambda_{d,k}^2),$$

as $k \to \infty$. Thus, condition (2.1) holds and hence, $\mathbb{E}(|X_t|^{\mathsf{r}}) < \infty$ and $\mathbb{E}(|\sigma_t|^{\mathsf{r}}) < \infty$, for all $\mathsf{r} > 0$. Corollary 2.2 shows that this result also holds if $Z_0 \sim \mathrm{GED}(\nu)$, for any $\nu > 0$.

Corollary 2.2. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be an SFIEGARCH$(p,d,q)_s$ process, with $d < 0.5$. Assume that θ and γ are not both equal to zero. Let $\{Z_t\}_{t\in\mathbb{Z}}$ be i.i.d. GED with zero mean, variance equal to one, and tail-thickness parameter $v > 1$. Then, $\mathbb{E}(X_t^{\mathsf{r}}) < \infty$ and $\mathbb{E}([\sigma_t^2]^{\mathsf{r}}) < \infty$, for all $t \in \mathbb{Z}$ and $\mathsf{r} > 0$.*

From expression (A.1.5) in Nelson (1991), if $Z_0 \sim \mathrm{GED}(\nu)$, with $\nu > 1$, and $b_k := \frac{\mathsf{r}}{2}\lambda_{d,k}$, for all $k \in \mathbb{N}$ and any $\mathsf{r} > 0$, then

$$\mathbb{E}(e^{b_k g(Z_0)}) = \exp\left\{-b_k\gamma\lambda 2^{1/\nu}\frac{\Gamma\left(\frac{2}{\nu}\right)}{\Gamma\left(\frac{1}{\nu}\right)}\right\}\sum_{j=0}^{\infty}\left(b_k\lambda 2^{1/\nu}\right)^j\left[(\gamma+\theta)^j + (\gamma-\theta)^j\right]\frac{\Gamma\left(\frac{j+1}{\nu}\right)}{2\Gamma\left(\frac{1}{\nu}\right)\Gamma(j+1)}, \qquad (2.3)$$

where $\lambda = \left[2^{1/\nu}\Gamma(1/\nu)/\Gamma(3/\nu)\right]^{1/2}$, for all $k \in \mathbb{N}$. From expression (2.3), it is easy to see that $\mathbb{E}(X_t^{\mathsf{r}})$ is symmetric in θ, for any $\mathsf{r} > 0$ and $\nu > 1$.

Example 2.1. Figures 2.1 - 2.3 consider SFIEGARCH$(0,d,0)_s$ processes, with $s = 2$. In these figures we analyze the behavior of $\mathbb{E}(X_t^2)$, with respect to the parameters θ, γ and d. We also study the behavior of $\mathbb{E}(X_t^2)$ with respect to the parameter ν, when $Z_0 \sim \mathrm{GED}(\nu)$. From expression (2.3) one observes that $\mathbb{E}(X_t^2)$ is symmetric in θ, whenever $Z_0 \sim \mathrm{GED}(\nu)$, for any $\nu > 1$. Therefore, for these figures we only consider positive values of θ. Figure 2.1 (a) shows the behavior of $\mathbb{E}(X_t^2)$ as a function of θ and γ, for $d = 0.25$. Figure 2.1 (b) presents $\mathbb{E}(X_t^2)$ as a function of θ and d, for $\gamma = 0.24$. Figure 2.1 (c) shows $\mathbb{E}(X_t^2)$ as a function of γ and d, for $\theta = 0.25$. For all graphs in Figure 2.1, $s = 2$, $\omega = 0$ and $Z_0 \sim \mathcal{N}(0,1)$.

From Figure 2.1, one observes that for θ or γ fixed, $\mathbb{E}(X_t^2)$ slowly decreases for $d \in [-0.45, 0]$ and increases for $d \in [0, 0.45]$. This behavior can be better observed in Figures 2.2 (a) and (b) where the values of $\mathbb{E}(X_t^2)$, as a function of d, are plotted for $\theta \in \{0, 0.15, 0.30\}$ and $\gamma \in \{-0.30, 0, 0.30\}$, respectively. Similarly, for each d fixed, the function $\mathbb{E}(X_t^2)$ is decreasing for $\theta, \gamma \in [-0.3, 0.0]$ (the function is symmetric in θ) and it is increasing for $\theta, \gamma \in [0.0, 0.3]$. This behavior can be observed in Figures 2.2 (c) and (d) where $\mathbb{E}(X_t^2)$ is given, respectively, as a function of θ and γ, for $d \in \{-0.40, 0, 0.45\}$.

Example 2.2. Figure 2.3 (a) shows the graph of $\mathbb{E}(X_t^2)$, as a function of ν and d, when $Z_0 \sim \mathrm{GED}(\nu)$, with $\nu > 1$. Figures 2.3 (b) and (c) present the graph of $\mathbb{E}(X_t^2)$, respectively, as a function of ν, for $d \in \{-0.45, 0, 0.45\}$ and as a function of d, for $\nu \in \{1.01, 3, 5\}$. For all graphs, $s = 2$, $\omega = 0$, $\theta = 0.25$ and $\gamma = 0.24$. From Figure 2.3 one concludes that $\mathbb{E}(X_t^2)$ is a decreasing function of ν and, as a function of d, $\mathbb{E}(X_t^2)$ presents the same behavior as in the Gaussian case, that is, it is decreasing for $d \in [-0.45, 0]$ and increasing for $d \in [0, 0.45]$.

The following proposition presents the kurtosis and the asymmetry measures for any stationary SFIEGARCH process.

Proposition 2.1. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be a stationary SFIEGARCH$(p,d,q)_s$ process with $\mathbb{E}(|X_t^4|) < \infty$. The asymmetry and kurtosis measures of $\{X_t\}_{t\in\mathbb{Z}}$ are given, respectively, by*

$$A_X = \mathbb{E}(Z_0^3)\frac{\displaystyle\prod_{k=0}^{\infty}\mathbb{E}\Big(\exp\Big\{\frac{3}{2}\lambda_{d,k}g(Z_t)\Big\}\Big)}{\displaystyle\prod_{k=0}^{\infty}\Big[\mathbb{E}\Big(\exp\{\lambda_{d,k}g(Z_t)\}\Big)\Big]^{3/2}} \quad and \quad K_X = \mathbb{E}(Z_0^4)\frac{\displaystyle\prod_{k=0}^{\infty}\mathbb{E}\Big(\exp\{2\lambda_{d,k}g(Z_t)\}\Big)}{\displaystyle\prod_{k=0}^{\infty}\Big[\mathbb{E}\Big(\exp\{\lambda_{d,k}g(Z_t)\}\Big)\Big]^2}.$$
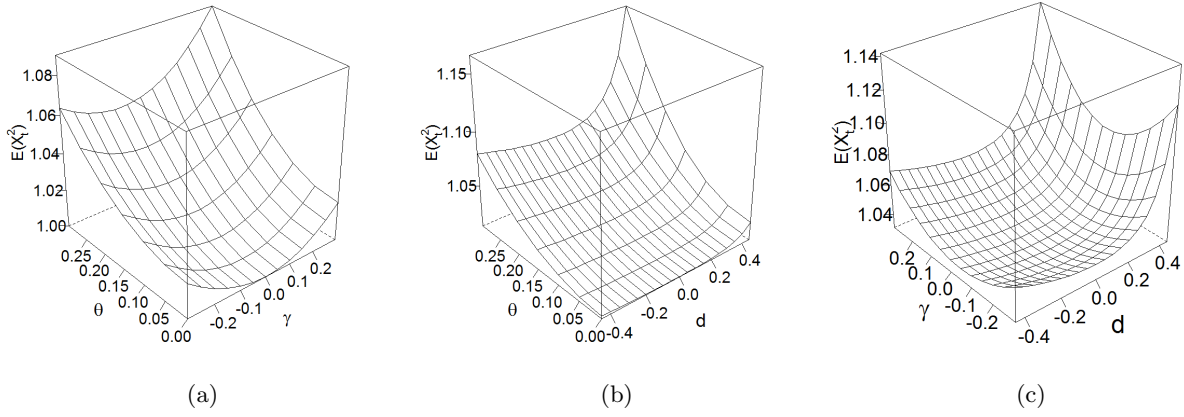
Figure 2.1: This figure illustrates the behavior of $\mathbb{E}(X_t^2)$ with respect to the parameters θ, γ and d when $\{X_t\}_{t\in\mathbb{Z}}$ is an SFIEGARCH$(0, d, 0)_s$ processes, with $s = 2$, $\omega = 0$ and $Z_0 \sim \mathcal{N}(0, 1)$. Panel (a) fixes $d = 0.25$ and shows $\mathbb{E}(X_t^2)$ as a function of θ and γ. Panel (b) fixes $\gamma = 0.24$ and presents $\mathbb{E}(X_t^2)$ as a function of θ and d. Panel (c) fixes $\theta = 0.25$ and considers $\mathbb{E}(X_t^2)$ as a function of γ and d.
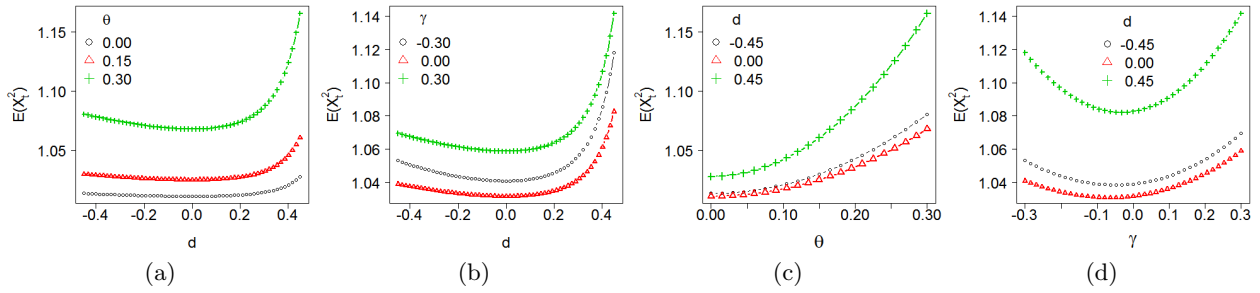
Figure 2.2: This figure shows the behavior of $\mathbb{E}(X_t^2)$ as a function of d, θ or γ, when $\{X_t\}_{t\in\mathbb{Z}}$ is an SFIEGARCH$(0, d, 0)_s$ processes with $s = 2$, $\omega = 0$ and $Z_0 \sim \mathcal{N}(0, 1)$. Panel (a) shows $\mathbb{E}(X_t^2)$ as a function of d, for $\theta \in \{0, 0.15, 0.30\}$ and $\gamma = 0.24$. Panel (b) considers $\mathbb{E}(X_t^2)$ as a function of d, for $\gamma \in \{-0.30, 0.15, 0.30\}$ and $\theta = 0.25$. Panel (c) gives $\mathbb{E}(X_t^2)$ as a function of θ, for $d \in \{-0.45, 0, 0.45\}$ and $\gamma = 0.24$. Panel (d) presents $\mathbb{E}(X_t^2)$ as a function of γ, for $d \in \{-0.45, 0, 0.45\}$ and $\theta = 0.25$.
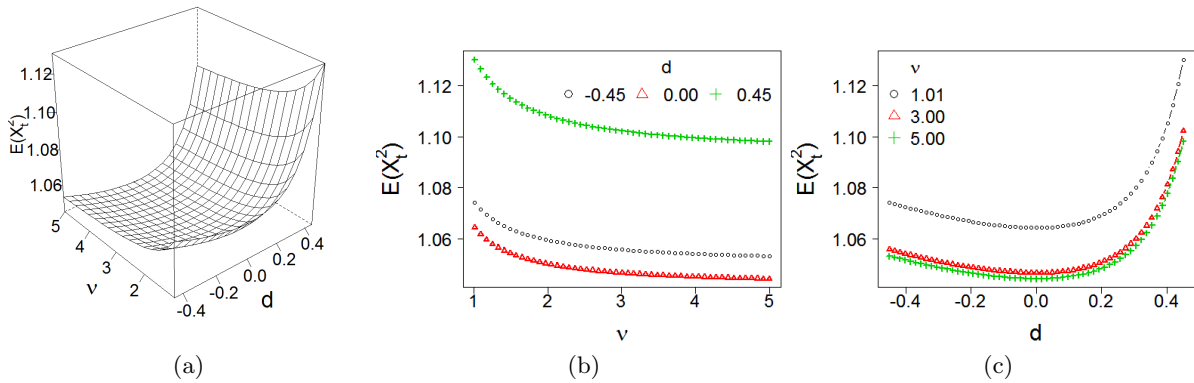

Figure 2.3: This figure illustrates the behavior of $\mathbb{E}(X_t^2)$ as a function of d and ν, when $\{X_t\}_{t\in\mathbb{Z}}$ is an SFIEGARCH$(0, d, 0)_s$ processes with $s = 2$, $\theta = 0.25$, $\gamma = 0.24$, $\omega = 0$ and $Z_0 \sim \text{GED}(\nu)$. Panel (a) shows $\mathbb{E}(X_t^2)$ as a function of ν and d. Panel (b) gives $\mathbb{E}(X_t^2)$ as a function of ν, for $d \in \{-0.45, 0, 0.45\}$. Panel (c) considers $\mathbb{E}(X_t^2)$ as a function of d, for $\nu \in \{1.01, 3, 5\}$.

Example 2.3. From expression (2.3), one easily concludes that the kurtosis measure K_X is symmetric in θ, whenever $Z_0 \sim \text{GED}(\nu)$, for any $\nu > 1$. Figure 2.4 (a) considers SFIEGARCH$(0, d, 0)_s$ processes and shows the behavior of K_X as a function of d and ν, for $s = 2$, $\omega = 0$, $\theta = 0.25$ and $\gamma = 0.24$. Figure 2.4 (b) presents the graph of K_X as a function of d, for $\nu = 2$. From the graphs in

Figure 2.4 one concludes that the kurtosis measure is a decreasing function of ν, for each d fixed. Moreover, for each ν fixed, it is decreasing for $d \in [-0.45, 0]$ and increasing for $d \in [0, 0.45]$.

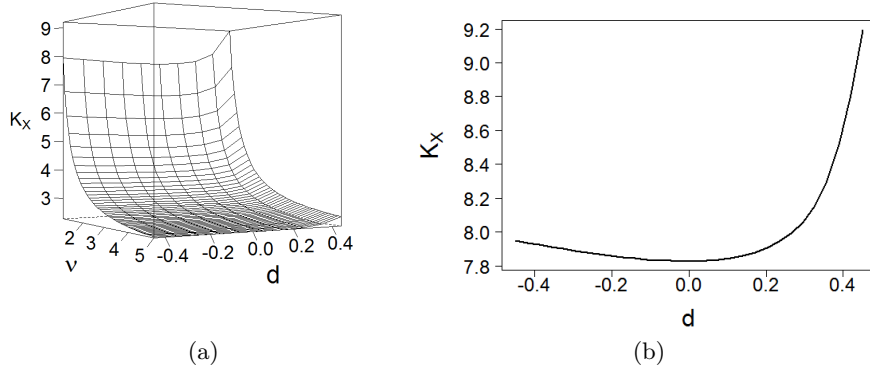

(a)   (b)

Figure 2.4: The kurtosis measure of an SFIEGARCH$(0, d, 0)_s$ process with $s = 2$, $\omega = 0$, $\theta = 0.25$, $\gamma = 0.24$ and $Z_0 \sim \text{GED}(\nu)$. Panel (a) shows the graph of the kurtosis measure as a function of d and ν. Panel (b) presents the kurtosis measure as a function of d, for $\nu = 2$ fixed.

Example 2.4. Figure 2.5 considers SFIEGARCH$(p, d, q)_s$ processes, with $p, q \in \{0, 1\}$, $d = 0.25$ and the same values of s, ω, θ and γ as in Figure 2.4. This figure presents the behavior of K_X as a function of α_1 and β_1. The cases $\alpha_1 = \beta_1$ (the polynomials have a common root) are actually equivalent to the case $\alpha_1 = 0 = \beta_1$ and, in this case, one has an SFIEGARCH$(0, d, 0)_s$ process. While Figure 2.5 (a) shows the graphs of K_X for $\alpha_1, \beta_1 \in [-0.8, 0.8]$, Figure 2.5 (b) considers only the interval $[-0.4, 0.4]$.

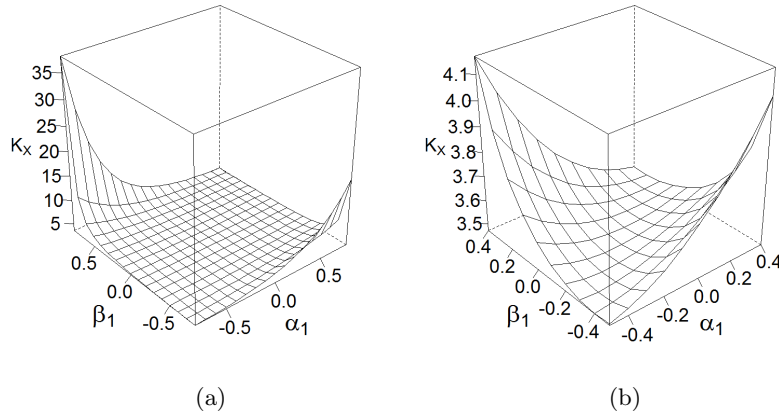(a)   (b)

Figure 2.5: The kurtosis measure of an SFIEGARCH$(p, d, q)_s$ process with $p, q \in \{0, 1\}$, $s = 2$, $\omega = 0$, $\theta = 0.25$, $\gamma = 0.24$ and $Z_0 \sim \text{GED}(\nu)$. Panel (a) shows the kurtosis measure as a function of α_1 and β_1 for $\alpha_1, \beta_1 \in [-0.8, 0.8]$. Panel (b) considers α_1 and β_1 only in the interval $[-0.4, 0.4]$ (for a better visualization).

From Figure 2.5 one observes that the behavior of K_X depends on the sign of both α_1 and β_1. Also, it goes from increasing (when $\beta_1 = -0.8$) to decreasing (when $\beta_1 = 0.8$) in α_1. Also, by comparing the graphs in Figures 2.5 (a) and (b), it is easy to see that the kurtosis measure presents small variation on its value for $(\alpha_1, \beta_1) \in [-0.4, 0.4] \times [-0.4, 0.4]$ (in this region $3 < K_X < 4.5$). Moreover, for $\alpha_1 \in [-0.8, -0.4]$ and $\beta_1 \in [-0.8, 0.8]$ the kurtosis measure varies faster than in the case $\beta_1 \in [-0.8, -0.4]$ and $\alpha_1 \in [-0.8, 0.8]$.

Lemma 2.2 and Corollary 2.3 present the autocovariance function of the process $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$ and its asymptotic behavior. Although, in practice, this stochastic process cannot be observed and hence, the sample autocovariance structure cannot be analysed, the results presented in these theorems are necessary to prove Theorem 2.3.

Lemma 2.2. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be an SFIEGARCH$(p,d,q)_s$ process, defined by (1.1) and (1.2), with $d < 0.5$. Suppose that γ and θ are not both equal to zero. Then, the autocovariance function $\gamma_{\ln(\sigma_t^2)}(\cdot)$ of the process $\{\ln(\sigma_t^2)\}_{t\in\mathbb{Z}}$ is given by*

$$\gamma_{\ln(\sigma_t^2)}(sh + r) = \sum_{k\in\mathbb{Z}}\gamma_A(sk + r)\gamma_V(sh - sk), \quad \text{for all } h \in \mathbb{Z} \quad \text{and } r \in \{0, \cdots, s-1\}, \qquad (2.4)$$

where $\gamma_A(\cdot)$ and $\gamma_V(\cdot)$ are given, respectively, by

$$\gamma_A(h) = \sum_{i=0}^{\infty} f_i f_{i+|h|}, \quad \text{for all } h \in \mathbb{Z}, \quad \text{with } f(z) := \frac{\alpha(z)}{\beta(z)} = \sum_{k=0}^{\infty} f_k z^k, \qquad (2.5)$$

and

$$\gamma_V(hs) = \sigma_g^2 \frac{(-1)^h \Gamma(1 - 2d)}{\Gamma(1 - d + h)\Gamma(1 - d - h)} \quad \text{and} \quad \gamma_V(hs + r) = 0, \qquad (2.6)$$

for all $h \in \mathbb{Z}$ and $r \in \{0, \cdots, s-1\}$, with σ_g^2 given by (1.14).

From Lemma 2.2 one easily concludes that, if $p = 0 = q$, then $\gamma_{\ln(\sigma_t^2)}(sh+r) \neq 0$ if an only if $r = 0$. Moreover, if $p > 0$ and $q = 0$, $\gamma_{\ln(\sigma_t^2)}(sh + r) = 0$, for all $h \in \mathbb{Z}$, whenever $\left\lceil -\frac{p+r}{s} \right\rceil > \left\lfloor \frac{p-r}{s} \right\rfloor$. This is so because $\gamma_A(sk + r) \neq 0$ if and only if $|sk + r| \leq p$, that is, $-\frac{p+r}{s} \leq k \leq \frac{p-r}{s}$ and $k \in \mathbb{Z}$. Thus, for $p \geq 0$ and $q = 0$, one can rewrite (2.4) as

$$\gamma_{\ln(\sigma_t^2)}(sh + r) = \begin{cases} \displaystyle\sum_{\left\lceil -\frac{p+r}{s} \right\rceil \leq k \leq \left\lfloor \frac{p-r}{s} \right\rfloor} \gamma_A(sk + r)\gamma_V(sh - sk), & \text{if} \quad \left\lceil -\frac{p+r}{s} \right\rceil \leq \left\lfloor \frac{p-r}{s} \right\rfloor; \\ 0, & \text{otherwise,} \end{cases}$$

for all $r \in \{0, \cdots, s-1\}$, and $h \in \mathbb{Z}$. In this case, it is obvious that $\sum_{h=0}^{\infty} |\rho_{\ln(\sigma_t^2)}(h)| < \infty$ if and only if, $\sum_{h=0}^{\infty} |\rho_V(h)| < \infty$. Corollary 2.3 presents the asymptotic behavior of $\gamma_{\ln(\sigma_t^2)}(h)$, as $h \to \infty$, which leads to the conclusion that this result actually holds for any p and q.

Corollary 2.3. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be an SFIEGARCH$(p,d,q)_s$ process, defined by (1.1) and (1.2), with $d < 0.5$. Suppose that γ and θ are not both equal to zero. Let $\gamma_{\ln(\sigma_t^2)}(\cdot)$ be the autocovariance function of the process $\{\ln(\sigma_t^2)\}_{t\in\mathbb{Z}}$. Then, for all $r \in \{0, \cdots, s-1\}$,*

$$\gamma_{\ln(\sigma_t^2)}(sh + r) = \gamma_V(sh)\mathscr{G}(sh + r) + o(h^{-\nu}), \quad \text{as} \quad h \to \infty, \qquad (2.7)$$

where ν is any positive number and $\mathscr{G}(\cdot)$ is a real function satisfying

$$\lim_{h\to\infty} \sum_{r=0}^{s-1} \mathscr{G}(sh + r) = \sum_{k\in\mathbb{Z}} \gamma_A(k), \quad \text{with } \gamma_A(\cdot) \text{ given by (2.5)}.$$

Theorem 2.3 presents the autocovariance function $\gamma_{\ln(X_t^2)}(\cdot)$ of the process $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$, where $\{X_t\}_{t\in\mathbb{Z}}$ is an SFIEGARCH process. This theorem also gives the asymptotic behavior of $\gamma_{\ln(X_t^2)}(sh+r)$, for all $r \in \{0, \cdots, s-1\}$, as h goes to infinity.

Theorem 2.3. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be an SFIEGARCH$(p,d,q)_s$ process, defined by (1.1) and (1.2), with $d < 0.5$. Suppose that γ and θ are not both equal to zero and $\text{Var}(\ln(Z_0^2)) := \sigma_{\ln(Z_t^2)}^2 < \infty$. Then, the autocovariance function of the process $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$ is given by*

$$\gamma_{\ln(X_t^2)}(h) = \gamma_{\ln(\sigma_t^2)}(h) + C_1 \lambda_{d,|h|-1} \mathbb{I}_{\mathbb{Z}^*}(h) + \sigma_{\ln(Z_t^2)}^2 \mathbb{I}_{\{0\}}(h), \quad \text{for all } h \in \mathbb{Z}, \qquad (2.8)$$

where $C_1 = \text{Cov}(g(Z_0), \ln(Z_0^2))$ and $\gamma_{\ln(\sigma_t^2)}(\cdot)$ is given in Lemma 2.2. Thus, for all $\nu > 0$,

$$\gamma_{\ln(X_t^2)}(sh + r) = \gamma_V(sh)\mathscr{G}(sh + r) + \pi_{d,s\lfloor \frac{sh+r-1}{s} \rfloor} \mathscr{K}(sh + r - 1) + o(h^{-\nu}), \quad \text{as } h \to \infty, \qquad (2.9)$$

for all $r \in \{0, \cdots, s-1\}$, where $\mathscr{G}(\cdot)$ and $\mathscr{K}(\cdot)$ are given, respectively, in Corollary (2.3) and Theorem 1.1.
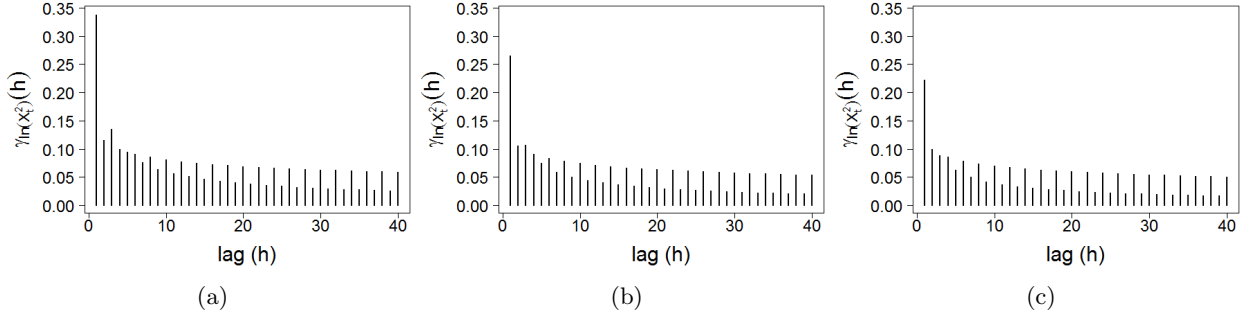
Figure 2.6: Theoretical autocovariance function $\gamma_{\ln(X_t^2)}(\cdot)$, for $h \in \{1, \cdots, 100\}$, corresponding to the process $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$, when $\{X_t\}_{t\in\mathbb{Z}}$ is an SFIEGARCH$(0,d,0)_s$ process with $d = 0.4$, $s = 2$, $\omega = 0$, $\theta = 0.25$, $\gamma = 0.24$ and $Z_0 \sim \text{GED}(\nu)$. Panel (a) considers $\nu = 1.01$. Panel (b) assumes $\nu = 2$ (Gaussian case). Panel (c) fixes $\nu = 5$.

Example 2.5. From expressions (2.4) and (2.8), one concludes that, if Z_0 is a symmetric random variable, then $\gamma_{\ln(X_t^2)}(\cdot)$ is symmetric in θ. Figures 2.6 (a) - (d) show the graphs of $\gamma_{\ln(X_t^2)}(h)$, for $h \in \{1, \cdots, 100\}$, where $\{X_t\}_{t\in\mathbb{Z}}$ is an SFIEGARCH$(0,d,0)_s$ process, with $d = 0.4$, $s = 2$, $\omega = 0$, $\theta = 0.25$, $\gamma = 0.24$ and $Z_0 \sim \text{GED}(\nu)$, for $\nu \in \{1.01, 2, 3, 5\}$, respectively. All graphs are presented in the same scale for a better visualization. The corresponding values of $\gamma_{\ln(X_t^2)}(0)$ are, respectively, 6.7228, 5.0978, 4.6445 and 4.3556. From Figure 2.6, one observes that, for each fixed h, $\gamma_{\ln(X_t^2)}(h)$ decreases as ν increases.

Example 2.6. Figure 2.7 (a) presents $\gamma_{\ln(X_t^2)}(0) = \text{Var}(\ln(X_t^2))$ as a function of ν and d, where $\{X_t\}_{t\in\mathbb{Z}}$ is an SFIEGARCH$(0,d,0)_s$ process, with $s = 2$, $\omega = 0$, $\theta = 0.25$ and $\gamma = 0.24$. Figure 2.7 (b) presents $\gamma_{\ln(X_t^2)}(0)$ as a function of d, for $\nu = 2$. From the graphs in Figure 2.7 one observes that the variance of $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$ decreases with ν. For each ν fixed, $\gamma_{\ln(X_t^2)}(0)$ is decreasing for $d \in [-0.45, 0]$ and increasing for $d \in [0, 0.45]$.

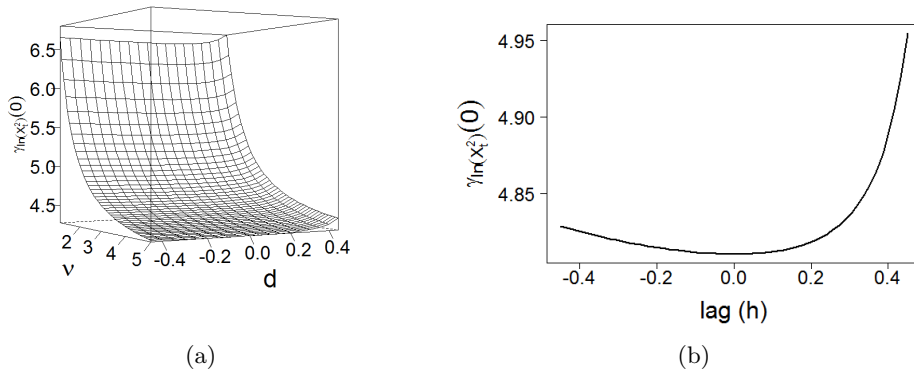Figure 2.7: This figure shows the graphs of $\sigma^2_{\ln(X_t^2)} := \gamma_{\ln(X_t^2)}(0)$, the variance of the process $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$, when $\{X_t\}_{t\in\mathbb{Z}}$ is an SFIEGARCH$(0,d,0)_s$ process with $s = 2$, $\omega = 0$, $\theta = 0.25$, $\gamma = 0.24$ and $Z_0 \sim \text{GED}(\nu)$. Panel (a) considers the variance as a function of d and ν. Panel (b) shows the variance as a function of d, for $\nu = 2$ fixed.

The following corollary compares the asymptotic behavior of $\sum_{r=0}^{s-1} \gamma_{\ln(X_t^2)}(sh+r)$ when $d < 0$ and $d > 0$.

Corollary 2.4. *Let $\gamma_{\ln(X_t^2)}(\cdot)$ be the autocovariance function of the process $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$, given in Theorem 2.3. Then,*

$$\sum_{r=0}^{s-1} \gamma_{\ln(X_t^2)}(sh+r) \sim \begin{cases} \mathscr{C}_1(h)h^{d-1}, & \text{if } d < 0; \\ \mathscr{C}_2(h)h^{2d-1}, & \text{if } d > 0, \end{cases} \qquad (2.10)$$

with

$$\mathscr{C}_1(h) := \sigma_g^2 \frac{\Gamma(1-2d)}{\Gamma(1-d)\Gamma(d)} \sum_{r=0}^{s-1} \mathscr{G}(sh+r) \quad and \quad \mathscr{C}_2(h) := \frac{1}{\Gamma(d)} \sum_{r=0}^{s-1} \mathscr{K}(sh+r-1), \qquad (2.11)$$

where $\mathscr{G}(\cdot)$ and $\mathscr{K}(\cdot)$ are given, respectively, in Corollary 2.3 and Theorem 1.1 and σ_g^2 is given by (1.14).

3 Spectral Representation

Recently economists have noticed that volatility of high frequency financial time series shows long range dependence merged with periodic behavior due to some operating features of financial markets. Periodic components are represented as marked peaks at some frequencies in the periodogram function. It is a well known result that the periodogram function is an estimator of the spectral density function. Therefore, in order to choose the best model for a time series, one should know how does the spectral density function behaves in order to gather information from the periodogram function.

Here we present the spectral density function of both processes $\{\ln(\sigma_t^2)\}_{t\in\mathbb{Z}}$ and $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$. It is easy to see that expression (3.1), in Theorem 3.1, is similar to the expression (2.5) from Hurvich et al. (2005). In this paper, the authors present the asymptotic properties of some semiparametric estimators for the long-memory parameter for a class of stochastic process which includes LMSV (*Long Memory Stochastic Volatility*) and FIEGARCH models.

Theorem 3.1. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be an SFIEGARCH$(p,d,q)_s$ process, defined by (1.1) and (1.2), with $d < 0.5$. Suppose that γ and θ, given in (1.3), are not both equal to zero and that $\alpha(z) \neq 0$ in the closed disk $\{z : |z| \leq 1\}$. If $\mathrm{Var}(\ln(Z_t^2)) := \sigma_{\ln(Z_t^2)}^2 < \infty$, for all $t \in \mathbb{Z}$, the spectral density function of $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$ is given by*

$$f_{\ln(X_t^2)}(\lambda) = f_{\ln(\sigma_t^2)}(\lambda) + \frac{C_1}{\pi}\Re\big(e^{-\mathrm{i}\lambda}\Lambda(\lambda)\big) + f_{\ln(Z_t^2)}(\lambda), \quad for \ all \ \lambda \in [0,\pi], \qquad (3.1)$$

where $f_{\ln(\sigma_t^2)}(\cdot)$ is given in (A.6), $f_{\ln(Z_t^2)}(\lambda) = \frac{\sigma_{\ln(Z_t^2)}^2}{2\pi}$ is the spectral density function of $\{\ln(Z_t^2)\}_{t\in\mathbb{Z}}$, $C_1 = \mathrm{Cov}\big(g(Z_0), \ln(Z_0^2)\big)$, $\Lambda(z) := \lambda(e^{-\mathrm{i}z})$ and $\lambda(\cdot)$ is defined in (1.9).

Notice that the spectral density function is symmetric around π. Hence, in what follows, although the graphs consider the interval $[0, 2\pi]$, one only needs to pay attention to the interval $[0, \pi]$. Moreover, all graphs are presented in the same scale and they are truncated in the y-axis for a better visualization.

Example 3.1. Figures 3.1 and 3.2 show the spectral density function of the process $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$, where $\{X_t\}_{t\in\mathbb{Z}}$ is an SFIEGARCH$(0,d,0)_s$ with different parameter values and $Z_0 \sim \mathcal{N}(0,1)$. Since Z_0 is a symmetric random variable, $\mathbb{E}(Z_0|Z_0|) = 0$ and the function $f_{\ln(X_t^2)}(\cdot)$ is symmetric in θ. Thus, in both figures, we fixed $\theta = 0.25$. In Figure 3.1 we fix $s = 2$ and, in Figure 3.2, we consider $s = 6$. For each figure, $\gamma \in \{-0.24, 0.24\}$ and $d \in \{0.1, 0.2, 0.3, 0.4\}$. From Figures 3.1 and 3.2, one observes that, for each fixed d and s, the behavior of the function completely changes as γ changes from -0.24 to 0.24 (left to right). While for $\gamma = -0.24$ the function attains its minimum in the region close to zero, for $\gamma = 0.24$ the minimum is attained close to π.

Example 3.2. Figures 3.3 - 3.5 present the spectral density function of $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$, where $\{X_t\}_{t\in\mathbb{Z}}$ is an SFIEGARCH$(p,d,q)_s$, with $Z_0 \sim \mathcal{N}(0,1)$, $p,q \in \{0,1\}$ (not both equal to zero), $s = 4$, $d = 0.25$, $\theta = 0.25$, $\gamma \in \{-0.24, 0.24\}$ and $\alpha_1, \beta_1 \in \{-0.9, -0.5, -0.1, 0.1, 0.5, 0.9\}$. For these figures the parameters values increase from left to right and from top to bottom.

Although Figure 3.5 presents only the graphs for $\alpha_1 \in \{-0.5, 0.5\}$, in the sequel we discuss the behavior of $f_{\ln(X_t^2)}(\cdot)$, for all $\alpha_1 \in \{-0.9, -0.5, -0.1, 0.1, 0.5, 0.9\}$. The remaining graphs are available upon request. From Figures 3.3 - 3.5, one concludes the following:
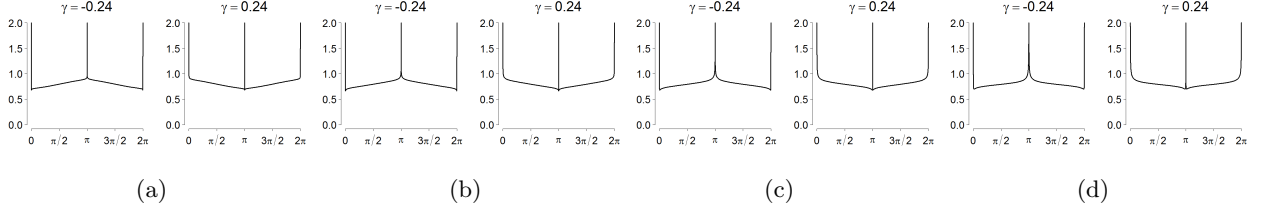
Figure 3.1: Theoretical spectral density function of the process $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$, when $\{X_t\}_{t\in\mathbb{Z}}$ is an SFIEGARCH$(0,d,0)_s$ process with $s=2$, $\theta=0.25$ and $\gamma\in\{-0.24,0.24\}$ (in each panel, from left to right). The parameter d is set as follows: in (a) $d=0.1$, in (b) $d=0.2$, in (c) $d=0.3$ and in (d) $d=0.4$.
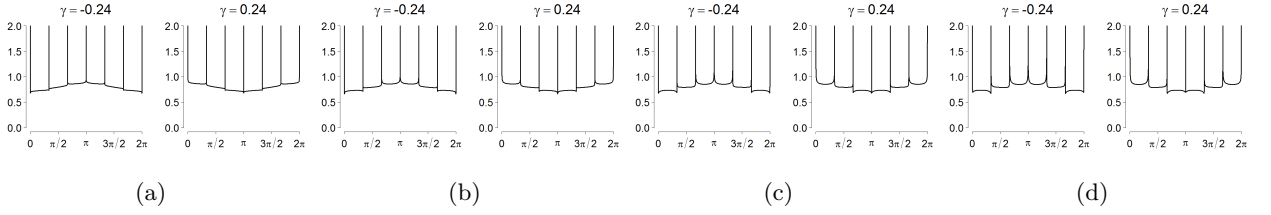

Figure 3.2: Theoretical spectral density function of the process $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$, when $\{X_t\}_{t\in\mathbb{Z}}$ is an SFIEGARCH$(0,d,0)_s$ process with $s=6$, $\theta=0.25$ and $\gamma\in\{-0.24,0.24\}$ (in each panel, from left to right). The parameter d is set as follows: in (a) $d=0.1$, in (b) $d=0.2$, in (c) $d=0.3$ and in (d) $d=0.4$.

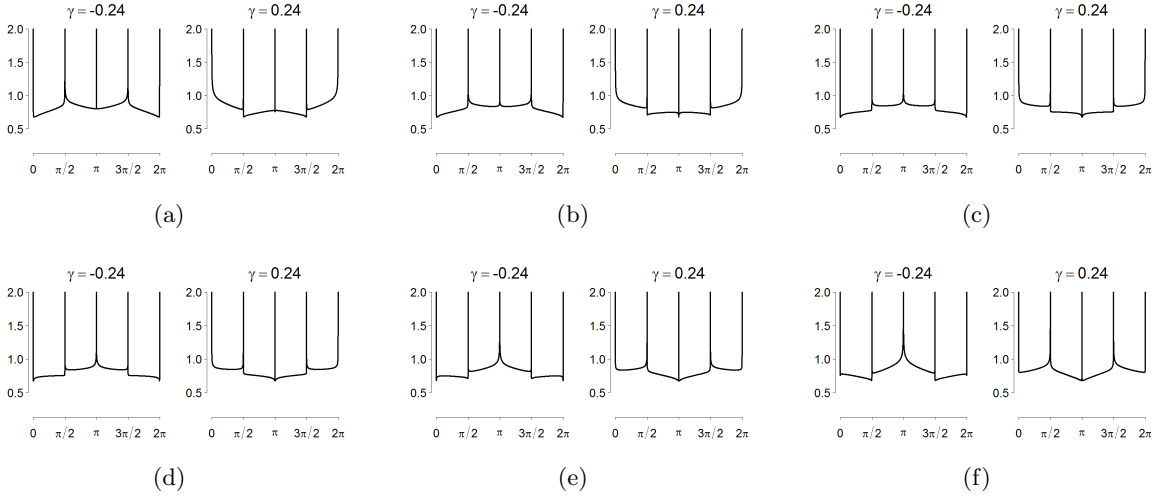Figure 3.3: Theoretical spectral density function of $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$, when $\{X_t\}_{t\in\mathbb{Z}}$ is an SFIEGARCH$(1,d,0)_s$, with $s=4$, $\omega=0$, $d=0.25$, $\theta=0.25$, $\gamma\in\{-0.24,0.24\}$ (in each panel, from left to right). The parameter α_1 is set as follows: in (a) $\alpha_1=-0.9$, in (b) $\alpha_1=-0.5$, in (c) $\alpha_1=-0.1$, in (d) $\alpha_1=0.1$, in (e) $\alpha_1=0.5$ and in (f) $\alpha_1=0.9$.

- if $p=1$ and $q=0$

  - the region where $f_{\ln(X_t^2)}(\cdot)$ attains its minimum depends not only on the sign of γ, but also on the sign of α_1. If $\gamma<0$, the minimum is attained either close to zero or close to $\pi/2$. If $\gamma>0$, the minimum is attained either close to $\pi/2$ or close to π;

  - for γ fixed, the graph of $f_{\ln(X_t^2)}(\cdot)$ slowly changes its behavior in the regions around the seasonal frequencies, as α_1 increases;

  - almost no difference is observed in the graphs of $f_{\ln(X_t^2)}(\cdot)$ when α_1 changes from -0.1 to 0.1. Actually, for these values of α_1, the function behaves as in the case $p=0=q$ (the

Figure 3.4: Theoretical spectral density function of $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$, when $\{X_t\}_{t\in\mathbb{Z}}$ is an SFIEGARCH$(1,d,0)_s$, with $s = 4$, $\omega = 0$, $d = 0.25$, $\theta = 0.25$, $\gamma \in \{-0.24, 0.24\}$ (in each panel, from left to right). The parameter β_1 is set as follows: in (a) $\beta_1 = -0.9$, in (b) $\beta_1 = -0.5$, in (c) $\beta_1 = -0.1$, in (d) $\beta_1 = 0.1$, in (e) $\beta_1 = 0.5$ and in (f) $\beta_1 = 0.9$.
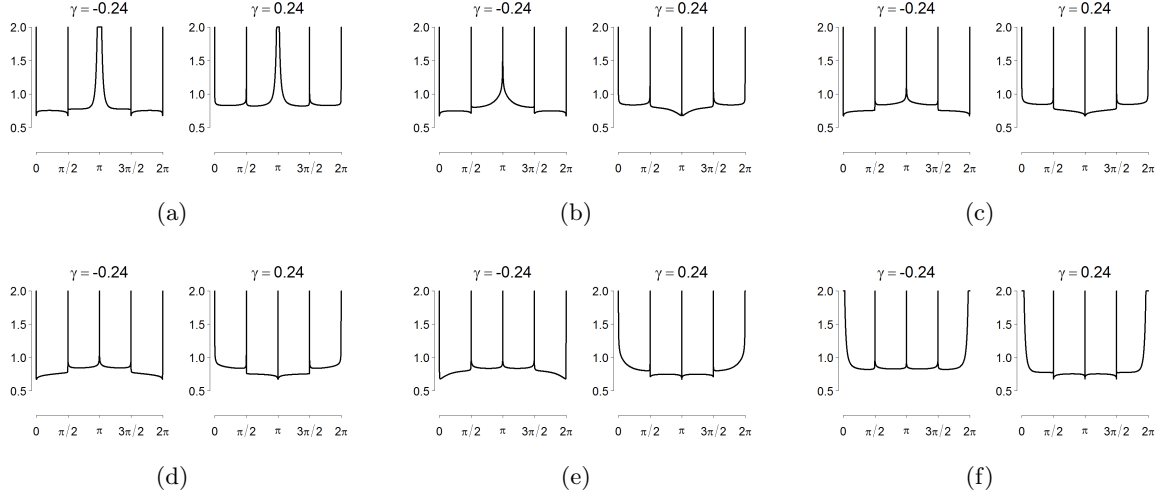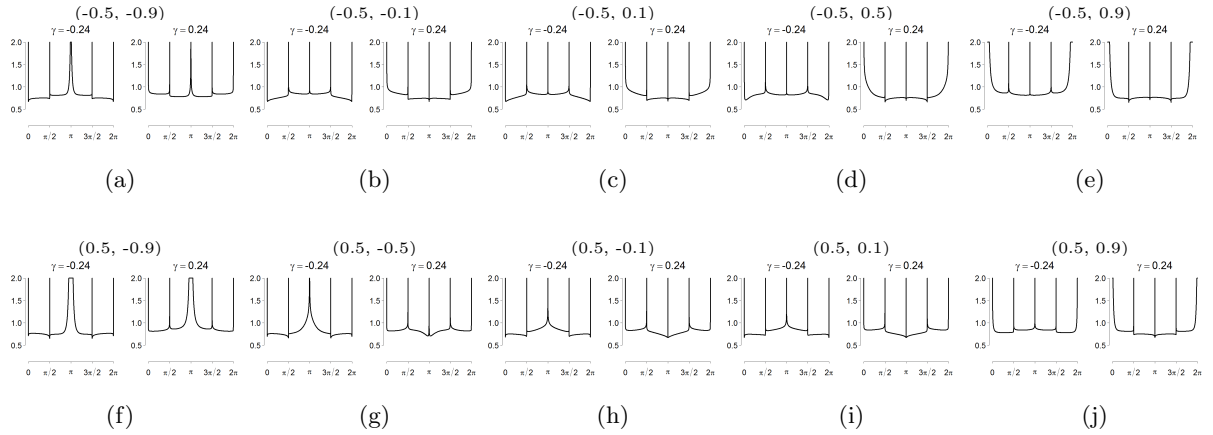

Figure 3.5: Theoretical spectral density function of $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$, when $\{X_t\}_{t\in\mathbb{Z}}$ is an SFIEGARCH$(1,d,1)_s$, with $s = 4$, $\omega = 0$, $d = 0.25$, $\theta = 0.25$. For each panel $\gamma \in \{-0.24, 0.24\}$ (from left to right). For all panels in the first row $\alpha_1 = -0.5$ and for all panels in the second row $\alpha = 0.5$. For each row $\beta_1 \in \{-0.9, -0.5, -0.1, 0.1, 0.5, 0.9\}$ (from left to right panel).

graphs can be obtained upon request);

– the graph of $f_{\ln(X_t^2)}(\cdot)$ for $\alpha_1 = -0.9$ and $\gamma = -0.24$ is very similar to the graph of the same function for $\alpha_1 = 0.9$, $\gamma = 0.24$.

• if $p = 0$ and $q = 1$

– the region where $f_{\ln(X_t^2)}(\cdot)$ attains its minimum depends on the sign of γ and on the sign of β_1. The behavior is similar to the case $p = 1$ and $q = 0$;

– for γ fixed, the changes in the graph of $f_{\ln(X_t^2)}(\cdot)$, in the regions around the seasonal frequencies, are much more visible than in the case $p = 1$ and $q = 0$;

– for $\beta_1 \in \{-0.1, 0.1\}$ the graphs are very similar to the case $p = 0$ and $q = 1$, with $\alpha_1 \in \{-0.1, 0.1\}$.

– The graphs of $f_{\ln(X_t^2)}(\cdot)$ for $\beta_1 = -0.5$ are almost identical to the graphs of this function in the case $p = 1$ and $q = 0$ with $\alpha_1 = 0.5$. The same similarity is observed between the graphs of this function for $\beta_1 = 0.5$ and $\alpha_1 = -0.5$.

- if $p = 1$ and $q = 1$

  - for γ and θ fixed, the changes in the graph of $f_{\ln(X_t^2)}(\cdot)$ are more visible as β_1 changes than when α_1 does;

  - for $\alpha_1, \beta_1 \in \{-0.1, 0.1\}$ the graphs are very similar to the case $p = 0$ and $q = 1$ (same occurs with $p = 1$ and $q = 0$);

  - generally, the graphs do not show any peculiar characteristic that is not present in the cases $p = 0$ or $q = 0$.

4 Forecasting

Let $\{X_t\}_{t \in \mathbb{Z}}$ be an SFIEGARCH$(p, d, q)_s$ process and $\{\mathcal{F}_t\}_{t \in \mathbb{Z}}$ be the filtration defined by $\mathcal{F}_t := \sigma(\{Z_s\}_{s \leq t})$. Notice that, by considering the same argument as in the proof of lemma 1 in Lopes and Prass (2013), one can show that $\{X_t\}_{t \in \mathbb{Z}}$ is a martingale difference with respect to $\{\mathcal{F}_t\}_{t \in \mathbb{Z}}$. In this case, the best predictor (in terms of the mean square error measure) for X_{t+h}, given $\mathcal{F}_t$, is $\mathbb{E}(X_{t+h}|\mathcal{F}_t) = 0$, for all $h > 0$ and $t \in \mathbb{Z}$.

Since the h-step ahead predictor for $\{X_t\}_{t \in \mathbb{Z}}$ is always zero (the mean of the process), the aim of this section is to derive expressions for the h-step ahead forecast for the processes $\{\sigma_t^2\}_{t \in \mathbb{Z}}$, $\{X_t^2\}_{t \in \mathbb{Z}}$, $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$ and $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$, for any $h > 0$. The approach considered in this work is slightly different from Lopes and Prass (2013). In Lopes and Prass (2013) two different h-step ahead predictors for σ_{t+h}^2 (consequently, for X_{t+h}^2) were proposed, both based on the h-step ahead predictor for $\ln(\sigma_{t+h}^2)$. Here we provide the exact formula for $\mathbb{E}(\sigma_{t+h}^2|\mathcal{F}_t)$, for any $h > 0$ and $t \in \mathbb{Z}$, and the relation between this expression and the ones in Lopes and Prass (2013).

Remark 4.1. In the sequel we consider the following notation, which is the same as in Lopes and Prass (2013). Let Y_t, for $t \in \mathbb{Z}$, denote any random variable. Then

- the symbol "^" denotes the h-step ahead forecast defined in terms of the conditional expectation, that is, $\hat{Y}_{t+h} = \mathbb{E}(Y_{t+h}|\mathcal{F}_t)$. Notice that this is the best linear (or non-linear) predictor in terms of mean square error value;

- the symbols "~" (e.g. $\tilde{Y}_{t+h}$) and "ˇ" (e.g. $\check{Y}_{t+h}$) denote alternative estimators;

- $\hat{\ln}(Y_{t+h})$ denotes the h-step ahead forecast of $\ln(Y_{t+h})$ (analogously, for "~" and "ˇ");

- we follow the approach usually considered in the literature and denote the h-step ahead forecast of Y_{t+h}^2 as $\hat{Y}_{t+h}^2$ instead of $\widehat{Y_{t+h}^2}$ and, to avoid confusion, we will denote the square of $\hat{Y}_{t+h}$ as $(\hat{Y}_{t+h})^2$ (analogously, for "~" and "ˇ").

To obtain the predictors for the processes $\{\sigma_t^2\}_{t \in \mathbb{Z}}$ and $\{X_t^2\}_{t \in \mathbb{Z}}$ observe that the i.i.d. property of $\{Z_t\}_{t \in \mathbb{Z}}$ implies that $\mathbb{E}(Z_{n+h}^2|\mathcal{F}_n) = \mathbb{E}(Z_{n+h}^2) = 1$, σ_{n+1}^2 is $\mathcal{F}_n$-measurable and σ_n^2 and Z_n^2 are independent, for all $n \in \mathbb{Z}$ and $h > 0$. Therefore, the h-step ahead forecast for X_{n+h}^2 given $\mathcal{F}_n$ is given by

$$\hat{X}_{n+h}^2 := \mathbb{E}(X_{n+h}^2|\mathcal{F}_n) = \mathbb{E}(\sigma_{n+h}^2|\mathcal{F}_n) := \hat{\sigma}_{n+h}^2, \quad \text{for all} \quad h > 0.$$

In particular, $\hat{\sigma}_{n+1}^2 = \sigma_{n+1}^2$ and hence the 1-step ahead forecast of X_{n+1}^2, given $\mathcal{F}_n$, is simply $\hat{X}_{n+1}^2 = \sigma_{n+1}^2$. Theorem 4.1 provides the general formula for $\hat{\sigma}_{n+h}^2$, when $h > 1$.

Theorem 4.1. *Let $\{X_t\}_{t \in \mathbb{Z}}$ be a stationary SFIEGARCH$(p, d, q)_s$ process with $\mathbb{E}(\sigma_t^2) < \infty$. Then, for any fixed $n \in \mathbb{Z}$, the h-step ahead forecast of σ_{n+h}^2 (consequently, for X_{n+h}^2), given $\mathcal{F}_n$, can be expressed as*

$$\hat{\sigma}_{n+h}^2 = e^\omega \prod_{k=h-1}^{\infty} \exp\{\lambda_{d,k} g(Z_{n+h-1-k})\} \prod_{\ell=0}^{h-2} \mathbb{E}(\exp\{\lambda_{d,\ell} g(Z_0)\}), \quad \text{for all } h > 1. \tag{4.1}$$

Moreover, if $\mathbb{E}(\sigma_{n+h}^4) < \infty$, the mean square errors of forecast for σ_{n+h}^2 and X_{n+h}^2, respectively denoted as $mse(\sigma_{n+h}^2)$ and $mse(X_{n+h}^2)$, are given by

$$mse(\sigma_{n+h}^2) = \left[\prod_{k=0}^{h-2} \mathbb{E}\big(\exp\{2\lambda_{d,k} g(Z_0)\}\big) - \prod_{\ell=0}^{h-2} \Big[\mathbb{E}\big(\exp\{\lambda_{d,\ell} g(Z_0)\}\big)\Big]^2 \right] \prod_{j=h-1}^{\infty} \mathbb{E}\big(\exp\{2\lambda_{d,j} g(Z_0)\}\big) \quad (4.2)$$

and

$$mse(X_{n+h}^2) = \mathbb{E}(\sigma_0^4)\big[\mathbb{E}(Z_0^4) - 1\big] + mse(\sigma_{n+h}^2), \quad \text{for all } h > 1.$$

Lopes and Prass (2013) propose two h-step ahead predictors for the process $\{\sigma_t^2\}_{t\in\mathbb{Z}}$ in the context of FIEGARCH processes. The first one, denoted by $\check{\sigma}_{n+h}^2$, was obtained through the relation $\check{\sigma}_{n+h}^2 := \exp\{\hat{\ln}(\sigma_{n+h}^2)\}$, where $\hat{\ln}(\sigma_{n+h}^2)$ is the h-step ahead predictor for $\ln(\sigma_{n+h}^2)$. The second predictor, denoted by $\tilde{\sigma}_{n+h}^2$, was derived upon considering an order 2 Taylor's expansion of the exponential function. The authors also showed that $\check{\sigma}_{n+h}^2$ and $\tilde{\sigma}_{n+h}^2$ satisfy

$$\tilde{\sigma}_{n+h}^2 = \begin{cases} \exp\{\hat{\ln}(\sigma_{n+h}^2)\} = \check{\sigma}_{n+h}^2, & \text{if } h = 1; \\ \exp\{\hat{\ln}(\sigma_{n+h}^2)\}\left[1 + \frac{1}{2}\sigma_g^2 \sum_{k=0}^{h-2} \lambda_{d,k}^2\right] = \check{\sigma}_{n+h}^2\left[1 + \frac{1}{2}\sigma_g^2 \sum_{k=0}^{h-2} \lambda_{d,k}^2\right], & \text{if } h > 1. \end{cases} \quad (4.3)$$

With obvious identifications, the same predictors $\check{\sigma}_{n+h}^2$ and $\tilde{\sigma}_{n+h}^2$ can be defined for SFIEGARCH processes. However, by following the same steps as in Lopes and Prass (2013), it can be shown that both $\check{\sigma}_{n+h}^2$ and $\tilde{\sigma}_{n+h}^2$ are biased estimators for σ_{n+h}^2. On the other hand $\hat{\sigma}_{n+h}^2$, given in (4.1), not only is an unbiased estimator but also it is the best predictor for σ_{n+h}^2 in terms of the mean square error measure.

Now, to obtain the h-step ahead predictor for $\ln(X_{n+h}^2)$ observe that, from Definition 1.1 and from the i.i.d. property of $\{Z_t\}_{t\in\mathbb{Z}}$,

$$\hat{\ln}(X_{n+h}^2) := \mathbb{E}\big(\ln(X_{n+h}^2)|\mathcal{F}_n\big) = \mathbb{E}\big(\ln(\sigma_{n+h}^2)|\mathcal{F}_n\big) + \mathbb{E}\big(\ln(Z_{n+h}^2)|\mathcal{F}_n\big) := \hat{\ln}(\sigma_{n+h}^2) + \mathbb{E}(\ln(Z_0^2)),$$

for all $n \in \mathbb{Z}$ and $h > 0$. The expressions for $\hat{\ln}(\sigma_{n+h}^2) := \mathbb{E}(\ln(\sigma_{n+h}^2)|\mathcal{F}_t)$ and for the mean square errors of forecast for $\ln(\sigma_{n+h}^2)$ and $\ln(X_{n+h}^2)$ are given in Theorem 4.2.

Theorem 4.2. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be an SFIEGARCH$(p,d,q)_s$ process. Then, for any fixed $n \in \mathbb{Z}$, the h-step ahead forecast $\hat{\ln}(\sigma_{n+h}^2)$ of $\ln(\sigma_{n+h}^2)$, given $\mathcal{F}_n$, can be expressed as*

$$\hat{\ln}(\sigma_{n+h}^2) = \omega + \sum_{k=0}^{\infty} \lambda_{d,k+h-1} g(Z_{n-k}), \quad \text{for all } h > 0. \quad (4.4)$$

Moreover, if $h = 1$, the mean square errors of forecast for $\ln(\sigma_{n+h}^2)$ and $\ln(X_{n+h}^2)$ are both equal to zero and, for any $h > 1$, are given, respectively, by

$$mse(\ln(\sigma_{n+h}^2)) = \sigma_g^2 \sum_{k=0}^{h-2} \lambda_{d,k}^2 \quad \text{and} \quad mse(\ln(X_{n+h}^2)) = mse(\ln(\sigma_{n+h}^2)), \quad (4.5)$$

where $\sigma_g^2 = \mathbb{E}\big([g(Z_0)]^2\big)$.

Remark 4.2. Lopes and Prass (2013) consider the h-step ahead predictor for $\ln(X_{n+h}^2)$ defined as $\check{\ln}(X_{n+h}^2) = \hat{\ln}(\sigma_{n+h}^2)$, for any $n \in \mathbb{Z}$ and $h > 0$. This is an unbiased estimator if and only if $\mathbb{E}(\ln(Z_0^2)) = 0$.

From (4.1) and (4.4), one concludes that, $\hat{\sigma}_{n+1}^2 = \check{\sigma}_{n+1}^2 = \tilde{\sigma}_{n+1}^2$ and

$$\hat{\sigma}_{n+h}^2 = \check{\sigma}_{n+h}^2 \prod_{\ell=0}^{h-2} \mathbb{E}\big(\exp\{\lambda_{d,\ell} g(Z_0)\}\big) = \tilde{\sigma}_{n+h}^2 \left[1 + \frac{1}{2}\sigma_g^2 \sum_{k=0}^{h-2} \lambda_{d,k}^2\right]^{-1} \prod_{\ell=0}^{h-2} \mathbb{E}\big(\exp\{\lambda_{d,\ell} g(Z_0)\}\big), \quad (4.6)$$

for any fixed $n \in \mathbb{Z}$ and $h > 1$.

Now, let $\{r_t\}_{t \in \mathbb{Z}}$ be the stochastic process defined by

$$r_t = \mu + \sum_{k=0}^{\infty} \psi_k X_{t-k} := \mu + \psi(\mathcal{B})X_t, \quad \text{for all } t \in \mathbb{Z}, \tag{4.7}$$

where $\mu \in \mathbb{R}$, $\{\psi_k\}_{t \in \mathbb{Z}}$ is a sequence of real numbers satisfying $\sum_{k=0}^{\infty} \psi_k^2 < \infty$ and $\{X_t\}_{t \in \mathbb{Z}}$ is an SFIEGARCH$(p, d, q)_s$ with $\sup_{t \in \mathbb{Z}}\{\mathrm{Var}(X_t^2)\} < \infty$. Notice that, $\sum_{k=0}^{\infty} \psi_k^2 < \infty$ and $\sup_{t \in \mathbb{Z}}\{\mathrm{Var}(X_t^2)\} < \infty$ imply $\mathrm{Cov}(X_k, X_j) = 0$, for all $k \neq j$, and

$$\mathbb{E}\left(\left|\sum_{k=m}^{n} \psi_k X_{t-k}\right|^2\right) \leq \sup_{t \in \mathbb{Z}}\{\mathrm{Var}(X_t^2)\} \sum_{k=m}^{n} \psi_k^2, \quad \text{for all } m \leq n.$$

Therefore, (4.7) converges in L^2 (Cauchy convergence criterion) and hence $\{r_t\}_{t \in \mathbb{Z}}$ is well defined.

Example 4.1. Let $\phi(\cdot)$ and $\varphi(\cdot)$ be the polynomials of order p and q, with no common roots, respectively defined by $\phi(z) := \sum_{k=0}^{p_1}(-\phi_k)z^k$ and $\varphi(z) := \sum_{j=0}^{q_1}(-\varphi_j)z^j$, with $\phi_0 = \varphi_0 = -1$. Given a weakly stationary SFIEGARCH$(p, d, q)_s$ process $\{X_t\}_{t \in \mathbb{Z}}$, define $\{r_t\}_{t \in \mathbb{Z}}$ by

$$\phi(\mathcal{B})(r_t - \mu) = \varphi(\mathcal{B})X_t, \quad \text{for all } t \in \mathbb{Z}.$$

Observe that $\{r_t\}_{t \in \mathbb{Z}}$ is an ARMA(p_1, q_1) process and hence it can be rewritten as in equation (4.7), whenever $\phi(z) \neq 0$ in the closed disk $\{z : |z| \leq 1\}$, with $\{\psi_k\}_{k \in \mathbb{Z}}$ uniquely defined through

$$\sum_{k=0}^{\infty} \psi_k z^k = \psi(z) := \frac{\varphi(z)}{\phi(z)}, \quad |z| \leq 1.$$

Theorem 4.3 provides the h-step ahead forecast for the process $\{r_t^2\}_{t \in \mathbb{Z}}$, with r_t defined in (4.7). Similar equations can be derived if the assumption that $\{X_t\}_{t \in \mathbb{Z}}$ is an SFIEGARCH$(p, d, q)_s$ is replaced by any other ARCH-type model. This result is applied in Section 5 to compare the forecasting performance of the different models considered in the time series analysis.

Theorem 4.3. *Let $\{r_t\}_{t \in \mathbb{Z}}$ be defined by (4.7). Then, for any fixed $n \in \mathbb{Z}$, the h-step ahead forecast $\hat{r}_{n+h}^2$ of r_{n+h}^2, given $\mathcal{F}_n$, can be expressed as*

$$\hat{r}_{n+h}^2 = \mu^2 + \sum_{k=0}^{h-1} \psi_k^2 \hat{\sigma}_{n+h-k}^2 + \sum_{j=h}^{\infty}\sum_{\ell=h}^{\infty} \psi_j \psi_\ell X_{n+h-j} X_{n+h-\ell} + 2\mu \sum_{i=h}^{\infty} \psi_i X_{n+h-i}, \quad \textit{for any } h > 0, \tag{4.8}$$

where $\hat{\sigma}_{n+1}^2 = \sigma_{n+1}^2$ is given by (1.2) and $\hat{\sigma}_{n+h}^2$ is the h-step ahead forecast of X_{n+h}^2 given in (4.1), for all $h > 1$.

5   An Application

In this section we analyze the behavior of the intraday volatility of the S&P500 US stock index log-return time series in the period from December 13, 2004 to October 10, 2009. The trading hours are from 8:30 am to 3:15 pm (Chicago time) and the intraday frequency of the original index time series is 15 minutes, which gives a total of 33993 observations (1259 days). The fifteen-minute log-returns (see Remark 5.1) are aggregated to obtain a one-hour[1] log-return time series $\{R_t\}_{t=1}^{8498}$.

---

[1]Up to this day, we only had access to the stock index time series with sampling frequency equal to 15 minutes. Therefore, a thorough study on the existence of microstructural noise could not be performed. With the data set available we create a signature volatility plot (see, for instance, Andersen et al., 2000) by considering only sample frequencies which are multiples of 15 minutes (e.g., 15, 30 and 45 minutes). The minimum and maximum sample frequency values considered were, respectively, 15 minutes (which is the original sampling frequency) and 405 minutes (which corresponds to one-day log-returns). Under this scenario, the average realized volatilities were all close to each other regardless the sampling frequencies considered.

Remark 5.1. The fifteen-minutes log-return time series $\{R_t^{(15)}\}_{t=1}^n$ is defined as

$$R_t^{(15)} = 100 \times \ln\left(\frac{P_t}{P_{t-1}}\right), \quad \text{for all } t \in \{1, \cdots, n\},$$

where $\{P_t\}_{t=0}^n$ is the index time series, with $n = 33992$ ($t = 0$ corresponds to the first available observation). The one-hour log-return time series $\{R_t\}_{t=1}^{8498}$ is obtained by letting $R_t = R_{\tau-3}^{(15)} + \cdots + R_\tau^{(15)}$, with $\tau = 4t$, for all $t \in \{1, \cdots, 8498\}$.

Figure 5.1 (a) and (b) show, respectively, the S&P500 US stock index log-return time series and the one-hour log-return time series $\{R_t\}_{t=1}^{8498}$, in the studied period. Notice that, for any $t \in \{1, \cdots, 1259\}$ the times 3:15 pm (closing time) from day t and 8:30 am (opening time) from day $t+1$ are equivalent (there is no trading between these two periods). Therefore, there are 27 index values and, consequently, 27 available 15-minutes log-returns for each trading day. It is easy to see that, by applying the aggregation equation described in Remark 5.1, the trading time associated to one-hour log-returns for two consecutive days are not necessarily the same. In particular, the following holds

- the first available index value corresponds to December 13, 2004 8:30 am (or equivalently, December 12, 2004 3:15 pm). Consequently, the first one-hour log-return for day 1 corresponds to 9:30 am;

- the last one-hour log-return for day 1 corresponds to 2:30 pm and, since the trading day ends at 3:15 pm, the next one-hour log-return will be the aggregation of 15-minutes log-returns for 2:30 pm, 2:45 pm, 3:00 pm, 3:15 pm (or equivalently, 8:30 am from day 2) and 8:45 am from day 2. Consequently, the first one-hour log-return for day 2 corresponds to 8:45 am;

- whenever the first one-hour log-return corresponds to 9:30 am, there are only 6 one-hour log-returns for the corresponding day;

- for every four days there is one day with 6 one-hour log-returns followed by 3 consecutive days with 7 one-hour log-returns. This fact may or may not induce a cyclical behavior.
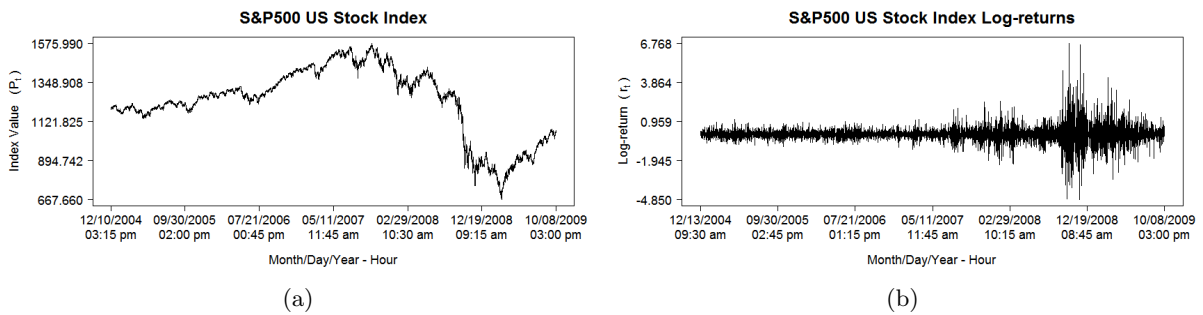
Figure 5.1: This figure considers the S&P500 US stock index time series in the period from December 13, 2004 to October 10, 2009. The intraday frequency of the index time series is 15 minutes, which gives a total of 33993 observations (1259 days). Panel (a) shows the original index time series. Panel (b) presents the one-hour log-return time series, with $n = 8498$ observations, obtained by aggregating the fifteen-minute returns.

From Figure 5.1 (b) it is clear that the volatility in the period from 2004 to 2007 is much lower[2] than in the period from 2007 to 2009. Given that high volatility is more concerning and more difficult to model than low volatility we shall discard the first 3996 observations (592 days). We also reserve the last 40 days of data (270 observations) to analyze the out-of-sample forecasting

---

[2]This fact is already known in the literature and it is beyond the scope of this work to discuss possible causes for this behavior.

performance of the fitted models. The remaining time series corresponds to the period from March 21, 2007 to August 13, 2008 and has 4232 observations. This time series shall be denoted by $\{r_t\}_{t=1}^{4232}$, where $r_t := R_{t+3996}$, for all $t = 1, \cdots, 4232$.

Remark 5.2. The two highest peaks in Figure 5.1 (a) correspond, respectively, to October 10, 2008 and November 21, 2008. The two lowest values in Figure 5.1 (a) correspond, respectively, to October 06, 2008 and November 20, 2008. October 10, 2008 is the day with the highest trading volume ever for the S&P 500 index. In this day, the trading volume for the SPY SPDR surpassed 871 million shares (see, for instance, AMEX:SPY daily prices for October 2008 from Yahoo! Finance). On November 20, 2008 the S&P 500 index closed at 752.44, its lowest since early 1997.

The descriptive statistics for the log-return time series $\{r_t\}_{t=1}^{4232}$ are given in Table 5.1. For comparison, this table also shows the descriptive statistics for time series $\{R_t\}_{t=1}^{8498}$. From Table 5.1 one observes that both time series $\{R_t\}_{t=1}^{8498}$ and $\{r_t\}_{t=1}^{4232}$ have mean approximately equal to zero but high skewness values, which usually indicates a non-symmetric distribution. However, we shall notice that, for these time series, the high skewness values could be due to the presence of some outliers instead of non-symmetry. In fact, upon replacing all values higher than eight standard deviations by the sample mean of the corresponding time series, the skewness values for $\{R_t\}_{t=1}^{8498}$ and $\{r_t\}_{t=1}^{4232}$ are, respectively, 0.0398 and -0.0018, which reinforces our claim. Nevertheless, the possible outliers are not removed in the analysis to be performed in the sequel. The aim of this approach is to observe whether the SFIEGARCH model captures or not this feature.

Table 5.1: Descriptive statistics for the S&P500 one-hour log-return time series $\{R_t\}_{t=1}^{8498}$ and for the time series $\{r_t\}_{t=1}^{4232}$, where $r_t := R_{t+3996}$, for all $t \in \mathbb{Z}$.

Period	n	mean	st. dev.	kurtosis	skewness
2007 - 2009	4232	-0.0078	0.6931	14.5986	0.4156
2004 - 2009	8498	-0.0013	0.5168	23.6251	0.4609

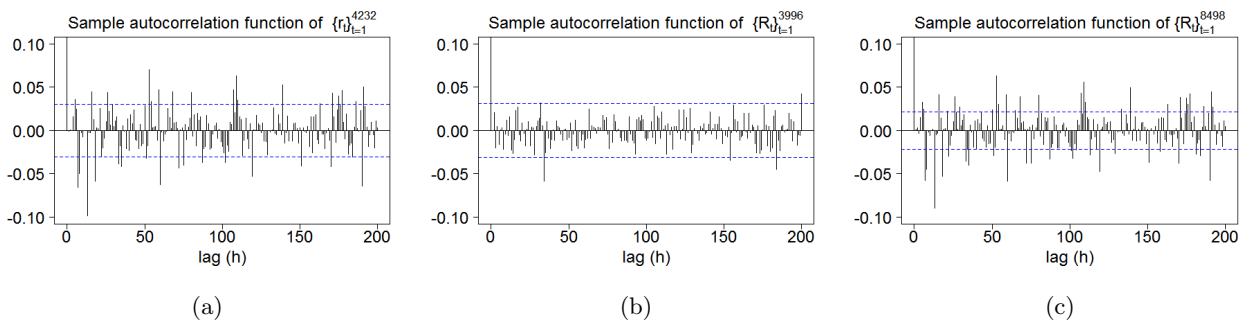**Note:** st. dev. := standard deviation.

(a)   (b)   (c)

Figure 5.2: This figure shows the sample autocorrelation function for the S&P500 US stock index one-hour log-return time series $\{R_t\}_{t=1}^{8498}$ and its sub-samples. For $\{R_t\}_{t=1}^{8498}$ the time index $t = 1$ corresponds to December 13, 2004 at 09:30 am, $t = 3997$ corresponds to March 21, 2007 at 09:30 am and $t = 8498$ corresponds to October 10, 2009 at 03:00 pm. Panel (a) considers the time series $\{r_t\}_{t=1}^{4232}$ with $r_t := R_{t+3996}$. Panel (b) corresponds to the sub-sample $\{R_t\}_{t=1}^{3996}$. Panel (c) shows the sample autocorrelation function of $\{R_t\}_{t=1}^{8498}$. For better visualization the y-axis only shows the interval $[-0.10, 0.10]$. In all graphs the dashed lines correspond to $\pm 1.96/\sqrt{n}$, where n is the sample size of the corresponding time series.

Figure 5.2 (a), (b) and (c) shows, respectively, the sample autocorrelation functions for the log-return time series $\{r_t\}_{t=1}^{4232}$, $\{R_t\}_{t=1}^{3996}$ and $\{R_t\}_{t=1}^{8498}$. From this figure we observe that the log-return time series presents small (notice the scales) autocorrelations (but significatively different from zero) for some lags $h > 0$. Upon comparing Figure 5.2 (a), (b) and (c), one observes that while the sample autocorrelation functions for $\{r_t\}_{t=1}^{4232}$ and $\{R_t\}_{t=1}^{8498}$ seem identical, the difference is remarkable when

considering the sample autocorrelation functions for $\{R_t\}_{t=1}^{3996}$ and $\{R_t\}_{t=1}^{8498}$. This indicates that the correlation in the series is mainly due to the last observations of the time series $\{R_t\}_{t=1}^{8498}$.

From Figures 5.2 (a) and (c) one observes that the autocorrelation value with higher magnitude is associated to the lag 13 (roughly 2 days). Next, in order of magnitude (including the values not reported in Figure 5.2), are the autocorrelations associated to lags 53, 7, 263, 190, 109, 60, 18 (roughly 2 and a half days), 203, 218, 119, 139, 191 and 8. The remaining autocorrelation values are all smaller (in magnitude) than the one associated to lag 8 and, therefore, very close to the confidence limits (this includes the autocorrelation values with lag higher than 200, which are not reported in Figure 5.2). Recall that the aggregation rule considered implies that for every 4 days there is one day with only 6 one-hour log-returns followed by 3 days with 7 one-hour log-returns. Moreover, notice that 53, 109, 139, 190, 191, 218 are very close to multiples of 27 (total number of observations in 4 days). Furthermore, while 60 and 263 differ from multiples of 27 by approximately 7, 119 and 203 differ from approximately 13.

The facts just mentioned, indicate a short-memory cyclical behavior of length 27. On the other hand, there is also evidence that a single seasonal polynomial may not be enough to remove the correlation. For this reason we shall consider a constrained[3] $\text{ARMA}(p_1, q_1)$ model for the log-return time series. Under this assumption we have

$$\phi(\mathcal{B})(r_t - \mu) = \varphi(\mathcal{B})X_t, \quad \text{for all } t \in \mathbb{Z}, \tag{5.1}$$

where $\mu \in \mathbb{R}$, $\phi(z) = \sum_{k=0}^{p_1}(-\phi_k)z^k$, $\varphi(z) = \sum_{j=0}^{q_1}(-\varphi_j)z^j$, with $\phi_0 = \varphi_0 = -1$, and $\{X_t\}_{t \in \mathbb{Z}}$ is a white noise process. Notice that, by letting p_1 and q_1 be large enough, equation (5.1) also covers the seasonal ARMA class of model, denoted by $\text{SARMA}(p_1, q_1) \times (P, Q)_s$ (see Remark 5.3).

Remark 5.3. For any $d, D \in \mathbb{N}$, let $(1 - \mathcal{B})^d$ and $(1 - \mathcal{B}^s)^D$ be, respectively, the non-seasonal and seasonal difference operators iterated, respectively, d and D times. Let $A(z) = \sum_{k=0}^{P}(-A_k)z^k$, $a(z) = \sum_{k=0}^{p}(-a_k)z^k$, $M(z) = \sum_{k=0}^{Q}(-M_k)z^k$ and $m(z) = \sum_{k=0}^{q}(-m_k)z^k$ be polynomials, respectively, of order P, p, Q and q, with $A_0 = a_0 = M_0 = m_0 = -1$. A seasonal autoregressive integrated moving average model, denoted by $\text{SARIMA}(p, d, q) \times (P, D, Q)_s$, is defined by (for more details and for the definition of a SARFIMA process, see Bisognin and Lopes, 2009)

$$A(\mathcal{B}^s)a(\mathcal{B})(1 - \mathcal{B})^d(1 - \mathcal{B}^s)^D(Y_t - \mu) = M(\mathcal{B}^s)m(\mathcal{B})\varepsilon_t, \quad \text{for all } t \in \mathbb{Z}, \tag{5.2}$$

where $\mu \in \mathbb{R}$ and $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ is a white noise process with zero mean and variance σ_ε^2. In particular, when $d = D = 0$, (5.2) is called a $\text{SARMA}(p, q) \times (P, Q)_s$ model. It is immediate that by letting $d = D = 0$, $\phi(z) = A(\mathcal{B}^s)a(\mathcal{B})$ and $\varphi(z) = M(\mathcal{B}^s)m(\mathcal{B})$, (5.2) can be rewritten as an $\text{ARMA}(p_1, q_1)$ model, given in (5.1), with $p_1 = P + p$ and $q_1 = Q + q$.

Figure 5.3 (a) and (b) present, respectively, the sample autocorrelation and the periodogram functions for the time series $\{\ln(r_t^2)\}_{t=1}^{4232}$. We observe that both functions indicate long-memory and cyclical behavior, with seasonal parameter $s = 7$ (one day cycle). To account for the long-memory cyclical behavior in the volatility, we shall consider an $\text{SFIEGARCH}(p_2, d, q_2)_s$, described in Section 1. To confirm the importance of including the seasonal effect associated to long-memory[4] we also consider a $\text{FIEGARCH}(p_2, d, q_2)$ (Bollerslev and Mikkelsen, 1996), and an $\text{EGARCH}(p_2, q_2)$ (Nelson, 1991) model.

Recall that, for all models just mentioned, $\{X_t\}_{t \in \mathbb{Z}}$ is written as $X_t = \sigma_t Z_t$, for any $t \in \mathbb{Z}$, where $\{Z_t\}_{t \in \mathbb{Z}}$ is an i.i.d. sequence of random variables with $\mathbb{E}(Z_0) = 0$ and $\text{Var}(Z_0) = 1$. For the $\text{SFIEGARCH}(p_2, d, q_2)_s$ model one has

$$\ln(\sigma_t^2) = \omega + \frac{\alpha(\mathcal{B})}{\beta(\mathcal{B})}(1 - \mathcal{B}^s)^{-d}g(Z_{t-1}), \quad \text{for all } t \in \mathbb{Z},$$

---

[3]By constrained we mean that some ϕ_i and θ_j will be fixed as zero.

[4]The FIEGARCH model captures non-seasonal long-memory and short-memory cyclical behaviors (if p_2 and q_2 are large enough). The EGARCH model is only able to describe short-memory cyclical behavior.

(a)   (b)
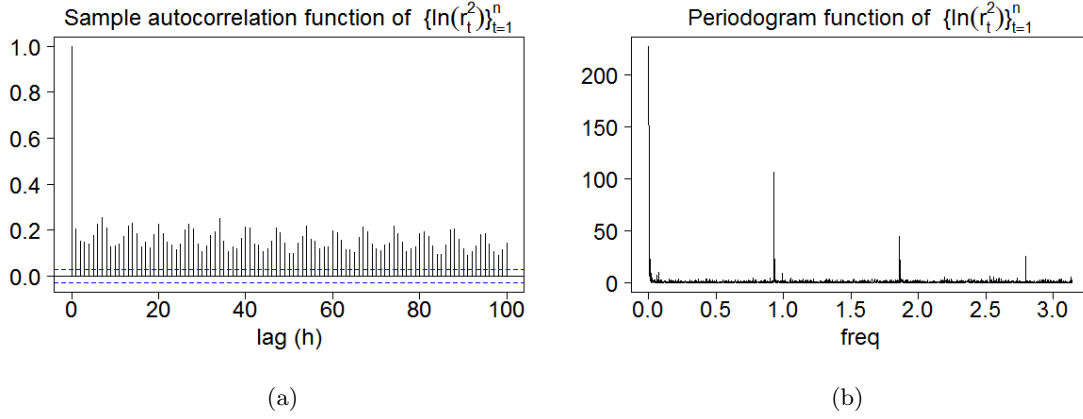
Figure 5.3: This figure considers the time series $\{\ln(r_t^2)\}_{t=1}^{4232}$, where $\{r_t\}_{t=1}^{4232}$ is the S&P500 US stock index log-return time series corresponding to the period from March 21, 2007 at 09:30 am to August 13, 2009 at 03:00 pm. Panel (a) shows the sample autocorrelation function of $\{\ln(r_t^2)\}_{t=1}^{4232}$. Panel (b) presents its periodogram function.

with $\alpha(\cdot)$, $\beta(\cdot)$ and $(1-\mathcal{B}^s)^{-d}$ be defined as in (1.4) and (1.5), and $g(Z_t) = \theta Z_t + \gamma\big[|Z_t| - \mathbb{E}\big(|Z_t|\big)\big]$. The FIEGARCH and EGARCH models are particular cases of the SFIEGARCH model obtained from (1.2), respectively, when $s = 1$ and $d = 0$.

Remark 5.4. By comparing the sample kurtosis values given in Table 5.1 with the theoretical kurtosis values of an SFIEGARCH$(p, d, q)_s$ process (see Figures 2.4 and 2.5) we conclude that the best SFIEGARCH fit for the data more likely will have $p, q > 0$.

Remark 5.5. Our first intention was to compare the performance of the SFIEGARCH model with the PLM-GARCH (Bordignon et al., 2007, 2009), since both models are able to describe long-memory cyclical behavior. Analogously to the SFIEGARCH case, upon considering the PLM-GARCH we would also include a HYGARCH (Davidson, 2004) and a GARCH model (Bollerslev, 1986). It turns out that we were not able to fit any PLM-GARCH model for which the squared residuals time series $\{\hat{z}_t^2\}_{t=1}^n$ shows no correlation and at the same time the positivity criteria for $\{\sigma_t^2\}_{t=1}^n$ would be satisfied. The number of cases for which $\sigma_t^2 < 0$ were always too high to be replaced by a constant or by $|\sigma_t^2|$.

5.1   Model Selection Procedure

Parameter estimation is carried out by applying the so-called quasi-likelihood method. In this method, the estimate $\hat{\boldsymbol{\eta}}$ for the vector of unknown parameters $\boldsymbol{\eta}$ is the value that maximizes

$$\mathcal{L}_n(\boldsymbol{\eta}) := -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^n\left[\ln(\sigma_t^2) + \frac{(r_t - \mu_t)^2}{\sigma_t^2}\right],$$

where $\mu_t := \mu + \sum_{k=1}^{p_1}\phi_k(r_{t-k} - \mu) + \sum_{j=1}^{q_1}\varphi_j X_{t-j}$, for all $t \in \mathbb{Z}$. The recursion starts by setting $r_t = \bar{r}$, where $\bar{r} = \frac{1}{n}\sum_{t=1}^n r_t$ is the sample mean of the log-return time series, and $X_t = g(Z_t) = 0$ and $X_t^2 = \sigma_X^2$, whenever $t \leq 0$, where σ_X^2 is the sample variance of $\{X_t\}_{t=1}^n$.

Since we intent to compare the performance of the SFIEGARCH model with other ARCH-type models which incorporate or not seasonal long-memory cyclical behavior in the volatility, we shall consider a two step estimation procedure. First an ARMA(p_1, q_1) model for the one-hour log-return time series is selected. The second step consists on fitting an SFIEGARCH$(p_2, d, q_2)_s$ (or any other ARCH-type) model to the residuals of the ARMA model.

The information obtained through the analysis of the autocorrelation function of $\{r_t\}_{t=1}^{4232}$ is applied to select the orders p_1 and q_1. Given the large number of possible combinations of p_1 and

q_1 we restrict our attention to ARMA$(263, 0)$ and ARMA$(0, 263)$ models with $\phi_i = \theta_j = 0$, whenever $i, j \notin \{7, 8, 13, 18, 53, 60, 109, 119, 139, 190, 191, 203, 218, 263\}$.

In order to select an SFIEGARCH$(p_2, d, q_2)_s$ model for the residuals we fix $s = 7$ (estimated from the periodogram and sample autocorrelation functions) and consider different combinations of p_2 and q_2. Once the right combination of p_2 and q_2 is found, the same values are consider to select the FIEGARCH and the EGARCH models.

The following criteria applies to both estimation steps.

1. For any combination of p_1, q_1 or p_2, q_2, we start with the full model and remove the non-significant parameters (one at a time) until all p-values are smaller than 0.05.

2. The standard deviations for the model parameters were obtained by considering the robust covariance matrix given by $n^{-1}H^{-1}(\hat{\boldsymbol{\eta}})B(\hat{\boldsymbol{\eta}})H^{-1}(\hat{\boldsymbol{\eta}})$, where $n^{-1}H(\hat{\boldsymbol{\eta}})$ and $n^{-1}B(\hat{\boldsymbol{\eta}})$ are, respectively, the Hessian and the outer product of the gradients (see Bollerslev and Wooldridge, 1992).

3. A model is considered to fit the data well if $\{\hat{x}_t\}_{t=1}^n$ (the residual of the ARMA model), $\{\hat{z}_t\}_{t=1}^n$ (the residual of the SFIEGARCH model) and $\{\hat{z}_t^2\}_{t=1}^n$ show no significant correlation. To test for correlation we consider both the Box-Pierce and Ljung-Box hypothesis tests (see Remark 5.6).

4. When more than one model satisfy the criteria in Step 3, model selection is performed based on the values of the log-likelihood, AIC, BIC and HQC criteria, obtained in Step 2.

5. In case two or more models present similar AIC, BIC, HQC and/or log-likelihood criteria values, we chose the more parsimonious one.

Remark 5.6. When applying the Box-Pierce (or the Ljung-Box) hypothesis test, if the null hypothesis is rejected for $\{\hat{x}_t\}_{t=1}^n$ but it is not reject for both $\{\hat{z}_t\}_{t=1}^n$ and $\{\hat{z}_t^2\}_{t=1}^n$, the cumulative periodogram (also known as Kolmogorov-Smirnov hypothesis test) is considered. If this test does not reject the null hypothesis that $\{\hat{x}_t\}_{t=1}^n$ is a white noise process, the model is not discarded.

Further residuals analysis is performed by following the same approach as Haas et al. (2004). The procedure consists on employing a density transformation, as presented in Diebold et al. (1998), to test the assumption $\hat{x}_t|\mathcal{F}_{t-1} \sim F_t$, for some given target distribution F_t.

5.2 Forecasting Procedure

Once the parameters of the ARMA(p_1, q_1)-SFIEGARCH$(p_2, d, q_2)_s$ model are estimated, out-of-sample forecasting is performed. To obtain the predicted values $\hat{r}_{n+h}$, $\hat{\sigma}_{n+h}^2$ and $\hat{x}_{n+h}^2$, given $\{r_t\}_{t=1}^n$, with $n = 4232$ and $h \in \{1, \cdots, 270\}$, we proceed as described in steps **1** - **9** below. We shall denote by $\boldsymbol{\eta}$ the true parameters, namely,

$$\boldsymbol{\eta} = (d, \theta, \gamma, \omega, \mu, \phi_1, \cdots, \phi_{p_1}, \varphi_1, \cdots, \varphi_{q_1}, \alpha_1, \cdots, \alpha_{p_2}, \beta_{q_1}, \cdots, \beta_{q_2})',$$

and by $\hat{\boldsymbol{\eta}}$ the estimated values. With obvious identifications, the forecasting considering the other ARCH-type models is analogous.

1. The true parameters values $(d, \alpha_1, \cdots, \alpha_{p_2}, \beta_1, \cdots, \beta_{q_2})'$ are replaced by the estimated ones, namely, $(\hat{d}, \hat{\alpha}_1, \cdots, \hat{\alpha}_{p_2}, \hat{\beta}_1, \cdots, \hat{\beta}_{q_2})'$, and the recurrence formula given in Proposition 1.1 is used to calculate the corresponding coefficients $\{\hat{\lambda}_{d,k}\}_{k=0}^{n+270}$. Notice that, with obvious identifications, this recurrence formula can be also used to calculate the coefficients $\{\hat{\psi}_{d,k}\}_{k=0}^{n+270}$ associated to $\hat{\phi}_1, \cdots, \hat{\phi}_{p_1}, \hat{\varphi}_1, \cdots, \hat{\varphi}_{q_1}$.

2. Upon setting $r_t = \hat{\mu}$ and $\hat{x}_t = g(\hat{z}_t) = 0$, whenever $t < 0$, the time series $\{\hat{x}_t\}_{t=1}^n$, $\{\hat{\sigma}_t\}_{t=1}^n$ and $\{\hat{z}_t\}_{t=1}^n$ are obtained recursively as follows[5]

$$\hat{x}_t = \sum_{k=0}^{p_1} \phi_k(r_{t-k} - \hat{\mu}) + \sum_{j=1}^{q_1} \varphi_j \hat{x}_{t-j},$$

$$\hat{\sigma}_t = \exp\left\{\frac{\hat{\omega}}{2} + \frac{1}{2}\sum_{k=0}^{n-1} \hat{\lambda}_{d,k}\left[\hat{\theta} z_{t-1-k} + \hat{\gamma}\left(|\hat{z}_{t-1-k}| - \sqrt{2/\pi}\right)\right]\right\} \quad \text{and} \quad \hat{z}_t = \frac{\hat{x}_t}{\hat{\sigma}_t},$$

for all $t \in \{1, \cdots, n\}$. Note that, in particular, $\hat{x}_1 = r_1 - \hat{\mu}$, $\hat{\sigma}_1 = e^{\hat{\omega}0.5}$ and $\hat{z}_1 = \hat{x}_1 \hat{\sigma}_1^{-1}$.

3. Since the h-step ahead predictor for X_{n+h} given $\mathcal{F}_n$ is zero, it is set $\hat{x}_{n+h} = 0$, for all $h > 0$.

4. The h-step ahead forecast $\hat{r}_{n+h}$ for r_{n+h} is given by (see, for instance Brockwell and Davis, 1991)

$$\hat{r}_{n+h} = \hat{\mu} + \sum_{k=1}^{p_1} \phi_k(\hat{r}_{n+h-k} - \hat{\mu}) + \sum_{j=1}^{q_1} \varphi_j \hat{x}_{n+h-j}, \quad \text{with } \hat{r}_t = r_t, \text{ if } t \leq n. \tag{5.3}$$

5. An estimate $\hat{\sigma}_g^2$ for σ_g^2 is obtained by replacing $\mathbb{E}(|Z_0|)$ and $\mathbb{E}(Z_0|Z_0|)$, in expression (1.14), by their respective sample estimates, that is,

$$\hat{\sigma}_g^2 = \hat{\theta}^2 + \hat{\gamma}^2 - (\hat{\gamma}\hat{\mu}_{|z|})^2 + 2\hat{\theta}\hat{\gamma}\left[\frac{1}{n}\sum_{t=1}^n \hat{z}_t|\hat{z}_t|\right], \quad \text{with} \quad \hat{\mu}_{|z|} := \frac{1}{n}\sum_{t=1}^n |\hat{z}_t|.$$

6. An estimative for $E_{\boldsymbol{\eta}}(\ell) := \mathbb{E}(\exp\{\lambda_{d,\ell}g(Z_0)\})$, given in (4.1), is obtained by considering the respective sample estimator

$$\hat{E}_{\hat{\boldsymbol{\eta}}}(\ell) = \frac{1}{n}\sum_{t=1}^n \exp\left\{\hat{\lambda}_{d,\ell}\left[\hat{\theta}z_t + \hat{\gamma}(|\hat{z}_t| - \hat{\mu}_{|z|})\right]\right\}, \quad \text{for any} \quad \ell \in \{0, \cdots, h-2\}.$$

7. Since σ_{n+1}^2 is $\mathcal{F}_n$-measurable, $\hat{\sigma}_{n+1}^2 = \sigma_{n+1}^2$ and it is computed as in step **2**.

8. The predictor $\check{\sigma}_{n+h}^2$ is obtained upon replacing the true parameter values by the estimated ones in (4.4), with the additional assumption $g(\hat{z}_t) = 0$, if $t < 0$. Then, from (4.6), $\hat{\sigma}_{n+h}^2$ and $\tilde{\sigma}_{n+h}^2$ are obtained by setting, respectively,

$$\hat{\sigma}_{n+h}^2 = \check{\sigma}_{n+h}^2 \prod_{\ell=0}^{h-2} \hat{E}_{\hat{\boldsymbol{\eta}}}(\ell) \quad \text{and} \quad \tilde{\sigma}_{n+h}^2 = \tilde{\sigma}_{n+h}^2\left[1 + \frac{1}{2}\hat{\sigma}_g^2\sum_{k=0}^{h-2} \hat{\lambda}_{d,k}^2\right]^{-1} \prod_{\ell=0}^{h-2} \hat{E}_{\hat{\boldsymbol{\eta}}}(\ell), \quad \text{for any} \quad h > 2.$$

9. The predictor $\hat{r}_{n+h}^2$ is obtained through (4.8), with the additional assumption $\hat{x}_t = 0$, if $t < 0$.

Remark 5.7. In the literature, the time series $\{\hat{\mu}_t\}_{t=1}^n$, given by,

$$\hat{\mu}_t = \sum_{k=1}^{p_1} \phi_k(r_{t-k} - \hat{\mu}) + \sum_{j=1}^{q_1} \varphi_j \hat{x}_{t-j}, \quad \text{for } t \in \{1, \cdots, n\},$$

is called *fitted values* or *in-sample forecasts*[6]. Consequently, the residuals time series $\{\hat{x}_t\}_{t=1}^n$ is also denoted *in-sample errors of forecast*. Furthermore, since $\hat{z}_t = \hat{x}_t\hat{\sigma}_t^{-1}$, for all $t \in \{1, \cdots, n\}$, the time series $\{\hat{z}_t\}_{t=1}^n$ is often called *standardized residuals*.

---

[5]If the pseudo-likelihood is used instead of the quasi-likelihood, the value $\sqrt{2/\pi}$ must be replace by the value of $\mathbb{E}(|Z_0|)$ associated to the distribution considered in the estimation procedure.

[6] From (5.3) it is clear that $\hat{\mu}_t$ is the 1-step ahead forecast for r_t, given $\mathcal{F}_{t-1}$, for any $t > 0$.

5.3   Forecasting Performance and Models Comparison

Without loss of generality, let $\{\hat{y}_t\}_{t=t_0}^{t_0-1+n_p}$ denote either the in-sample ($t_0 = 1$ and $n_p = 4232$) or the out-of-sample ($t_0 = 4233$ and $n_p = 270$) forecast values corresponding to the time series $\{y_t\}_{t=t_0}^{t_0-1+n_p}$, where y_t is either r_t or r_t^2 and n_p is the number of predicted values. Denote by $\mathcal{M}$ any model used to obtain $\{\hat{y}_t\}_{t=t_0}^{t_0-1+n_p}$. The forecasting performance of model $\mathcal{M}$ is evaluated by computing the mean absolute error (mae), the mean percentage error (mpe) and the maximum absolute error (max_{ae}) measures, namely,

$$mae(\mathcal{M}) = \frac{1}{n_p}\sum_{t=1}^{n_p}|e_t|, \quad mpe(\mathcal{M}) := \frac{1}{n_p}\sum_{t=1}^{n_p}\frac{|e_t|}{|y_t|} \quad \text{and} \quad max_{ae}(\mathcal{M}) := \max_{t \in \{1,\cdots,n_p\}}\{|e_t|\}$$

where $e_t := y_t - \hat{y}_t$ denotes the forecasting error at step t. The statistical significance of the out-of-sample forecasting performance is evaluated by using the so called Diebold and Mariano hypothesis test (see Diebold and Mariano, 1995).

Remark 5.8. The mpe is an interesting measure since it considers not only the magnitude of the error (as does the mae) but also the proportion between the error and the true values so it is easier to decide whether the error is small or not. A drawback of the mpe is that this measure is highly affected when observations are too close to zero.

The predictive performance of model $\mathcal{M}$ is also evaluated by measuring the quality of the one-step ahead density forecasts (see, for instance, Paolella, 2013). The measure used for this analysis is the normalized sum of the realized predictive log-likelihood, given by

$$S(\mathcal{M}) = \frac{1}{n_p}\sum_{t=n+1}^{n+n_p}\ln\big(f_{t|\mathcal{F}_{t-1}}^{\mathcal{M}}(y_t; \hat{\boldsymbol{\eta}})\big),$$

where $\{y_t\}_{t=1}^{n+n_p}$ is a sample from $\{Y_t\}_{t\in\mathbb{Z}}$, n is the size of the sample used to estimate the parameters for model $\mathcal{M}$, n_p is the number of predicted values, $f_{t|\mathcal{F}_{t-1}}^{\mathcal{M}}(\cdot; \hat{\boldsymbol{\eta}})$ denotes the conditional probability density function of Y_t given $\mathcal{F}_{t-1}$ and $\boldsymbol{\eta}$ is the parameter vector for model $\mathcal{M}$.

Forecast efficiency regressions (see Mincer and Zarnowitz, 1969) are also used to compare the quality of the volatility forecasts among the different models fitted to the data. The standard Mincer-Zarnowitz regression for forecast efficiency is given by

$$y_{t+h} = \gamma_0 + \gamma_1\hat{y}_{t+h} + \varepsilon_t, \tag{5.4}$$

where y_{t+h} is the variable of interest and $\hat{y}_{t+h}$ is the h-step ahead forecast for y_{t+h} given $\mathcal{F}_t$. Under the null hypothesis of forecast efficiency $\gamma_0 = 0$ and $\gamma_1 = 1$. The coefficients γ_0 and γ_1 in (5.4) are obtained by ordinary least square (OLS) estimation. The standard errors of the estimates are corrected for heteroskedasticity and autocorrelation by using the HAC estimator (see Newey and West, 1987). Since the forecasts $\hat{y}_{n+h}$ are obtained from a model $\mathcal{M}$ for which the true parameter values are unknown, the uncertainty concerning parameter estimation is corrected by multiplying the Newey-West standard errors by $\lambda = \sqrt{1 + n_p/n}$ (see West and McCraken, 1998), where n is the sample size used to fit the model and n_p is the number of predicted values.

Since the true volatility cannot be directly measured, the forecast efficiency regression is performed by considering the realized volatility instead. The ideas for this analysis were adapted from Klaassen (2002), where a slightly different definition[7] for the "observed volatility" was considered. The definitions adopted here are given below.

---

[7]Klaassen (2002) considers daily log-returns and the models fitted to the data do not include the ARMA regression. So, r_t was written as $r_t = \mu + X_t$, where X_t follows an ARCH-type model. In this case, the author replaced $\bar{y}_t$ by $\hat{\mu}$ in the definition of v_t. In our case, μ is replaced by μ_t, which may vary according to each ARCH-type model associated to X_t. Therefore we shall use the traditional definition of realized volatility to let v_t be model free.

Definition 5.1. Let $r_{(t-1)M+k}$ be the log-return value corresponding to the k-th period of day t, for $k \in \{1, \cdots, M\}$ and $t = 1, \cdots, N$, where M is the number of intraday periods and N is the number of observed days.

a) The daily log-return, denoted by $r_t^{(d)}$, is defined through $r_t^{(d)} = \sum_{k=1}^{M} r_{(t-1)M+k}$, for all $t \in \{1, \cdots, N\}$.

b) The daily realized volatility, denoted by v_t, is given by

$$v_t = \sum_{k=1}^{M} (r_{(t-1)M+k} - \bar{r}_t)^2, \quad \text{where} \quad \bar{r}_t := \frac{1}{M} \sum_{j=1}^{M} r_{(t-1)M+j}, \quad \text{for all } t \in \{1, \cdots, N\}.$$

c) The log-return over the period of h days, denoted by $r_{t-1}^{(d)}[h]$, is given by

$$r_{t-1}^{(d)}[h] = \sum_{j=0}^{h-1} r_{t+j}^{(d)} = \sum_{j=0}^{h-1} \left[\sum_{k=1}^{M} r_{(t+j-1)M+k} \right] = \sum_{k=1}^{hM} r_{(t-1)M+k}, \quad \text{for any } h \geq 1 \text{ and } t \in \{1, \cdots, N\}.$$

In particular, $r_{t-1}^{(d)}[1] = r_t^{(d)}$ so the log-return of period 1-day is simply the daily log-return.

d) The realized volatility over the period of h days, denoted by $v_{t-1}[h]$, is given by

$$v_{t-1}[h] = \sum_{j=0}^{h-1} \left[\sum_{k=1}^{M} (r_{(t+j-1)M+k} - \bar{r}_{t+j})^2 \right] = \sum_{j=0}^{h-1} v_{t+j}, \quad \text{for any } h \geq 1 \text{ and } t \in \{1, \cdots, N\}.$$

In particular, $v_{t-1}[h] = v_t$ so the realized volatility over the period of 1-day is simply the daily volatility.

By following the same steps as in the proof of proposition 2.2 in Prass and Lopes (2013), one can show that, if $\{r_t\}_{t \in \mathbb{Z}}$ follows an ARMA(p_1, q_1)-SFIEGARCH$(p_2, d, q_2)_s$ model then the forecast for the log-return over the period of h days, given the information up to day[8] $t-1$, is given by

$$\hat{r}_{t-1}^{(d)}[h] = \sum_{j=0}^{h-1} \hat{r}_{t+j}^{(d)} = \sum_{k=1}^{hM} \hat{r}_{(t-1)M+k}, \quad \text{for any } h \geq 1 \text{ and } t \in \{1, \cdots, N\}. \tag{5.5}$$

where $\hat{r}_{(t-1)M+k} = \mathbb{E}(r_{(t-1)M+k}|\mathcal{F}_{(t-1)M})$ is the k-step ahead forecast for $r_{(t-1)M+k}$, obtained from the ARMA model. Moreover, the forecast for the conditional variance of log-return over the period of h, given the information up to day $t-1$, namely $\sigma_{t-1}^{2\,(d)}[h] = \text{Var}(r_{t-1}^{(d)}[h]|\mathcal{F}_{(t-1)M})$, is given by

$$\hat{\sigma}_{t-1}^{2\,(d)}[h] = \sum_{j=1}^{hM} \Psi_j \hat{\sigma}_{(t-1)M+j}^2, \quad \text{with } \Psi_j := \left[\sum_{k=0}^{hM-j} \psi_k \right]^2, \tag{5.6}$$

where $\hat{\sigma}_{(t-1)M+k}^2 = \mathbb{E}(X_{(t-1)M+k}^2|\mathcal{F}_{(t-1)M})$ is the k-step ahead forecast for $X_{(t-1)M+k}^2$, obtained from the SFIEGARCH model. Equivalent result is derived upon replacing the SFIEGARCH by any other ARCH-type model.

Remark 5.9. Notice that Definition 5.1 and expressions (5.5) and (5.6) assume M constant. When M varies over time, similar results are derived upon making the following adjustments: replace $(t-1)M$ by the number of intraday log-returns available up to day $(t-1)$ (included); replace hM by the number of intraday log-returns over the period from t to $t+h$; replace $\sum_{k=1}^{M}$ and $1/M$, respectively, by $\sum_{k=1}^{M_t}$ (in Definition 5.1 c) and d), by $\sum_{k=1}^{M_{t+j}}$) and $1/M_t$, where M_t is the number of intraday log-returns available for day t.

Remark 5.10. For any $h \geq 1$ fixed, the forecast efficiency regression is obtained upon replacing, in (5.4), y_{t+1} and $\hat{y}_{t+1}$, respectively, by $v_t[h]$ and $\hat{\sigma}_t^{2(d)}[h]$, for $t = n, \cdots, N-h$, where N is the sample size of the log-return time series and n is the size of the sample used to fit the model.

---

[8]Note that, since we are considering intraday log-returns, the information up to day $t-1$ corresponds to $\mathcal{F}_{(t-1)M}$, where M is number of the intraday periods.

5.4   Results

Constrained ARMA(263, 0) and ARMA(0, 263) models, with $\phi_i = \theta_j = 0$, whenever $i, j \notin \{7, 8, 13, 18, 53, 60, 109, 119, 139, 190, 191, 203, 218, 263\}$, where analyzed. It turns out that several initial considered parameters where not significant and were removed from the models. The most parsimonious found model is given by (the number in parenthesis is the robust standard error)

$$r_t = X_t - \underset{(0.0128)}{0.0391} X_{t-7} - \underset{(0.0169)}{0.0834} X_{t-13} + \underset{(0.0204)}{0.0803} X_{t-53} + \underset{(0.0215)}{0.0748} X_{t-109} - \underset{(0.0093)}{0.0433} X_{t-190}$$
$$+ \underset{(0.0221)}{0.0742} X_{t-203} + \underset{(0.0167)}{0.0680} X_{t-218} - \underset{(0.0111)}{0.0421} X_{t-263}$$
$$:= X_t + \mu_t,$$

for all $t \in \{1, \cdots, 4232\}$, with $X_t = 0$, if $t \leq 0$.

The observed time series $\{r_t\}_{t=1}^{4232}$ and the corresponding fitted values $\{\hat{\mu}_t\}_{t=1}^n$, obtained from the ARMA model, are given in Figure 5.4 (a). Figure 5.4 (b) shows the residuals time series $\{\hat{x}_t\}_{t=1}^{4232}$. The p-values for the Box-Pierce and Ljung-Box test statistic for $\{\hat{x}_t\}_{t=1}^n$, that is, the residuals of the ARMA model, were smaller than 0.05 for any lag higher than 15. On the other hand, upon applying the cumulative periodogram test, the null hypothesis that $\{\hat{x}_t\}_{t=1}^n$ is white noise process, was not rejected (the cumulative periodogram figure was omitted to save space and may be obtained from the authors upon request). As expected, both the Box-Pierce (or Ljung-box) and the cumulative periodogram tests reject the null hypothesis that $\{\hat{x}_t^2\}_{t=1}^n$ is a white noise process.

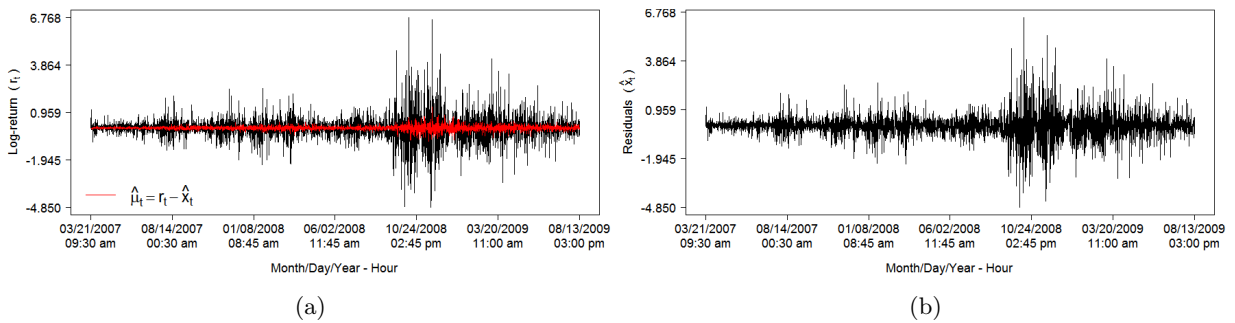(a)   (b)

Figure 5.4: This figure shows the S&P500 US stock index log-return time series observed in the period from from March 21, 2007 at 9:30 am to August 13, 2009 at 03:00 pm and the corresponding fitted values and residuals obtained form the constrained ARMA model. Panel (a) gives the observed time series $\{r_t\}_{t=1}^{4232}$ (in black) and the fitted values $\{\hat{\mu}_t\}_{t=1}^{4232}$ (in red). Panel (b) shows the residuals time series $\{\hat{x}_t\}_{t=1}^{4232}$.

The estimated values and the corresponding standard errors for the parameters of considered ARCH-type models are given in Table 5.2. For any model in Table 5.2, the p-values for both the Box-Pierce and Ljung-Box test statistics corresponding to the time series $\{\hat{z}_t\}_{t=1}^{4232}$, that is, the residual of the ARCH-type model, were always higher than 0.05, for any lag $h > 0$. The same applies to the time series $\{\hat{z}_t^2\}_{t=1}^{4232}$. The histogram (or the kernel density) and the QQ-Plot (both omitted here to save space but available from the authors upon request) indicate that, although symmetric, the distribution of $\{\hat{z}_t\}_{t=1}^{4232}$ is not Gaussian.

Remark 5.11. The SFIEGARCH model given in Table 5.2 is the most parsimonious model such that $\beta(\cdot)$ has no roots in the closed disk $\{z : |z| \leq 1\}$. The p-value for α_3, in the FIEGARCH model, is 0.1141. On the other hand, the model fitted without this coefficient does not lead to uncorrelated residuals. The polynomial $\beta(\cdot)$ associated to the FIEGARCH model has two roots with absolute value 1.0023 and two other roots with absolute value 1.0024. Therefore, very close to the unit circle. Analogously, the polynomial $\beta(\cdot)$ associated to the EGARCH model has two roots with absolute value 1.0024 and another one with absolute value 1.0001. Despite this fact, the EGARCH model presents sligltly better performance in terms of log-likelihood, AIC, BIC and HQC criteria.

Table 5.2: Estimated values and the corresponding robust standard errors (in parenthesis) for the parameters of the SFIEGARCH, FIEGARCH and EGARCH models fitted to the S&P500 US stock index log-return time series observed in the period from March 21, 2007 at 9:30 am to August 13, 2009 at 03:00 pm. This table also presents the corresponding log-likelihood, AIC, BIC and HQC criteria values.

Parameter	SFIEGARCH	FIEGARCH	EGARCH
d	0.4532 (0.0104)	0.4529 (0.0057)	-
ω	-1.1810 (0.0129)	-1.6972 (0.0049)	-0.8601 (0.0058)
θ	-0.0820 (0.0115)	-0.1100 (0.0017)	-0.0954 (0.0016)
γ	0.2127 (0.0213)	0.2393 (0.0053)	0.2197 (0.0025)
α_1	0.1655 (0.0040)	0.0657 (0.0065)	-0.2718 (0.0015)
α_2	0.1963 (0.0263)	0.3021 (0.0088)	0.3761 (0.0019)
α_3	0.1821 (0.0045)	0.0151 (0.0095)	-0.2503 (0.0021)
α_4	-0.3095 (0.0048)	-0.4206 (0.0079)	-0.4603 (0.0039)
α_5	-0.4115 (0.0094)	-0.5962 (0.0136)	-0.5355 (0.0038)
α_6	-0.3139 (0.0112)	0.4580 (0.0023)	0.9040 (0.0019)
β_1	0.3151 (0.0095)	0.1378 (0.0039)	0.3271 (0.0013)
β_2	-0.0668 (0.0174)	0.1231 (0.0025)	0.0779 (0.0016)
β_3	0.2092 (0.0198)	-0.0081 (0.0030)	0.1545 (0.0025)
β_4	-0.3621 (0.0079)	-0.2854 (0.0042)	-0.4983 (0.0014)
β_5	0.0785 (0.0124)	-0.1842 (0.0020)	0.1895 (0.0023)
β_6	0.6534 (0.0115)	0.8896 (0.0029)	0.7389 (0.0026)
log-likelihood	-3053.9066	-2936.1820	-2878.2630
AIC	6139.8131	5904.3641	5786.5260
BIC	6241.4201	6005.9709	5881.7824
HQC	6175.7272	5940.2781	5820.1954

To apply the density transform procedure (for details, see Haas et al., 2004; Diebold et al., 1998) the GED distribution, with different values for the tail-thickness parameter ν. Under this scenario, the null hypothesis to be tested is

$$H_0 : \hat{x}_t | \mathcal{F}_{t-1} \sim \text{GED}(\nu, 0, \sigma_t) \quad \text{or, equivalently,} \quad H_0 : F_t^{-1}(\hat{x}_t) \sim \mathcal{U}(0, 1),$$

where $\text{GED}(\nu, 0, \sigma_t)$ denotes the generalized error distribution with tail-thickness parameter ν, mean zero and standard deviation σ_t, and $F_t(\cdot)$ is the corresponding cumulative distribution function. In particular, when $\nu = 2$, we have the Gaussin distribution. The time series $\{x_t\}_{t=1}^{4232}$ corresponds to the residuals of the ARMA model fitted to the one-hour log-return time series $\{r_t\}_{t=1}^{4232}$ and $\{\hat{\sigma}_t^2\}_{t=1}^{4232}$ denotes the conditional variance of the log-returns, obtained from the SFIEGARCH, FIEGARCH or from the EGARCH model. Table 5.3 reports the results for the Kolmogorov-Smirnov (K-S) hypothesis test used to compare the sample $\{F^{-1}(\hat{x}_t)\}_{t=1}^{4232}$ with the uniform distribution.

Table 5.3 confirms the results obtained with the QQ-Plot, that is, $\{\hat{z}_t\}_{t=1}^n$ does have Gaussian distribution. This table also indicates that the assumption that $\{\hat{z}_t\}_{t=1}^n$ follows a $\text{GED}(\nu)$ distribution holds for more than one value of ν. The next step in this analysis would be to replace the QMLE by the log-likelihood estimation procedure, using the GED distribution, and estimate ν alongside with the other parameters. The information on the ν parameter could then be incorporated in the forecasting equation to see whether forecast efficiency improves or not. This analysis shall be performed in a future work.

The mean absolute error (mae), the mean percentage error (mpe) and the maximum absolute error (max_{ae}) measures for the selected models are reported in Tables 5.4 and 5.5. For the in-sample analysis, the mae, mpe and max_{ae} values were obtained by letting $e_t = r_t - \hat{\mu}_t$, for all $t \in \{1, \cdots, 4232\}$ (see Section 5.3). For the out-of-sample comparison, we consider not only the forecasts for r_{t+h} but also for r_{t+h}^2. The out-of-sample mae, mpe and max_{ae} values were obtained

Table 5.3: Results for the Kolmogorov-Smirnov (K-S) hypothesis test used to compare the sample $\{F_t^{-1}(\hat{x}_t)\}_{t=1}^{4232}$ with the uniform distribution. The values reported are the p-value for the K-S test statistic. The null hypothesis considered is $H_0 : \hat{x}_t|\mathcal{F}_{t-1} \sim \text{GED}(\nu, 0, \sigma_t)$, for different values of ν. The time series $\{x_t\}_{t=1}^{4232}$ corresponds to the residuals of the ARMA model fitted to the one-hour log-return time series $\{r_t\}_{t=1}^{4232}$ and $\{\hat{\sigma}_t^2\}_{t=1}^{4232}$ denotes the conditional variance of the log-returns, obtained from the SFIEGARCH, FIEGARCH or from the EGARCH model.

ν	SFIEGARCH	FIEGARCH	EGARCH
1.45	0.56	0.03	0.25
1.50	0.45	0.09	0.19
1.55	0.36	0.21	0.14
1.70	0.11	0.28	0.04
2.00	0.00	0.00	0.00

by letting $e_{t+h} = r_{t+h} - \hat{r}_{t+h}$, for $t = 2432$ (fixed) and $h = 1, \cdots, 270$; by setting $h = 1$ fixed and letting $e_{t+1} = r_{t+1} - \hat{r}_{t+1}$ for $t = 4232, \cdots, 4501$; by letting $e_{t+h} = r_{t+h}^2 - \hat{r}_{t+h}^2$, for $t = 2432$ (fixed) and $h = 1, \cdots, 270$; and also by letting $h = 1$ fixed and considering $e_{t+1} = r_{t+1}^2 - \hat{r}_{t+1}^2$, for $t = 4232, \cdots, 4501$. For any ARCH-type model, $\hat{r}_{t+h}^2$, for any $h \geq 1$ and $t \in \mathbb{Z}$, was obtained according to Theorem 4.3. While Table 5.4 reports the *mae*, *mpe* and *max_{ae}* associated to the in-sample and out-of-sample forecasts for r_{t+h}, Table 5.5 gives the values associated to r_{t+h}^2. Notice that, since the ARMA model was selected independently of the ARCH-type models, all values in Table 5.4 do not depend on the model for the conditional variance σ_t^2.

Table 5.4: The mean absolute error (*mae*), the mean percentage error (*mpe*) and the maximum absolute error (*max_{ae}*) measures for the out-of-sample forecasts for r_{t+h}. Case 1 are the values corresponding to the in-sample forecasts, with $e_t = r_t - \hat{\mu}_t$, for $t \in \{1, \cdots, 4232\}$. Case 2 and 3 correspond to the out-of-sample forecasts. For Case 2, $e_{t+h} = r_{t+h} - \hat{r}_{t+h}$, for $t = 4232$ and $h \in \{1, \cdots, 270\}$. Case 3 assumes $e_{t+1} = r_{t+1} - \hat{r}_{t+1}$, for $t \in \{4232, \cdots, 4501\}$.

Measure	Case 1	Case 2	Case 3
mae	0.4309	0.2611	0.1525
mape	1.6291	1.3480	1.7712
max_{ae}	6.4438	1.7892	0.9703

Table 5.5: The mean absolute error (*mae*), the mean percentage error (*mpe*) and the maximum absolute error (*max_{ae}*) measures for the out-of-sample forecasts for r_{t+h}^2. For each model the left column considers $e_t = r_{t+h}^2 - \hat{r}_{t+h}^2$, for $t = 4232$ and $h \in \{1, \cdots, 270\}$, and the right column assumes $e_{t+1} = r_{t+1}^2 - \hat{r}_{t+1}^2$, for $t \in \{4232, \cdots, 4501\}$.

Measure	SFIEGARCH		FIEGARCH		EGARCH	
mae	0.2845	0.1865	0.2829	0.1724	0.2272	0.1687
mape	970.1824	531.1481	684.6831	409.1600	417.9443	372.8434
max_{ae}	2.9271	2.7457	2.9426	2.9844	2.9597	3.0164

Table 5.5 indicates that the SFIEGARCH model presents the best performance, among all models, only in terms of *max_{ae}*. Figures 5.5 and 5.6 show, respectively, the observed values r_{t+h}, r_{t+h}^2 and the fitted ones, obtained from the ARCH-type models, for $h = 1, \cdots, 270$. These figures help to explain the reason why the SFIEGARCH model have higher *mae* and *mape* than the other ARCH-type models.

From Figure 5.5 it is clear that r_{t+h}^2 is very close to zero, for several values of $h \in \{1, \cdots, 270\}$. From Figure 5.6 (first row, from top to bottom) it is clear that the h-step ahead forecasts obtained from the SFIEGARCH model converge to a fixed value, which is expected since $\hat{\sigma}_{t+h}^2$ converges to the unconditional variance as h goes to infinity. From Figure 5.6 (first row, from top to bottom) is also evident that the forecasts for the FIEGARCH and EGARCH models do not converge at all.

(a)   (b)
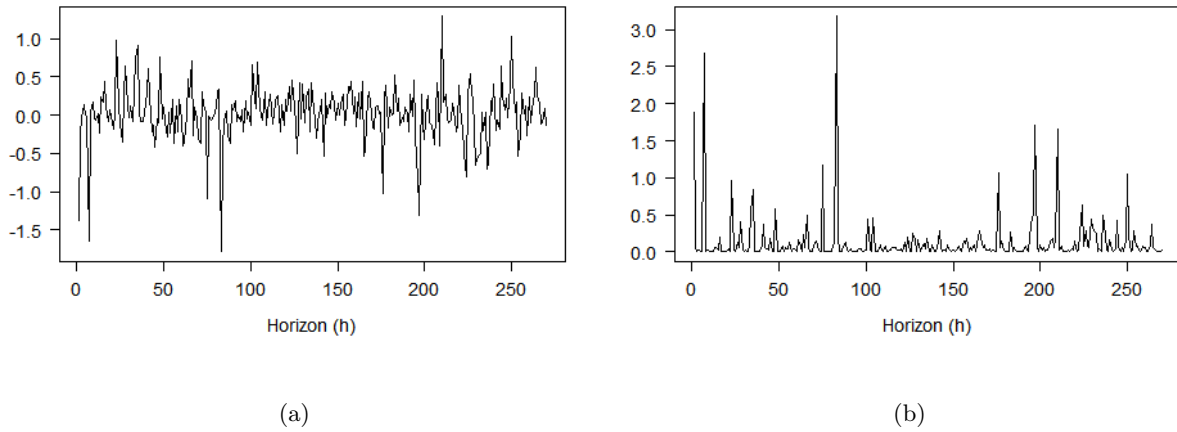
Figure 5.5: Sample from the one-hour log-return time series corresponding to the period from August 14, 2009 at 09:15 am to October 8, 2009 at 03:00 pm. Panel (a) gives the log-returns time series $\{r_{t+h}\}_{h=1}^{270}$. Panel (b) gives the squared log-returns $\{r_{t+h}^2\}_{h=1}^{270}$. The time index $t = 4232$ corresponds to August 13, 2009 at 03:00 pm.

(a)   (b)   (c)

(d)   (e)   (f)
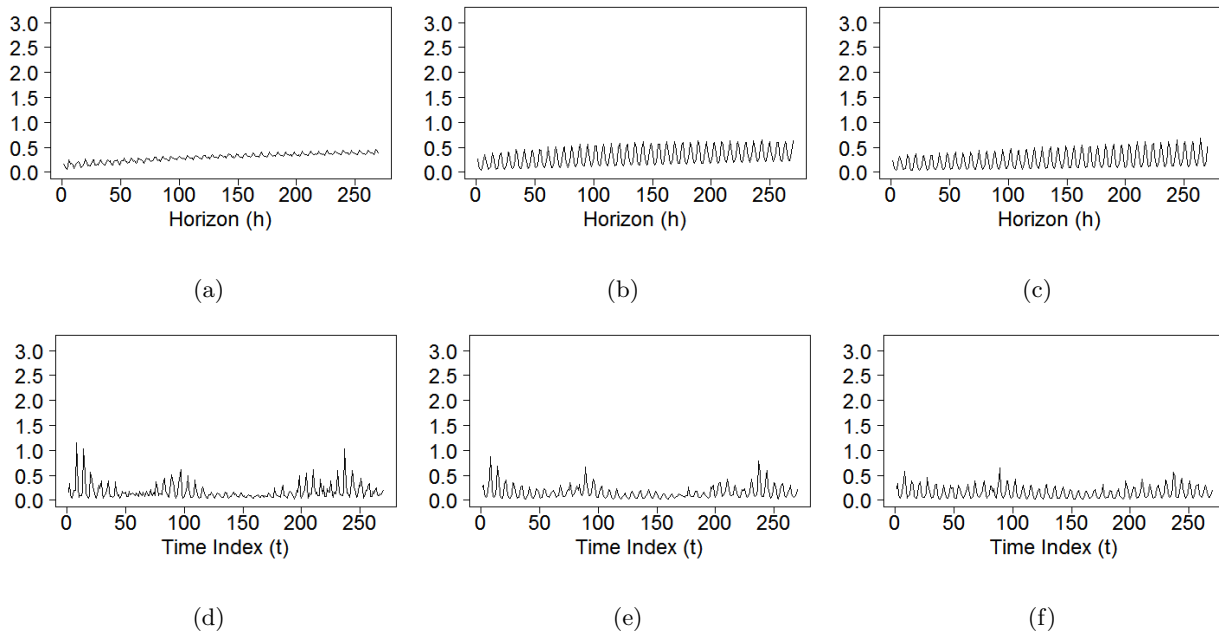
Figure 5.6: Out-of-sample forecasts for the sample from the one-hour log-return time series corresponding to the period from August 14, 2009 at 09:15 am to October 8, 2009 at 03:00 pm. Panels (a), (b) and (c) give the h-step ahead forecasts $\{\hat{r}_{t+h}^2\}_{h=1}^{270}$, with forecasting origin $t = 2432$, for the SFIEGARCH, FIEGARCH and EGARCH models, respectively. Panels (d), (e) and (b) give the one-step ahead forecast $\{\hat{r}_{t+1}^2\}_{t=4232}^{4501}$, obtained by considering, respectively, the SFIEGARCH, FIEGARCH and EGARCH models fitted to the sample $\{r_t\}_{t=1}^{4232}$.

On the contrary, the forecast values for these two models seem to oscillate around a curve which converges to the same value as the forecasts from the SFIEGARCH model. Moreover, the amplitude of these oscillations increases over time. This should be expected since both the FIEGARCH and the EGARCH models are such that the polynomial $\beta(\cdot)$ has at least one unit root (or a root close enough to the unit root).

Figure 5.6 (second row, from top to bottom) also shows that the one-step ahead forecasts for the SFIEGARCH model were able to capture the peaks in the square log-returns much better than the other two models. Although this is a good feature of the SFIEGARCH model it also makes the *mae* and the *mape* values increase since the forecast values tend to increase in the region around the

peaks while the observed time series shows several values close to zero in the same region. Also, notice that, since the forecasts for r_{t+h}^2 are always positive, it is evident that a model which oscilates as the FIEGARCH and the EGARCH do will provide h-step ahead forecasts close to zero more often than a model for which the forecast value converges to a non-zero constant.

Remark 5.12. We apply the Diebold and Mariano hypothesis test (see Diebold and Mariano, 1995) to verify the statistical significance of the out-of-sample forecasting performance. We consider the absolute error as loss function so the loss-differential series is given by $\{d_t\}$, with $d_t = |y_{t_0+t+1} - \hat{y}_{t_0+t+1}| - |y_{t_o+t+1}|$, for $t_0 = 2432$ and $t = 0, \cdots, 269$. The variable y_{t_0+t+1} denotes either the log-returns r_{t_0+t+1} or the squared log-returns $r_{t_0+t+1}^2$ and $\hat{y}_{t_0+t+1}$ are the corresponding one-step ahead forecasts obtained from model $\mathcal{M}$. For all models the p-value for the test statistic was smaller than 0.0002. Therefore, the null hypothesis that $\hat{y}_{t_0+t+1} = 0$, for all $t > 0$, was always rejected.

Table 5.6 reports the values of the normalized sum of the realized predictive likelihood for each ARCH-type models. The statistic was obtained by considering the GED(ν) probability density function, for different values of ν. In particular, for $\nu = 2$ we have the Gaussian case. The results in 5.6 show that $\nu = 1.45$ provides slightly better density forecasts for all three models, compared to the other values of $\nu < 2$. Moreover, density forecasts from the EGARCH model are slightly better than for the other two models.

Table 5.6: Normalized sum values of the realized predictive likelihood for each ARCH-type models considering the GED(ν) probability density function for different values of ν.

ν	SFIEGARCH	FIEGARCH	EGARCH
1.45	1.0050	1.0959	1.1331
1.50	0.9973	1.0893	1.1269
1.55	0.9899	1.0829	1.1209
1.70	0.9688	1.0649	1.1040
2.00	0.9318	1.0338	1.0749

Table 5.7 shows the estimated values of the coefficients γ_0 and γ_1 in (5.4). From Table 5.7 one concludes that, in all cases, the null hypothesis of forecast efficiency ($\gamma_0 = 0$ and $\gamma_1 = 1$) is rejected. Table 5.7 also indicates that the three models have a very similar performance. On the other hand, as we mentioned earlier, the log-return time series has several values very close to zero. Moreover, the volatility forecasts obtained from the SFIEGARCH model converge to a constant while the forecasts obtained from the FIEGARCH and EGARCH models show an oscillating behavior leading to forecasts close to zero more often than the SFIEGARCH model.

Table 5.7: Estimated values of the coefficients γ_0 and γ_1 defined in the forecast efficiency regression. The number in parenthesis are the standard errors corrected for heteroskedasticity and autocorrelation (by using the HAC estimator) and uncertainty concerning parameter estimation (upon multiplying by $\lambda = \sqrt{1 + n_p/n}$, where n is the sample size used to fit the model and n_p is the number of predicted values).

h	SFIEGARCH		FIEGARCH		EGARCH	
	γ_0	γ_1	γ_0	γ_1	γ_0	γ_1
1	0.8447 (0.2668)	-0.1087 (0.2043)	0.8018 (0.2453)	-0.0789 (0.1796)	0.7636 (0.4685)	-0.0458 (0.4522)
2	1.5941 (0.6158)	-0.0942 (0.1874)	1.5559 (0.5474)	-0.0840 (0.1799)	1.8090 (0.9887)	-0.2285 (0.4327)
3	2.2665 (1.1055)	-0.0657 (0.1933)	2.2181 (0.9726)	-0.0538 (0.1758)	2.6404 (1.8851)	-0.2181 (0.5425)
4	3.1568 (1.3799)	-0.0906 (0.2194)	3.1464 (1.1611)	-0.0918 (0.1865)	3.8928 (2.4312)	-0.3186 (0.5390)
5	4.0411 (2.3326)	-0.1074 (0.2713)	4.0263 (2.2802)	-0.1077 (0.2753)	5.0712 (4.0225)	-0.3666 (0.7009)
6	5.0353 (2.6067)	-0.1345 (0.2473)	5.0360 (2.3323)	-0.1374 (0.2297)	6.3321 (3.4384)	-0.4120 (0.4943)
7	6.0834 (3.4448)	-0.1598 (0.2711)	6.1792 (3.4706)	-0.1743 (0.2811)	7.7567 (4.0976)	-0.4684 (0.5108)
8	7.2811 (3.2498)	-0.1938 (0.2080)	7.3698 (3.2783)	-0.2050 (0.2247)	9.1453 (3.1329)	-0.5013 (0.3597)
9	8.7566 (4.5220)	-0.2508 (0.2554)	8.7339 (3.9764)	-0.2485 (0.2409)	10.4351 (3.0866)	-0.5120 (0.3238)
10	9.8906 (2.9644)	-0.2636 (0.2124)	9.7976 (2.6865)	-0.2548 (0.2050)	11.4789 (3.4867)	-0.4921 (0.4433)

6　Conclusions

In this work we presented several theoretical results regarding seasonal FIEGARCH (SFIEGARCH) processes. The similarities/differences between this model and the PLM-EGARCH model, introduced by Bordignon et al. (2009), were also discussed.

We proved here that $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$ is a SARFIMA$(0, d, 0) \times (p, 0, q)_s$ process. With this result we provided a complete description of the process $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$ given that necessary and sufficient conditions for the existence, stationarity and ergodicity, as well as the autocovariance structure and spectral representation of the SARFIMA processes are well known. These results were used to establish the conditions for the existence, stationarity and ergodicity of the process $\{X_t\}_{t \in \mathbb{Z}}$ itself. We also provided conditions for the existence of the r-th moment of the random variables $\{\sigma_t^2\}_{t \in \mathbb{Z}}$ and $\{X_t^2\}_{t \in \mathbb{Z}}$ when the underlying distribution is GED. Expressions for the asymmetry and kurtosis measures of any stationary SFIEGARCH process were also derived.

In this paper we also contributed to the theory of SARFIMA$(0, d, 0) \times (p, 0, q)_s$ processes by extending the range of the parameter d for the invertibility property and by providing an alternative asymptotic expression for the autocovariance function $\gamma_{\ln(\sigma_t^2)}(\cdot)$ of the process $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$. We also derived the exact and the asymptotic expressions for the autocovariance and spectral density functions of the process $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$.

As an illustration, we analyzed the behavior of the intraday volatility of the S&P500 US stock index log-returns in the period from December 13, 2004 to October 10, 2009. To account for serial correlation in the log-return time series we considered a constrained ARMA model. An SFIEGARCH model was used to account for both the long memory and seasonal behavior for the volatility. FIEGARCH and EGARCH models were also considered in order to analyze the influence of including or not the seasonal parameter in the volatility equation. We conclude that, for this particular time series, not including the seasonal parameter (FIEGARCH model) or ignoring the long-memory behavior (EGARCH model) lead to models which are close enough to the non-stationary region.

Acknowledgments

Appendix A: Proofs

In this section we provide the proofs of all propositions, lemmas, corollaries and theorems stated in Sections 1 - 3, in the same order as they appear in the text.

Proof of Proposition 1.2:

Straightforward.　　　　　　　　　　　　　　　　　　　　　　　　　　　　　■

Proof of Theorem 1.1:

From (1.5) and (1.9), one has

$$\sum_{k=0}^{\infty} \lambda_{d,k} z^k = \lambda(z) = \frac{\alpha(z)}{\beta(z)}(1 - z^s)^{-d} = \left(\sum_{k=0}^{\infty} f_k z^k \right)\left(\sum_{j=0}^{\infty} \pi_{d,j} z^j \right) = \sum_{k=0}^{\infty} \left(\sum_{j=0}^{k} f_j \pi_{d,k-j} \right) z^k.$$

It follows that, $\lambda_{d,k} = \sum_{j=0}^{k} f_j \pi_{d,k-j}$ or, equivalently,

$$\lambda_{d,sk+r} = \sum_{j=0}^{sk+r} f_j \pi_{d,sk+r-j} = \sum_{j=0}^{k} f_{sj+r} \pi_{d,sk-sj}, \quad \text{for all } k \geq 0 \text{ and } r \in \{0,\cdots,s-1\}. \quad (A.1)$$

Let $\mathscr{K}(\cdot)$ be defined as

$$\mathscr{K}(sk+r) = \sum_{j=0}^{\lfloor\sqrt{k}\rfloor} f_{sj+r} \frac{\pi_{d,sk-sj}}{\pi_{sk}}, \quad \text{for all } k \in \mathbb{N} \quad \text{and all } r \in \{0,\cdots,s-1\}.$$

From expression (1.10), $\pi_{d,sk-sj} \sim \pi_{d,sk}$, uniformly, for all $0 \leq j \leq \lfloor\sqrt{k}\rfloor$. Also, since $\beta(\cdot)$ has no roots in the closed disk $\{z : |z| \leq 1\}$, there exist constants $a > 0$ and $B > 0$ such that $|f_k| < Be^{-ak}$, for all $k \in \mathbb{N}$ (see Kokoszka and Taqqu, 1994). Hence, $\sum_{j>k_0} |f_j| \leq \frac{Be^{-a(k_0+1)}}{1-e^{-a}}$, for all $k_0 \in \mathbb{N}$. It follows that,

$$\lim_{k\to\infty} \sum_{r=0}^{s-1} \mathscr{K}(sk+r) = \sum_{r=0}^{s-1} \left[\lim_{k\to\infty} \sum_{j=0}^{\lfloor\sqrt{k}\rfloor} f_{sj+r} \frac{\pi_{d,sk-sj}}{\pi_{sk}} \right] = \sum_{j=0}^{\infty} f_j = \frac{\alpha(1)}{\beta(1)}.$$

Moreover,

$$\left| \sum_{j=\lfloor\sqrt{k}\rfloor+1}^{k} f_{sj+r} \pi_{d,sk-sj} \right| \leq \sup_{k\in\mathbb{N}} \{|\pi_{d,k}|\} \sum_{j>\lfloor\sqrt{k}\rfloor} |f_{sj+r}| = o(k^{-\nu}), \quad \text{for any } \nu > 0.$$

Thus,

$$\lambda_{d,sk+r} = \pi_{d,sk} \sum_{j=0}^{\lfloor\sqrt{k}\rfloor} f_{sj+r} \frac{\pi_{d,sk-sj}}{\pi_{sk}} + \sum_{j=\lfloor\sqrt{k}\rfloor+1}^{k} f_{sj+r} \pi_{d,sk-sj} = \pi_{d,sk} \mathscr{K}(sk+r) + o(h^{-\nu}),$$

for any $\nu > 0$, and

$$\lim_{k\to\infty} \left[\left(\sum_{r=0}^{s-1} \lambda_{d,sk+r} \right) \left(\pi_{sk} \frac{\alpha(1)}{\beta(1)} \right)^{-1} \right] = \frac{\beta(1)}{\alpha(1)} \lim_{k\to\infty} \left[\sum_{r=0}^{s-1} \mathscr{K}(sk+r) + \frac{o(h^{-\nu})}{\pi_{sk}} \right] = 1.$$

Therefore, the result holds. ■

Proof of Theorem 1.2:

From expression (A.1), for all $r \in \{0,\cdots,s-1\}$,

$$\lambda_{d,sk+r} - \frac{1}{\Gamma(d)k^{1-d}} \frac{\alpha(1)}{\beta(1)} = \sum_{j=0}^{sk+r} f_j \left[\pi_{d,sk+r-j} - \frac{1}{\Gamma(d)k^{1-d}} \right] - \frac{1}{\Gamma(d)k^{1-d}} \sum_{j>sk+r} f_j$$

$$= \sum_{j=0}^{k} f_{sj+r} \left[\pi_{d,sk-sj} - \frac{1}{\Gamma(d)k^{1-d}} \right] - \frac{1}{\Gamma(d)k^{1-d}} \sum_{j=0}^{sk+r} f_j \mathbb{I}_{\mathbb{R}\setminus\mathbb{N}} \left(\frac{|j-r|}{s} \right)$$

$$- \frac{1}{\Gamma(d)k^{1-d}} \sum_{j>sk+r} f_j.$$

Since $\sum_{j=0}^{\infty} |f_j| \leq \infty$, there exist $a > 0$ and $B > 0$ such that $|f_j| < Be^{-aj}$, for all $j \in \mathbb{N}$. Consequently,

$$\frac{1}{\Gamma(d)k^{1-d}} \sum_{j>sk+r} f_j = o(k^{-\nu}) \quad \text{and} \quad \frac{1}{\Gamma(d)k^{1-d}} \sum_{j=0}^{sk+r} f_j \mathbb{I}_{\mathbb{R}\setminus\mathbb{N}} \left(\frac{|j-r|}{s} \right) = O(k^{d-1}) \mathbb{I}_{\mathbb{N}\setminus\{0,1\}}(s),$$

for any $\nu > 0$, as k goes to infinity. From Theorem 1.1, one concludes that

$$\sum_{j=0}^{k} f_{sj+r} \left[\pi_{d,sk-sj} - \frac{1}{\Gamma(d)k^{1-d}} \right] = \sum_{j=0}^{\lfloor\sqrt{k}\rfloor} f_{sj+r} \left[\pi_{d,sk-sj} - \frac{1}{\Gamma(d)k^{1-d}} \right] + o(k^{-\nu}) - \frac{1}{\Gamma(d)k^{1-d}} \sum_{j=\lfloor\sqrt{k}\rfloor+1}^{k} f_{sj+r}$$

$$= \sum_{j=0}^{\lfloor\sqrt{k}\rfloor} f_{sj+r} \left[\pi_{d,sk-sj} - \frac{1}{\Gamma(d)k^{1-d}} \right] + O(k^{2-d}),$$

for any $r \in \{0, \cdots, s-1\}$. Since

$$\int_1^{\sqrt{k}} e^{-a(sx+r)} \left[\left(\frac{k-x}{k}\right)^{d-1} - 1 \right] \mathrm{d}x = e^{-ar} \int_1^{\sqrt{k}} e^{-asx} \left[\left(1 - \frac{x}{k}\right)^{d-1} - 1 \right] \mathrm{d}x \quad \text{(setting } x = ky\text{)}$$

$$= e^{-ar} k \int_{1/k}^{\sqrt{k}/k} e^{-asky} \left[(1-y)^{d-1} - 1\right] \mathrm{d}y$$

$$\leq (1-d)2^{1-d} e^{-ar} k \int_0^{\sqrt{k}/k} e^{-asky} y \, \mathrm{d}y \quad \text{(setting } u = asky\text{)}$$

$$= (1-d)2^{1-d} \frac{e^{-ar}}{ask} \int_0^{as\sqrt{k}} u e^{-u} \mathrm{d}u = O(k^{-1}), \quad \text{as } k \to \infty,$$

by using equality (1.5), one concludes that

$$\sum_{j=0}^{\lfloor \sqrt{k} \rfloor} f_{sj+r} \left[\pi_{d,sk-sj} - \frac{1}{\Gamma(d)k^{1-d}} \right] = O(k^{d-2}) \sum_{j=0}^{\lfloor \sqrt{k} \rfloor} f_{sj+r} + \frac{1}{\Gamma(d)k^{1-d}} \sum_{j=0}^{\lfloor \sqrt{k} \rfloor} f_{sj+r} \left[\left(\frac{k-j}{k}\right)^{d-1} - 1 \right]$$

$$= O(k^{d-2}), \quad \text{as } k \to \infty.$$

Therefore, equation (1.12) holds.   ∎

Proof of Proposition 1.1:

By definition,

$$\frac{\alpha(z)}{\beta(z)} (1-z^s)^{-d} = \sum_{k=0}^{\infty} \lambda_{d,k} z^k \quad \Longrightarrow \quad \alpha(z) = \beta(z)(1-z^s)^d \left(\sum_{k=0}^{\infty} \lambda_{d,k} z^k \right). \tag{A.2}$$

Set

$$\alpha_m^* = \begin{cases} \alpha_m, & \text{if } 0 \leq m \leq p, \\ 0, & \text{if } m > p, \end{cases} \quad \text{and} \quad \beta_m^* = \begin{cases} \beta_m, & \text{if } 0 \leq m \leq q, \\ 0, & \text{if } m > q. \end{cases} \tag{A.3}$$

Then,

$$\beta(z)(1-z^s)^d = \left(\sum_{k=0}^{\infty} -\beta_k^* z^k \right) \left(\sum_{k=0}^{\infty} \delta_{d,k} \mathcal{B}^{sk} \right) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} -\beta_i^* \delta_{d,j} \mathcal{B}^{sj+i} = \sum_{k=0}^{\infty} \tau_k \mathcal{B}^k, \tag{A.4}$$

where $\tau_k = \sum_{i=0}^{k} -\beta_i^* \delta_{d,\frac{k-i}{s}}^*$, for all $k \in \mathbb{N}$, and δ_m^* is defined in (1.13).

Thus, (A.2) can be rewritten as

$$\beta(\mathcal{B})(1-\mathcal{B}^s)^d \left(\sum_{k=0}^{\infty} \lambda_{d,k} \mathcal{B}^k \right) = \left(\sum_{k=0}^{\infty} \tau_k \mathcal{B}^k \right) \left(\sum_{k=0}^{\infty} \lambda_{d,k} \mathcal{B}^k \right) = \sum_{k=0}^{\infty} \left(\sum_{i=0}^{k} \lambda_{d,i} \tau_{k-i} \right) \mathcal{B}^k$$

$$= \sum_{k=0}^{\infty} \left[\sum_{i=0}^{k} \lambda_{d,i} \left(-\sum_{j=0}^{k-i} \beta_j^* \delta_{d,\frac{k-i-j}{s}}^* \right) \right] \mathcal{B}^k$$

$$= \lambda_{d,0} + \sum_{k=1}^{\infty} \left[\lambda_{d,k} - \sum_{i=0}^{k-1} \lambda_{d,i} \left(\sum_{j=0}^{k-i} \beta_j^* \delta_{d,\frac{k-i-j}{s}}^* \right) \right] \mathcal{B}^k. \tag{A.5}$$

Therefore, from (1.13) and (A.5), expression (A.2) holds if and only if

$$-\alpha_0^* = \lambda_{d,0} \quad \text{and} \quad -\alpha_k^* = \lambda_{d,k} - \sum_{i=0}^{k-1} \lambda_{d,i} \left(\sum_{j=0}^{k-i} \beta_j^* \delta_{d,\frac{k-i-j}{s}}^* \right), \quad \text{for all} \quad k \geq 1,$$

and the result holds.   ∎

Proof of Lemma 1.1:

From Theorem 1.1, $\sum_{k=0}^{\infty} \lambda_{d,k}^2 < \infty$ if and only if $d < 0.5$. Therefore, if (1.15) is a.s. convergent, by applying the three series theorem (see Billingsley, 1995), one concludes that, necessarily, $d < 0.5$. On the other hand, if $d < 0.5$, by applying Kolmogorov's convergence criteria (Billingsley, 1995, theorem 22.6), one concludes that (1.15) is a.s. convergent. Finally, from Theorem 1.1, if $d < 0$, $\sum_{k=0}^{\infty} |\lambda_{d,k}| < \infty$ and, from proposition 3.1.1 in Brocwel and Davis (1991), the series (1.15) is absolutely a.s. convergent. ∎

Proof of Corollary 1.1:

The result follows immediately from Proposition 1.2, Lemma 1.1 and the definition of SARFIMA processes in Bisognin and Lopes (2009). In particular, if $s = 1$, it is an ARFIMA(q, d, p) process (see Brockwell and Davis, 1991). ∎

Proof of Corollary 1.2:

From Lemma 1.1, if $d < 0.5$, the random variable σ_t^2 is finite with probability one, for all $t \in \mathbb{Z}$. Since, by assumption $\mathbb{E}(Z_t^2) = \mathbb{E}(Z_0^2) = 1$, the random variable Z_t is finite with probability one, for all $t \in \mathbb{Z}$. Therefore, from expression (1.1), the result follows. ∎

Proof of Theorem 1.3:

By Hölder's inequality, it suffices to prove the result for $\mathfrak{p} = 2$. From Bisognin and Lopes (2009), the spectral density function of $\{Y_t\}_{t \in \mathbb{Z}}$ is given by

$$f_{\ln(\sigma_t^2)}(\lambda) = \frac{\sigma_g^2}{2\pi} \frac{|\alpha(e^{-i\lambda})|^2}{|\beta(e^{-i\lambda})|^2} |1 - e^{-is\lambda}|^{-2d} = \frac{\sigma_g^2}{2\pi} \frac{|\alpha(e^{-i\lambda})|^2}{|\beta(e^{-i\lambda})|^2} \left[2\sin\left(\frac{s\lambda}{2}\right) \right]^{-2d}, \quad \text{(A.6)}$$

for all $\lambda \in [0, \pi]$. The authors also show that, for all $k = 0, \cdots, \lfloor \frac{s}{2} \rfloor$,

$$f_Y(\lambda) \sim \frac{\sigma_\varepsilon^2}{2\pi} \left| \frac{\alpha(e^{-i\lambda_k})}{\beta(e^{-i\lambda_k})} \right|^2 s^{-2d} |\lambda - \lambda_k|^{-2d}, \quad \text{as} \quad \lambda \to \lambda_k := \frac{2\pi k}{s}. \quad \text{(A.7)}$$

Suppose first that the L^2 convergence holds. Notice that, there exists a real constant $c > 0$ such that $f_Y(\lambda) \sim c\lambda^{-2d}$, as $\lambda \to 0^+$. Consequently, from proposition 1.5.8 in Bingham et al. (1987), if $d \geq 0.5$, the function $f_Y(\cdot) \notin L^1$ and hence cannot be a spectral density function. From Theorem 1.1, if $d \leq -1$, then $\tilde{\lambda}_{d,k} \nrightarrow 0$, when $k \to \infty$. Consequently, $\tilde{\lambda}_{d,k} Y_{t-k} \nrightarrow 0$ in $L^{\mathfrak{p}}$-norm, when $k \to \infty$, and the series representation $\sum_{k=0}^{\infty} \tilde{\lambda}_{d,k} Y_{t-k}$ cannot converge in $L^{\mathfrak{p}}$-norm, for any $0 < \mathfrak{p} \leq 2$. Therefore, necessarily, $d \in (-1, 0.5)$.

Suppose now that $d \in (-1, 0.5)$ and assume $s > 1$ (the case $s = 1$ can be found in Bondon and Palma, 2007). Notice that $d = d_1 + d_2$, where $d_1 \in (-0.5, 0)$ and $d_2 \in (-0.5, 0.5)$. Define the functions $f_1(\cdot)$ and $f_2(\cdot)$ as follows,

$$f_1(\lambda) := \frac{f_Y(\lambda)}{f_2(\lambda)} \quad \text{and} \quad f_2(\lambda) := |1 - e^{is\lambda}|^{-2d_2}, \quad \text{for all } \lambda \in [0, \pi].$$

Since $d_1 \in (-0.5, 0)$, from expressions (A.6) and (A.7), it is obvious that, $f_1(\cdot) \in L^\infty$ and

$$f_1(\lambda) \sim l\left(\frac{1}{\lambda - 2\pi k/s}\right) \left| \lambda - \frac{2\pi k}{s} \right|^{-2d_1}, \quad \text{as } \lambda \to \frac{2\pi k}{s}, \quad \text{for all } 0 \leq k \leq \left\lfloor \frac{s}{2} \right\rfloor,$$

where $l(\cdot)$ is a slowly varying function at infinity (for details, see Bingham et al., 1987). Since $d_2 \in (-0.5, 0.5)$, from expression (A.6), one easily concludes that, for any $\epsilon \in (0, 2\pi/s)$, the function

$f_1^{-1}(\cdot)$ is bounded in the interval $[2\pi k/s + \epsilon, 2\pi(k+1)/s - \epsilon]$, for all $0 \leq k \leq \lfloor s/2 \rfloor$. Moreover, from expression (A.7), there exists an $\epsilon \in (0, 2\pi/s)$, such that

$$f_1^{-1}(\lambda) \leq 2l_1\left(\frac{1}{\lambda - 2\pi k/s}\right)\left|\lambda - \frac{2\pi k}{s}\right|^{2d_1}, \quad \text{for all} \quad \lambda \in \left(\frac{2\pi k}{s} - \epsilon, \frac{2\pi k}{s}\right) \cup \left(\frac{2\pi k}{s}, \frac{2\pi k}{s} + \epsilon\right)$$

and $k \in \{0, \cdots, \lfloor s/2 \rfloor\}$, where $l_1(\cdot) = l^{-1}(\cdot)$, is also a slowly varying function. It follows that $\int_0^\pi f_1^{-1}(\lambda)d\lambda$ can be written as

$$\int_0^\pi f_1^{-1}(\lambda)d\lambda = \begin{cases} S_1, & \text{if } s \text{ is even;} \\ S_1 + S_2, & \text{if } s \text{ is odd,} \end{cases}$$

where

$$S_1 := \sum_{k=0}^{\lfloor s/2 \rfloor - 1} \left[\int_{2\pi k/s}^{2\pi k/s + \epsilon} f_1^{-1}(\lambda)d\lambda + \int_{2\pi k/s + \epsilon}^{2\pi(k+1)/s - \epsilon} f_1^{-1}(\lambda)d\lambda + \int_{2\pi(k+1)/s - \epsilon}^{2\pi(k+1)/s} f_1^{-1}(\lambda)d\lambda \right],$$

$$S_2 := \int_{2\pi \lfloor s/2 \rfloor/s}^{2\pi \lfloor s/2 \rfloor/s + \epsilon} f_1^{-1}(\lambda)d\lambda + \int_{2\pi \lfloor s/2 \rfloor/s + \epsilon}^{\pi} f_1^{-1}(\lambda)d\lambda,$$

and it satisfies

$$\int_0^\pi f_1^{-1}(\lambda)d\lambda \leq \begin{cases} S_1^*, & \text{if } s \text{ is even;} \\ S_1^* + S_2^*, & \text{if } s \text{ is odd,} \end{cases}$$

where

$$S_1^* := \sum_{k=0}^{\lfloor s/2 \rfloor - 1} \left[2\int_{2\pi k/s}^{2\pi k/s + \epsilon} l_1\left(\frac{1}{\lambda - 2\pi k/s}\right)\left|\lambda - \frac{2\pi k}{s}\right|^{2d_1}d\lambda + \int_{2\pi k/s + \epsilon}^{2\pi(k+1)/s - \epsilon} f_1^{-1}(\lambda)d\lambda \right.$$

$$\left. + 2\int_{2\pi(k+1)/s - \epsilon}^{2\pi(k+1)/s} l_1\left(\frac{1}{\lambda - 2\pi k/s}\right)\left|\lambda - \frac{2\pi k}{s}\right|^{2d_1}d\lambda \right],$$

$$S_2^* := 2\int_{2\pi \lfloor s/2 \rfloor/s}^{\pi \lfloor s/2 \rfloor 2/s + \epsilon} l_1\left(\frac{1}{\lambda - 2\pi k/s}\right)\left|\lambda - \frac{2\pi k}{s}\right|^{2d_1}d\lambda + \int_{2\pi \lfloor s/2 \rfloor/s + \epsilon}^{\pi} f_1^{-1}(\lambda)d\lambda.$$

Thus, since $d_1 > -0.5$, by proposition 1.5.10 in Bingham et al. (1987) we conclude that $f_1^{-1}(\cdot) \in L^1$. By comparing the function $f_2(\cdot)$ with the corresponding one in Bondon and Palma (2007), we conclude that $f_2(\lambda) = |1 - e^{is\lambda}|$ satisfies condition A_p, with $p = 2$, from theorem 3 in Bloomfield (1985). Therefore, taking $p = 1$ in theorem 4 from Bloomfield (1985) we conclude that $\tilde{\lambda}(\cdot)$ has a Fourier series that converges in $L^2(f)$, where $F(\cdot)$ is the spectral distribution function of Y_t, for all $t \in \mathbb{Z}$, and the result follows.   ∎

Proof of Lemma 2.1:

See Bisognin and Lopes (2009).   ∎

Proof of Corollary 2.1:

It follows immediately from Lemma 2.1.   ∎

Proof of Theorem 2.1

Suppose $d < 0.5$. (i) By hypothesis, Z_t is finite with probability one, for all $t \in \mathbb{Z}$. From Corollary 1.1, the random variables $\ln(\sigma_t^2) - \omega$ is finite with probability one, for all $t \in \mathbb{Z}$, so it is σ_t. Thus, from theorem 3.5.8 in Stout (1974), $\{X_t\}_{t \in \mathbb{Z}}$ is a strictly stationary and ergodic process.

(ii) Assume that $\mathbb{E}(|\ln(Z_0^2)|) < \infty$. It follows that the random variable $|\ln(Z_t^2)|$ is finite with probability one, for all $t \in \mathbb{Z}$. From Corollary 1.1, $\mathbb{E}(|\ln(\sigma_t^2)|) < \infty$, for all $t \in \mathbb{Z}$. By expression (1.2), $\ln(X_t^2) = \ln(\sigma_t^2) + \ln(Z_t^2)$, for all $t \in \mathbb{Z}$. It follows that the random variable $\ln(X_t^2)$ is finite with probability one, for all $t \in \mathbb{Z}$, and hence the stochastic process $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$ is well defined. From theorem 3.5.8 in Stout (1974), the stochastic process $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$ is strictly stationary and ergodic. Moreover, if $\mathbb{E}([\ln(Z_0^2)]^2) < \infty$ then $\mathrm{Var}\big(\ln(X_t^2)\big) = \mathrm{Var}\big(\ln(X_0^2)\big) = \mathrm{Var}(\ln(\sigma_0^2)) + \mathrm{Var}(\ln(Z_0^2)) < \infty$, for all $t \in \mathbb{Z}$. Therefore, $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$ is weakly stationary. ∎

Proof of Theorem 2.2:

From Corollary 2.1 and Theorem 2.1 both processes $\{X_t\}_{t\in\mathbb{Z}}$ and $\{\sigma_t\}_{t\in\mathbb{Z}}$ are strictly stationary and hence, any existing moments are time invariant. Let $r > 2$ be any real number such that $\mathbb{E}(|Z_t|^r) < \infty$. By the independence hypothesis $\mathbb{E}(|X_t|^r) = \mathbb{E}(|\sigma_t|^r)\mathbb{E}(|Z_t|^r)$, for all $t \in \mathbb{Z}$. Since $\mathbb{E}(|Z_t|^r) < \infty$, one has to show that $\mathbb{E}(|\sigma_t|^r) < \infty$. From expression (1.2), and from the i.i.d. hypothesis on the random variables Z_t, for all $t \in \mathbb{Z}$, one has

$$\mathbb{E}(|\sigma_t|^r) = \mathbb{E}\Big(\Big|\exp\Big\{\frac{\omega}{2} + \frac{1}{2}\sum_{k=0}^{\infty}\lambda_{d,k}g(Z_{t-1-k})\Big\}\Big|^r\Big) = e^{\frac{r\omega}{2}}\prod_{k=0}^{\infty}\mathbb{E}\Big(\exp\Big\{\frac{r}{2}\lambda_{d,k}g(Z_0)\Big\}\Big). \qquad (A.8)$$

From (2.1), expression (A.8) converges to a non-zero constant (see, section 0.25 in Gradshteyn and Ryzhik, 2000). By Hölder's inequality the result follows for all $0 < m < r$. ∎

Proof of Corollary 2.2:

By hypothesis, $d < 0.5$ and $\beta(z)$ has no roots in the closed disk $\{z : |z| \leq 1\}$. From Theorem 1.1, it follows that $\sum_{k=0}^{\infty}\lambda_{d,k}^2 < \infty$. Moreover, from expression A2.4 in Nelson (1991), for all $r > 0$, $\mathbb{E}(\exp\{r\lambda_{d,k}g(Z_0)\}) = 1 + O(\lambda_{d,k}^2)$, as $k \to \infty$. Thus, the result follows immediately from Theorem 2.2. ∎

Proof of Proposition 2.1:

Let $\{X_t\}_{t\in\mathbb{Z}}$ be any stationary SFIEGARCH process. Let $\lambda(\cdot)$ be the polynomial defined by (1.9). Since $\mathbb{E}(X_t^r) = \mathbb{E}(\sigma_t^r Z_t^r)$, for all $r > 0$, $\mathbb{E}(X_t) = 0$ and $\mathbb{E}(Z_t^2) = 1$, for all $t \in \mathbb{Z}$, the asymmetry and kurtosis measures of $\{X_t\}_{t\in\mathbb{Z}}$ are given, respectively, by

$$A_X = \frac{\mathbb{E}(X_t^3)}{\big(\mathbb{E}(X_t^2)\big)^{3/2}} = \frac{\mathbb{E}(\sigma_t^3)\mathbb{E}(Z_t^3)}{\big(\mathbb{E}(\sigma_t^2)\big)^{3/2}} \quad \text{and} \quad K_X = \frac{\mathbb{E}(X_t^4)}{\big(\mathbb{E}(X_t^2)\big)^2} = \frac{\mathbb{E}(\sigma_t^4)\mathbb{E}(Z_t^4)}{\big(\mathbb{E}(\sigma_t^2)\big)^2}. \qquad (A.9)$$

Upon replacing expression (A.8) in (A.9) the result follows. ∎

Proof of Lemma 2.2:

See Bisognin and Lopes (2009). ∎

Proof of Corollary 2.3:

Notice that, for any $\eta \in \mathbb{N}$ one can write

$$\gamma_{\ln(\sigma_t^2)}(sh + r) = \sum_{|k| \leq \eta}\gamma_A(sk + r)\gamma_V(sh - sk) + \sum_{|k| > \eta}\gamma_A(sk + r)\gamma_V(sh - sk).$$

Since $\sum_{h\in\mathbb{Z}}|\gamma_A(h)| < \infty$, there exist $a > 0$ and $B > 0$ such that $|\gamma_A(h)| < Be^{-ah}$, for all $h \in \mathbb{N}$ (see Kokoszka and Taqqu, 1994). Also, $\gamma_A(sk + r) = \gamma_A(-s|k| + r) = \gamma_A(s|k| - r)$, for all $k < -\eta$. Thus,

$$\Big|\sum_{|k| > \eta}\gamma_A(sk + r)\gamma_V(sh - sk)\Big| \leq \gamma_V(0)\sum_{k > \eta}\Big(|\gamma_A(sk + r)| + |\gamma_A(sk - r)|\Big)$$

$$\leq \frac{B\gamma_V(0)(e^{-ar} + e^{ar})}{1 - e^{-a}} e^{-as(\sqrt{h}+1)} = o(h^{-\nu}), \quad \text{for all } \nu > 0.$$

On the other hand,

$$\sum_{|k| \leq \eta} \gamma_A(sk + r)\gamma_V(sh - sk) = \gamma_V(sh) \sum_{|k| \leq \eta} \gamma_A(sk + r)\frac{\gamma_V(sh - sk)}{\gamma_V(sh)} := \gamma_V(sh)\mathscr{G}(sh + r).$$

Let $\eta = \lfloor \sqrt{h} \rfloor$. Then, for any $|k| \leq \eta = \lfloor \sqrt{h} \rfloor$, one has $(h - k)^{2d-1} \sim h^{2d-1}$, uniformly, as $h \to \infty$. From expression (2.6),

$$\gamma_V(sh) = \sigma_g^2 \frac{(-1)^h \Gamma(1 - 2d)}{\Gamma(1 - d + h)\Gamma(1 - d - h)} = \sigma_g^2 \frac{\Gamma(1 - 2d)}{\Gamma(1 - d)\Gamma(d)}\frac{\Gamma(h + d)}{\Gamma(h + 1 - d)} \sim \sigma_g^2 \frac{\Gamma(1 - 2d)}{\Gamma(1 - d)\Gamma(d)}h^{2d-1}, \quad \text{as} \quad h \to \infty,$$
(A.10)

where σ_g^2 is given by (1.14). Hence, $\gamma_V(sh - sk) = \gamma_V(s(h - k)) \sim \gamma_V(sh)$, uniformly, for all $|k| \leq \eta = \lfloor \sqrt{h} \rfloor$. Since $\sum_{h \in \mathbb{Z}} |\gamma_A(h)| < \infty$, it follows that

$$\lim_{h \to \infty} \sum_{r=0}^{s-1} \mathscr{G}(sh + r) = \lim_{h \to \infty} \sum_{r=0}^{s-1} \sum_{|k| \leq \eta} \gamma_A(sk + r)\frac{\gamma_V(sh - sk)}{\gamma_V(sh)} = \sum_{k \in \mathbb{Z}} \gamma_A(k),$$

and expression (2.7) holds.   ∎

Proof of Theorem 2.3:

From Theorem 2.1, $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$ is a stationary (weakly and strictly) process. Thus,

$$\gamma_{\ln(X_t^2)}(h) = \text{Cov}(\ln(X_t^2), \ln(X_{t+h}^2)) = \text{Cov}(\ln(X_0^2), \ln(X_h^2)), \quad \text{for all } t, h \in \mathbb{Z}$$

From expression (1.1), one concludes that

$$\gamma_{\ln(X_t^2)}(h) = \text{Cov}(\ln(\sigma_0^2), \ln(\sigma_h^2)) + \text{Cov}(\ln(\sigma_0^2), \ln(Z_h^2)) + \text{Cov}(\ln(Z_0^2), \ln(\sigma_h^2)) + \text{Cov}(\ln(Z_0^2), \ln(Z_h^2)),$$

for all $h \in \mathbb{Z}$. Notice that $\text{Cov}(\ln(\sigma_0^2), \ln(\sigma_h^2)) = \gamma_{\ln(\sigma_t^2)}(h)$ is the autocovariance function of $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$, given in Lemma 2.2, and $\text{Cov}(\ln(Z_0^2), \ln(Z_h^2)) = \sigma_{\ln(Z_t^2)}^2 \mathbb{I}_{\{0\}}(h)$. Moreover, from expression (1.2),

$$\text{Cov}(\ln(\sigma_h^2), \ln(Z_0^2)) = \text{Cov}\left(\sum_{k=0}^{\infty} \lambda_{d,k} g(Z_{h-1-k}), \ln(Z_0^2)\right) = \begin{cases} 0, & \text{if } h \leq 0; \\ C_1 \lambda_{d,h-1}, & \text{if } h > 0, \end{cases} \quad (A.11)$$

where $C_1 = \text{Cov}(\ln(Z_0^2), g(Z_0)) = \theta\,\mathbb{E}(Z_0 \ln(Z_0^2)) + \gamma\mathbb{E}(|Z_0| \ln(Z_0^2)) - \gamma\mathbb{E}(|Z_0|)\mathbb{E}(\ln(Z_0^2))$. Thus, from expression (A.11), $\text{Cov}(\ln(\sigma_0^2), \ln(Z_h^2)) + \text{Cov}(\ln(Z_0^2), \ln(\sigma_h^2)) = C_1\lambda_{d,|h|-1}\mathbb{I}_{\mathbb{Z}^*}(h)$, for all $h \in \mathbb{Z}$, and expression (2.8) holds. Expression (2.9) follows directly from Corollary 2.3 and Theorem 1.1.   ∎

Proof of Corollary 2.4:

From Corollary 2.3 and equation (A.10),

$$\sum_{r=0}^{s-1} \gamma_{\ln(\sigma_t^2)}(sh + r) = \gamma_V(sh) \sum_{r=0}^{s-1} \mathscr{G}(sh + r) + o(h^{-\nu}) \sim \mathscr{C}_1 h^{2d-1}, \quad \text{as } h \to \infty.$$

From Theorem 1.1 and equation (1.10) one concludes that $\mathscr{K}(sh - 1) = \mathscr{K}(s(h - 1) + (s - 1))$, for all $h \geq 1$, and

$$\sum_{r=0}^{s-1} \lambda_{d,sh+r-1} = \pi_{d,s(h-1)}\mathscr{K}(sh - 1) + \pi_{d,sh} \sum_{r=1}^{s-1} \mathscr{K}(sk + r - 1) + o(h^{-\nu}) \sim \mathscr{C}_2 k^{d-1}, \quad \text{as } h \to \infty,$$

for any $\nu > 0$, where $\mathscr{C}_1(\cdot)$ and $\mathscr{C}_2(\cdot)$ are given in equation (2.11). Since

$$\lim_{h \to \infty} \frac{h^{2d-1}}{h^{d-1}} = \lim_{h \to \infty} h^d = 0, \quad \text{if } d < 0, \qquad \lim_{h \to \infty} \frac{h^{d-1}}{h^{2d-1}} = \lim_{h \to \infty} h^{-d} = 0, \quad \text{if } d > 0,$$

$$\lim_{h \to \infty} \mathscr{C}_1(h) = \sigma_g^2 \frac{\Gamma(1 - 2d)}{\Gamma(1 - d)\Gamma(d)} \sum_{k \in \mathbb{Z}} \gamma_A(k) \qquad \text{and} \qquad \lim_{h \to \infty} \mathscr{C}_2(h) = \frac{1}{\Gamma(d)}\frac{\alpha(1)}{\beta(1)},$$

expression (2.10) holds.   ∎

Proof of Theorem 3.1:

From Theorem 2.1, if $d < 0.5$, the stochastic process $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$ is strictly stationary and ergodic. Moreover, if $\mathrm{Var}(\ln(Z_0^2)) < \infty$, then it is weakly stationary and hence, it has a spectral distribution function. Thus, from Herglotz's theorem (see Brockwell and Davis, 1991), it suffices to show that $f_{\ln(X_t^2)}(\cdot)$, given by (3.1), is a continuous, non-negative function and it satisfies

$$\gamma_{\ln(X_t^2)}(h) = \int_{(-\pi,\pi]} e^{ih\lambda} f_{\ln(X_t^2)}(\lambda) d\lambda, \quad \text{for all } h \in \mathbb{Z},$$

with $\gamma_{\ln(X_t^2)}(\cdot)$ given in Theorem 2.3.

The continuity of $f_{\ln(X_t^2)}(\cdot)$ follows immediately from its definition. To prove non-negativity notice that, from expression (A.6) and from the i.i.d. hypothesis on the $\{Z_t\}_{t\in\mathbb{Z}}$ process, one has

$$f_{\ln(\sigma_t^2)}(\lambda) = \frac{\sigma_g^2}{2\pi}|\Lambda(\lambda)|^2 \quad \text{and} \quad f_{\ln(Z_t^2)}(\lambda) = \frac{\sigma_{\ln(Z_t^2)}^2}{2\pi}, \quad \text{for all } \lambda \in [0,\pi],$$

where $\Lambda(z) := \lambda(e^{-iz})$ and $\lambda(\cdot)$ is defined in (1.9). Moreover, $|C_1| \le \sigma_g \sigma_{\ln(Z_t^2)}$ and $|\Re(z)| \le |z|$, for all $z \in \mathbb{C}$. Thus,

$$f_{\ln(\sigma_t^2)}(\lambda) + \frac{C_1}{\pi}\Re\big(e^{-i\lambda}\Lambda(\lambda)\big) + f_{\ln(Z_t^2)}(\lambda) \ge f_{\ln(\sigma_t^2)}(\lambda) - \frac{\sigma_g \sigma_{\ln(Z_t^2)}}{\pi}\big|e^{-i\lambda}\Lambda(\lambda)\big| + f_{\ln(Z_t^2)}(\lambda)$$

$$= \left(\frac{\sigma_g}{\sqrt{2\pi}}|\Lambda(\lambda)| - \frac{\sigma_{\ln(Z_t^2)}}{\sqrt{2\pi}}\right)^2 \ge 0, \quad \text{for all } \lambda \in [0,\pi].$$

To complete the proof, observe that

$$\gamma_{\ln(\sigma_t^2)}(h) = \int_{(-\pi,\pi]} e^{ih\lambda} f_{\ln(\sigma_t^2)}(\lambda) d\lambda, \qquad \gamma_{\ln(Z_t^2)}(h) = \int_{(-\pi,\pi]} e^{ih\lambda} f_{\ln(Z_t^2)}(\lambda) d\lambda$$

and

$$\int_{(-\pi,\pi]} e^{ih\lambda}\frac{C_1}{\pi}\Re\big(e^{-i\lambda}\Lambda(\lambda)\big) d\lambda = \frac{C_1}{\pi}\sum_{k=0}^{\infty}\lambda_{d,k}\int_{(-\pi,\pi]} e^{ih\lambda}\cos((k+1)\lambda) d\lambda$$

$$= \frac{C_1}{\pi}\sum_{k=0}^{\infty}\lambda_{d,k}\pi\mathbb{I}_{\{0\}}(k+1-|h|) = C_1\lambda_{d,|h|-1}\mathbb{I}_{\mathbb{N}^*}(|h|),$$

for all $h \in \mathbb{Z}$. Therefore, the result holds. ∎

Proof of Theorem 4.1

Let $\mathscr{S}_1 := \prod_{k=0}^{h-2}\exp\{\lambda_{d,k}g(Z_{n+h-1-k})\}$ and $\mathscr{S}_2 := \prod_{k=h-1}^{\infty}\exp\{\lambda_{d,k}g(Z_{n+h-1-k})\}$, for any $n \in \mathbb{N}$ and $h > 1$ fixed. Notice that, from expression (1.2), one can write

$$\sigma_{n+h}^2 = e^\omega \prod_{k=0}^{\infty}\exp\{\lambda_{d,k}g(Z_{n+h-1-k})\} := e^\omega \mathscr{S}_1\mathscr{S}_2, \quad \text{for all } n \in \mathbb{N} \text{ and } h > 1. \tag{A.12}$$

Also, observe that the hypothesis $\mathbb{E}(\sigma_t^2) < \infty$ implies $0 < \mathscr{S}_1, \mathscr{S}_2 < \infty$ with probability one.

The $\mathcal{F}_n$-measurability of $Z_{n+h-1-k}$, when $k \ge h-1$, and the the i.i.d. property of $\{Z_t\}_{t\in\mathbb{Z}}$ imply that

$$\mathbb{E}(\sigma_{n+h}^2|\mathcal{F}_n) = e^\omega \mathscr{S}_2\mathbb{E}(\mathscr{S}_1) = e^\omega \prod_{k=h-1}^{\infty}\exp\{\lambda_{d,k}g(Z_{n+h-1-k})\}\prod_{k=0}^{h-2}\mathbb{E}\big(\exp\{\lambda_{d,k}g(Z_0)\}\big),$$

and expression (4.1) holds.

Now, the independence of $\mathscr{S}_1$ and $\mathscr{S}_2$ implies that

$$\mathbb{E}\big(\big[\sigma_{n+h}^2 - \hat{\sigma}_{n+h}^2\big]^2\big) = \mathbb{E}\big(\big[e^\omega\mathscr{S}_1\mathscr{S}_2 - e^\omega\mathscr{S}_2\mathbb{E}(\mathscr{S}_1)\big]^2\big) = e^{2\omega}\mathbb{E}(\mathscr{S}_2^2)\mathbb{E}\big(\big[\mathscr{S}_1 - \mathbb{E}(\mathscr{S}_1)\big]^2\big), \quad \text{for any } h > 1.$$

Since $\mathbb{E}\big(\big[\mathscr{S}_1 - \mathbb{E}(\mathscr{S}_1)\big]^2\big) = \mathbb{E}(\mathscr{S}_1^2) - \big[\mathbb{E}(\mathscr{S}_1)\big]^2$, expression (4.2) holds.

To conclude the proof observe that

$$\mathbb{E}\big(\big[X_{n+h}^2 - \hat{X}_{n+h}^2\big]^2\big) = \mathbb{E}\big(\big[X_{n+h}^2 - \hat{\sigma}_{n+h}^2\big]^2\big) = \mathbb{E}\big(\big[e^\omega \mathscr{S}_1 \mathscr{S}_2 Z_{n+h}^2 - e^\omega \mathscr{S}_2 \mathbb{E}(\mathscr{S}_1)\big]^2\big), \quad \text{for any } h > 1.$$

Then, by adding and subtracting $\mathbb{E}\big(\big[e^\omega \mathscr{S}_1 \mathscr{S}_2\big]^2\big)$ to the right hand side of the above equation, one can rewrite

$$\mathbb{E}\big(\big[X_{n+h}^2 - \hat{X}_{n+h}^2\big]^2\big) = \mathbb{E}\big(\big[e^\omega \mathscr{S}_1 \mathscr{S}_2\big]^2\big)\big[\mathbb{E}(Z_{n+h}^4) - 1\big] + e^{2\omega}\mathbb{E}(\mathscr{S}_2^2)\big(\mathbb{E}(\mathscr{S}_1^2) - \big[\mathbb{E}(\mathscr{S}_1)\big]^2\big)$$
$$= \mathbb{E}(\sigma_0^4)\big[\mathbb{E}(Z_0^4) - 1\big] + mse(\sigma_{n+h}^2), \quad \text{for any } h > 1,$$

and the result follows. ∎

Proof of Theorem 4.2

Expression (4.4) and the first equation in (4.5) follow immediately by mimicking the proof of proposition 4 in Lopes and Prass (2013). The second equation in (4.5) is obtained by replacing $\check{\ln}(X_{n+h}^2)$ by $\hat{\ln}(X_{n+h}^2)$ and mimicking the proof of proposition 5 in Lopes and Prass (2013). ∎

Proof of Theorem 4.3

From (4.7), for any fixed $n \in \mathbb{Z}$,

$$r_{n+h}^2 = \mu^2 + 2\mu \sum_{i=0}^\infty \psi_i X_{n+h-i} + \sum_{k=0}^\infty \sum_{j=0}^\infty \psi_k \psi_j X_{n+h-k} X_{n+h-j}, \quad \text{for all } h > 0.$$

Thus the result follows by observing that $\mathbb{E}(X_{n+h-i}|\mathcal{F}_n) = X_{n+h-i}$, if $i \geq h$, and 0 otherwise, and that

$$\mathbb{E}(X_{n+h-k} X_{n+h-j}|\mathcal{F}_n) = \begin{cases} \mathbb{E}(X_{n+h-k}^2|\mathcal{F}_n) = \hat{\sigma}_{n+h-k}^2, & \text{for any } k,j < h, \text{ with } k = j; \\ 0, & \text{for any } k,j < h, \text{ with } k \neq j; \\ X_{n+h-k} X_{n+h-j}, & \text{for any } k,j \geq h, \text{ with } k \neq j. \end{cases}$$

for any $k, j \in \mathbb{N}$ and $h > 0$. ∎

References

T. Andersen, T. Bollerslev, F. Diebold and P. Labys (2000). "Great realizations". *Risk*, vol. 13, 105-108,

R. Baillie, T. Bollerslev and H. Mikkelsen (1996). "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity". *Journal of Econometrics*, vol. 74, 3-30.

P. Billingsley (1995). *Probability and Measure*. New York: Wiley.

N.H. Bingham, C.M. Goldie and J.L. Teugels (1987). *Regular Variation*. New York: Cambridge University Press.

C. Bisognin and S.R.C. Lopes (2009). "Properties of seasonal long memory processes." *Mathematical and Computer Modelling*, vol. 49, 1837-1851.

P. Bloomfield (1985). "On Series Representations for Linear Predictors." *The Annals of Probability*, vol. 13(1), 226-233.

T. Bollerslev (1986). "Generalized Autoregressive Conditional Heteroskedasticity". *Journal of Econometrics*, vol. 31, 307-327.

T. Bollerslev and H.O. Mikkelsen (1996). "Modeling and Pricing Long Memory in Stock Market Volatility". *Journal of Econometrics*, vol. 73, 151-184.

T. Bollerslev and J.M. Wooldridge (1992). "Quasi-maximum Likelihood Estimation and Inference in Dynamic Models with Time-varying Covariances". *Econometric Reviews*, vol. 11, 143-172.

P. Bondon and W. Palma (2007). "A class of antipersistent processes". *Journal of Time Series Analysis*, vol. 28, 261-273.

S. Bordignon, M. Caporin and F. Lisi (2007). "Generalized long memory GARCH for intradaily volatility modelling". *Computational Statistics & Data Analysis*, vol. 51, 5900-5912.

S. Bordignon, M. Caporin and F. Lisi (2009). "Periodic Long-Memory GARCH models." *Econometric Reviews*, vol. 28(1), 60-82.

P.J. Brockwell and R.A. Davis (1991). *Time Series: Theory and Methods*. Second Edition. New York: Springer-Verlag.

J. Davidson (2004). "Moment and Memory Properties of Linear Conditional Heteroscedasticity Model, and a New Model." *Journal of Business & Economic Statistics*, vol. 22, 16-29.

F.X. Diebold and R.S. Mariano (1995). "Comparing Predictive Accuracy". *Journal of Business & Economic Statistics*, vol. 13(3), 134-144.

F.X. Diebold, T.A. Gunther and A.S. Tay (1998). "Evaluating Density Forecasts, with Applications to Financial Risk Management". *International Economic Review*, vol. 39, 863-883.

I.S. Gradshteyn and I.M. Ryzhik (2000). *Table of Integrals, Series, and Products*. New York: Academic Press.

M. Haas, S. Mittnik and M. S. Paolella (2004). "A New Approach to Markov-Switching GARCH Models". *Journal of Financial Econometrics*, vol. 2(4), 493-530.

J. Hosking (1981). "Fractional differencing". *Biometrika*, vol. 68, 165-176.

C.M. Hurvich, E. Moulines and P. Soulier (2005). "Estimating Long Memory in Volatility". *Econometrica*, vol. 73(4), 1283-1328.

F. Klaassen (2002). "Improving GARCH Volatility Forecasts with Regime-Switching GARCH". *Empirical Economics*, vol. 27, 363-394.

P.S. Kokoszka and M.S. Taqqu (1994). "Infinite variance stable ARMA processes". *Journal of Time Series Analysis*, vol. 15, 203-220.

P.S. Kokoszka and M.S. Taqqu (1995). "Fractional ARIMA with stable Innovations." *Stochastic Processes and their Applications*, vol. 60, 19-47.

S.R.C. Lopes (2008). "Long-range Dependence in Mean and Volatility: Models, Estimation and Forecasting", In: V. Sidoravicius and M.E. Vares (eds.), *In and Out of Equilibrium 2 (Progress in Probability)*, vol. 60, 497-525, Birkhäuser.

S.R.C. Lopes and T.S. Prass (2013). "Theoretical Results on Fractionally Integrated Exponential Generalized Autoregressive Conditional Heteroskedastic Processes." In Revision.

(1969) J.A. Mincer and V. Zarnowitz. "The evaluation of economic forecasts", In: J. Mincer (Eds.) *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, New York: Columbia University Press, 3-46.

D.B. Nelson (1991). "Conditional Heteroskedasticity in Asset Returns: A New Approach". *Econometrica*, vol. 59, 347-370.

W.K. Newey and K.D. West (1987). "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix". *Econometrica*, vol. 55(3), 703-708.

W. Palma (2007). *Long-Memory Time Series: Theory and Methods.* New Jersey: John Wiley.

M. Paolella (2013). "Multivariate Asset Return Prediction with Mixture Models". Forthcoming in: *European Journal of Finance.*

T.S. Prass and S.R.C. Lopes (2012). "Risk Measures Estimation on FIEGARCH Processes." Working Paper.

T.S. Prass and S.R.C. Lopes (2013). "VaR, Teste de Estresse e MaxLoss na Presença de Heteroscedasticidade e Longa Dependência na Volatilidade". Forthcoming in: *Revista Brasilieira de Estatística.*

W.F. Stout (1974). *Almost Sure Convergence.* New York: Academic Press.

K.D. West and M.W. McCracken (1998). "Regression-based Tests of Predictive Ability". *International Economic Review*, vol. 39(4), 817-840.