Mao, X., Czellar, V., Ruiz, E., & Veiga, H. (2020). Asymmetric stochastic volatility models: Properties and particle filter-based simulated maximum likelihood estimation. *Econometrics and Statistics*, 13, pp. 84-105.
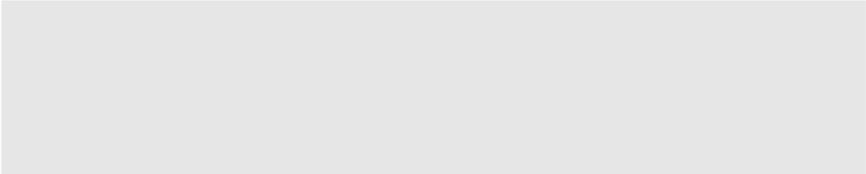
Asymmetric stochastic volatility models: Properties and particle filter-based simulated maximum likelihood estimation

Xiuping Mao [a], Veronika Czellar [b], Esther Ruiz [c, *], Helena Veiga [c, d]

[a] *School of Finance, Zhongnan University of Economics and Law, China*
[b] *EDHEC Business School, France*
[c] *Department of Statistics and Instituto Flores de Lemus, Universidad Carlos III de Madrid, Spain*
[d] *BRU-UNIDE, Portugal*

ABSTRACT

The statistical properties of a general family of asymmetric stochastic volatility (A-SV) models which capture the leverage effect in financial returns are derived providing analytical expressions of moments and autocorrelations of power-transformed absolute returns. The parameters of the A-SV model are estimated by a particle filter-based simulated maximum likelihood estimator and Monte Carlo simulations are carried out to validate it. It is shown empirically that standard SV models may significantly underestimate the value-at-risk of weekly S&P 500 returns at dates following negative returns and overestimate it after positive returns. By contrast, the general specification proposed provide reliable forecasts at all dates. Furthermore, based on daily S&P 500 returns, it is shown that the most adequate specification of the asymmetry can change over time.

1. Introduction

Stochastic Volatility (SV) models, originally proposed by Taylor (1982, 1986) and popularized by Harvey et al. (1994), specify the volatility of financial returns as a latent stochastic process. The original SV models were soon extended to cope with the asymmetric response of volatility to past standardized returns of different signs (leverage effect). For instance, Taylor (1994) and Harvey and Shephard (1996) independently incorporate the leverage effect through correlation between the level and log-volatility innovations. Alternatively, Breidt (1996) and So et al. (2002) introduce asymmetry by allowing the parameters of the log-volatility equation to be different depending on the sign of lagged returns while Asai and McAleer (2011) specify the leverage effect as an EGARCH component; see also McAleer (2005), Yu et al. (2006), Asai (2008), Tsiotas (2012) and Yu (2012a) for alternative asymmetric specifications or combinations of the three basic specifications. With respect to the shape of the conditional distribution, even though normality of standardized returns is still a common assumption, SV models with heavy-tailed and/or asymmetric distributions such as symmetric or skewed Student-t, Generalized Error Distribution (GED) and scale mixtures of normals are increasingly popular; see, for example, Asai and McAleer (2011), Wang et al. (2011, 2013), Feunou and Tédongap (2012), Nakajima and Omori (2012), Tsiotas (2012) and Jensen and Maheu (2014) for some recent references. Note that the GED distribution with parameter ν is described by Harvey (1990) and has the attractiveness of including distributions with different tail thickness as, for example, the extended exponential power

---

distribution of Choy and Walker (2003) when $\nu < 1$, the double exponential when $\nu = 1$, the Gaussian when $\nu = 2$ and the uniform distribution when $\nu = \infty$. The GED distribution has heavy tails if $\nu < 2$.

Although the literature on asymmetric SV models is rather extensive, statistical properties such as existence and analytical expressions of moments are only available in specific SV models with Gaussian standardized returns. The statistical properties of non-normal asymmetric SV models are generally unknown. The main objective of this paper is to fill this gap by deriving the properties of a very general family of asymmetric SV models with non-Gaussian standardized returns, from now on referred to as Generalized Asymmetric SV (GA-SV) family. In particular, we provide general conditions for stationarity and for the existence of integer moments of returns and absolute returns. In addition, we derive analytical expressions of the marginal variance and kurtosis, the autocorrelations of power-transformed absolute returns and the cross-correlations between returns and future power-transformed absolute returns for a given general specification of the volatility with GED standardized returns. Knowledge of these moments could be important to develop tests, as those proposed by Horváth et al. (2006) in the context of GARCH-type models, which are based on the comparison of sample and analytical plug-in moments. Expressions of analytical population moments are also important, for instance, when checking necessary conditions for the asymptotic validity of many estimators.

When dealing with real time series of financial returns, plug-in analytical moments can only be obtained after estimating the parameters. Due to the intractability of the likelihood, parameter estimation in GA-SV models is a rather difficult task. Asymmetric SV models have been estimated via a variety of methods including Efficient Importance Sampling (EIS), Monte Carlo Likelihood (MCL) and Monte Carlo Markov Chain (MCMC) methods. The feasibility and performance of available estimators of parameters and volatilities in the context of the general specification considered in this paper is to be analysed. Meanwhile, the second contribution of this paper is the proposal of a particle filter-based Simulated Maximum Likelihood (SML) procedure with continuous resampling introduced by Malik and Pitt (2011) to estimate the parameters of GA-SV models. We show by Monte Carlo simulations that this estimator provides very accurate estimates of the GA-SV parameters even in moderate sample sizes as small as five hundred observations. Furthermore, the proposed SML estimator is simple to implement and relatively fast on a standard PC.

Finally, the results are illustrated in an empirical application by estimating the volatility of a series of weekly S&P 500 returns. We show that standard SV models significantly underestimate the volatility at dates following negative returns and overestimate it after positive returns. By contrast, GA-SV models accurately estimate the volatility after both negative and positive returns and provide reliable value-at-risk (VaR) forecasts at all dates. Furthermore, we also analyze daily S&P 500 returns and show that the most adequate specification of the log-volatility may be different depending on the particular period of time considered.

The paper is organized as follows. In Section 2, we describe the GA-SV family and derive its statistical properties. Section 3 presents several popular asymmetric SV models contained in the GA-SV family and derives analytical expressions of their moments. Section 4 describes the particle filter-based SML and reports its finite sample moments in Monte Carlo simulations. Section 5 presents an empirical application to weekly S&P 500 returns. Finally, the main conclusions and some guidelines for future research are summarized in Section 6.

2. Statistical properties of GA-SV models

2.1. Model description

Let y_t be the return at time t, σ_t^2 its volatility, $h_t \equiv \log \sigma_t^2$ and ϵ_t be an independent and identically distributed (IID) sequence with mean zero and variance one. The GA-SV family is given by

$$y_t = \exp(h_t/2)\epsilon_t, \qquad t = 1, \ldots, T, \tag{1}$$

$$h_t - \mu = \phi(h_{t-1} - \mu) + f(\epsilon_{t-1}; \theta) + \eta_{t-1}, \tag{2}$$

where the log-volatility disturbance η_t is a Gaussian white noise with variance σ_η^2, $f(\epsilon_t; \theta)$ is a real parametric function of ϵ_t which is independent of η_t for all leads and lags and θ is a vector of parameters. Normality of η_t when $f(\epsilon_t; \theta) = 0$ is justified in Andersen et al. (2001b, 2001a, 2003). The scale parameter μ is related to the marginal variance of returns, while ϕ controls the rate of decay towards zero of the autocorrelations of power-transformed absolute returns, hence, the persistence of volatility shocks. Note that, in Eqs. (1) and (2), the standardized return at time $t - 1$, ϵ_{t-1}, is correlated with the volatility at time t. Furthermore, if $f(\cdot; \theta)$ is not an even function, then positive and negative past standardized returns with the same magnitude have different effects on volatility.

Although the specification of the log-volatility in Eq. (2) considers dependence on only one lag of past log-volatilities, $f(\epsilon_{t-1}; \theta)$ and η_{t-1}, the results obtained in this paper can be extended to models with further lags which can be written as follows $h_t - \mu = \sum_{i=1}^{\infty} \beta_i(f(\varepsilon_{t-i}; \theta) + \eta_{t-i})$ where β_i, $i = 1, 2, \ldots$, are the parameters of the lag polynomial $\beta(L) = (\Phi(L))^{-1}$ with $\Phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$. Moreover, general results can also be obtained when σ_η^2 changes; see, for example, Mao et al. (2017) for a particular specification in which σ_η^2 is allowed to change according to the sign of past returns. We do not consider changes in the variance of the volatility noise as they do not generate leverage effect. Finally, note that the specification of the log-volatility in (2) rules out models in which the persistence, ϕ, is time-varying, such as, the time-varying specification of Yu (2012a).

As mentioned in the Introduction, the statistical properties of models (1) and (2) are only known for particular specifications of $f(\epsilon_t; \theta)$ when ϵ_t is Gaussian. Taylor (1994, 2007) provide expressions for the variance and kurtosis of returns and the autocorrelation function of squared observations for the asymmetric autoregressive SV (A-ARSV) model of Harvey and Shephard (1996) which is obtained when $f(\epsilon_{t-1}; \theta) = \gamma_1 \epsilon_{t-1}$. For the same model and still assuming normality of standardized returns, Pérez et al. (2009) derive the autocorrelation and cross-correlation functions of the power-transformed absolute returns. Asai and McAleer (2011) derive the first and second order moments of returns and Yu (2012a) derives the moments of returns and the conditions for stationarity, strict stationarity and ergodicity in extended specifications of the A-ARSV model. Mao et al. (2017) derive the moments of returns when $f(\epsilon_{t-1}; \theta) = \gamma_1 \epsilon_{t-1} + \alpha I(\epsilon_{t-1} < 0)$ where $I(\cdot)$ is an indicator function that takes value one when the argument is true and zero otherwise. In the next subsection, we provide the moments for the GA-SV family.

2.2. Moments of returns

Theorem 2.1 establishes sufficient conditions for the stationarity of y_t and derives the expression of $E(y_t^c)$ and $E(|y_t|^c)$ for any positive integer c.

Theorem 2.1. *Consider the GA-SV process y_t defined in Eqs. (1) and (2). The process y_t is strictly stationary and ergodic if $|\phi| < 1$ and the variance of $f(\epsilon_t; \theta)$ is finite. Furthermore, if ϵ_t follows a distribution such that both $E(\exp(0.5cf(\epsilon_t; \theta)))$ and $E(|\epsilon_t|^c)$ exist and are finite for some positive integer c, then $\{y_t\}$ and $\{|y_t|\}$ have finite, time-invariant moments of order c given by*

$$E(y_t^c) = E(\epsilon_t^c) E\left(\exp\left(\frac{ch_t}{2} \right) \right), \tag{3}$$

and

$$E(|y_t|^c) = E(|\epsilon_t|^c) E\left(\exp\left(\frac{ch_t}{2} \right) \right), \tag{4}$$

where $E(\exp(\frac{ch_t}{2})) = \exp(\frac{c\mu}{2}) \exp(\frac{c^2 \sigma_\eta^2}{8(1-\phi^2)}) P(0.5c, \phi)$ with $P(a, b) \equiv \prod_{i=1}^{\infty} E(\exp(ab^{i-1} f(\epsilon_{t-i}; \theta)))$.

Proof. See Appendix A.1. □

Theorem 2.1 establishes the strict stationarity of y_t if $|\phi| < 1$ and the variance of $f(\epsilon_t; \theta)$ is finite. Existence of $E(y_t^2)$ is guaranteed if, in addition, $E(\exp(f(\epsilon_t; \theta))) < \infty$. Consequently, under these conditions, y_t is also weakly stationary.

Note that according to expression (3), if ϵ_t has a symmetric distribution, then all odd moments of y_t are zero. Furthermore, from expression (4), it is straightforward to obtain expressions of the marginal variance and kurtosis of y_t as the following corollaries show.

Corollary 2.1.1. *Under the conditions of Theorem 2.1 with $c = 2$ and taking into account that $E(y_t) = 0$, the marginal variance of y_t is directly obtained from (4) as follows*

$$\sigma_y^2 = \exp\left(\mu + \frac{\sigma_\eta^2}{2(1-\phi^2)} \right) P(1, \phi). \tag{5}$$

Corollary 2.1.2. *Under the conditions of Theorem 2.1 with $c = 4$, the kurtosis of y_t can be obtained as $E(y_t^4)/(E(y_t^2))^2$ using expression (4) with $c = 4$ and $c = 2$ as follows*

$$\kappa_y = \kappa_\epsilon \exp\left(\frac{\sigma_\eta^2}{1 - \phi^2} \right) \frac{P(2, \phi)}{(P(1, \phi))^2}, \tag{6}$$

where κ_ϵ is the kurtosis of ϵ_t.

The kurtosis of the basic symmetric Autoregressive SV (ARSV) model considered by Taylor (1982, 1986) and Harvey et al. (1994) is given by $\kappa_\epsilon \exp\left(\frac{\sigma_\eta^2}{1-\phi^2} \right)$. Therefore, this kurtosis is multiplied by the factor $r = \frac{P(2, \phi)}{(P(1, \phi)^2)}$ in the GA-SV family.

Note that the expression of $E(|y_t|^c)$ in (4) depends on both the distribution of ϵ_t and on $f(\cdot; \theta)$. Therefore, in order to obtain closed-form expressions of the variance and kurtosis of returns, one needs to assume a particular distribution of ϵ_t and specify $f(\epsilon_t; \theta)$. We provide analytical expressions for some popular distributions and specifications in Section 3.

2.3. Dynamic dependence

It is easy to see that returns generated from (1) and (2) are a martingale difference. However, they are not serially independent as the conditional heteroscedasticity generates non-zero autocorrelations of power-transformed absolute returns. The following theorem derives the autocorrelation function (acf) of power transformed absolute returns.

Theorem 2.2. *Consider a stationary process y_t defined in Eqs. (1) and (2) with $|\phi| < 1$ and $Var(f(\epsilon_t; \theta)) < \infty$. If ϵ_t follows a distribution such that $E(\exp(0.5cf(\epsilon_t; \theta))) < \infty$ and $E(|\epsilon_t|^c) < \infty$ for some positive integer c, then the τ-th order autocorrelation of $|y_t|^c$ is finite and given by*

$$\rho_c(\tau) = \frac{E(|\epsilon_t|^c)E(|\epsilon_t|^c \exp(0.5c\phi^{\tau-1}f(\epsilon_t; \theta)))\exp\left(\frac{\phi^{\tau}c^2\sigma_\eta^2}{4(1-\phi^2)}\right)P(0.5c(1+\phi^{\tau}), \phi)T_\tau(0.5c, \phi) - [E(|\epsilon_t|^c)P(0.5c, \phi)]^2}{E(|\epsilon_t|^{2c})\exp\left(\frac{c^2\sigma_\eta^2}{4(1-\phi^2)}\right)P(c, \phi) - [E(|\epsilon_t|^c)P(0.5c, \phi)]^2},$$

(7)

where $T_\tau(a, b) \equiv \prod_{i=1}^{\tau-1} E(\exp(ab^{i-1}f(\epsilon_{-i}; \theta)))$ if $\tau > 1$ while $T_1(a, b) \equiv 1$.

Proof. See Appendix A.2. □

Notice that, in practice, most authors dealing with real time series of financial returns focus on the autocorrelations of squares, $\rho_2(\tau)$, which can be obtained from (7) when $c = 2$.

The leverage effect is reflected in the cross-correlations between power-transformed absolute returns and lagged returns. The following theorem gives general expressions of these cross-correlations.

Theorem 2.3. *Consider a stationary process y_t defined in Eqs. (1) and (2) with $|\phi| < 1$ and $Var(f(\epsilon_t; \theta)) < \infty$. If ϵ_t follows a distribution such that $E(\exp(0.5cf(\epsilon_t; \theta))) < \infty$ and $E(|\epsilon_t|^{2c}) < \infty$ for some positive integer c, then the τ-th order cross-correlation between y_t and $|y_{t+\tau}|^c$ for $\tau > 0$ is finite and given by*

$$\rho_{c1}(\tau) = \frac{E(|\epsilon_t|^c)\exp\left(\frac{2c\phi^{\tau}-1}{8(1-\phi^2)}\sigma_\eta^2\right)E(\epsilon_t \exp(0.5c\phi^{\tau-1}f(\epsilon_t; \theta)))P(0.5(1+c\phi^{\tau}), \phi)\frac{T_\tau(0.5c, \phi)}{\sqrt{P(1,\phi)}}}{\sqrt{E(|\epsilon_t|^{2c})\exp\left(\frac{c^2\sigma_\eta^2}{4(1-\phi^2)}\right)P(c, \phi) - [E(|\epsilon_t|^c)P(0.5c, \phi)]^2}}.$$

(8)

Proof. See Appendix A.3. □

3. Nested SV models

Closed-form expressions of the moments of returns can be obtained for appropriate choices of the function $f(\cdot; \theta)$ and of the distribution of ϵ_t. In this section, we derive and illustrate these moments for a very general specification of the volatility when ϵ_t follows a GED distribution. The GED distribution has often been considered in the context of symmetric SV models; see, for example, Andersen (1996), Liesenfeld and Jung (2000) and Cappuccio et al. (2004, 2006). Note that these two latter works consider a skewed GED distribution. Another very popular distribution in the context of SV models is the Student-ν distribution; see, for example, Wang et al. (2011) and Nakajima and Omori (2012). However, closed-form expressions of the moments of returns cannot be obtained in this case given that $\exp(-\mu h_t)$ and $\exp(-\mu/2)y_t$ fail to have finite unconditional means and variances; see Nelson (1991) for the same problem in the context of EGARCH models. Furthermore, we also consider some of the most popular specifications nested in the GA-SV model and discuss their ability to match the empirical properties often observed in time series of financial returns.

Consider the GA-SV model defined in Eqs. (1) and (2) with ϵ_t having a centered and standardized *GED(ν)* distribution and

$$f(\epsilon_t; \theta) = \alpha(I(\epsilon_t < 0) - 0.5) + \gamma_1\epsilon_t + \gamma_2(|\epsilon_t| - E(|\epsilon_t|)).$$

(9)

We denote this model as Threshold GA-SV (TGA-SV); see Asai et al. (2012) for a similar specification allowing for long-memory. The TGA-SV model nests several popular models capturing asymmetric volatility in the context of SV models. For example, when $\alpha = \gamma_2 = 0$ and ϵ_t follows a Gaussian distribution, we obtain the A-ARSV model of Taylor (1994) and Harvey and Shephard (1996). On the other hand, when $\alpha = 0$, the model reduces to the EGARCH plus error model of Demos (2002) and Asai and McAleer (2011), denoted as E-SV and it is similar to that proposed by Asai and McAleer (2005). Asai and McAleer (2005) consider the specification in Eq. (9) with $\alpha = 0$ defined in terms of the observables y_t instead of the unobservables ϵ_t to facilitate estimation. However, this specification makes it difficult to derive the statistical properties of financial returns. When $\gamma_2 = 0$, the specification is similar to the asymmetric SV model proposed by Asai and McAleer (2006). Finally, when $\gamma_1 = \gamma_2 = 0$, the specification gives the threshold model where only the constant changes depending on the sign of past returns; see the recent discussion by Mao et al. (2017) on Threshold SV models.

Note that, if $\gamma_1 < 0$ and $\alpha > 0$, the leverage effect is always satisfied as, in this case, $2\gamma_1 < \alpha$ and, consequently, a positive shock to returns always has a smaller effect on the log-volatility than a negative shock of the same magnitude.

The second moment of a *GED(ν)* variable is finite when $0 < \nu < \infty$. Consequently, when ϵ_t is *GED(ν)*, the variance of $f(\epsilon_t; \theta)$ is finite. Therefore, according to Theorem 2.1, if $|\phi| < 1$, the return process defined in (1) and (2) with $f(\epsilon_t; \theta)$ defined in (9) and ϵ_t having a GED distribution is strictly stationary and ergodic. Furthermore, if $\nu > 1$, then a closed-form expression of $E(|y_t|^c)$ can be derived; see Appendix B.1 for the corresponding expectations. In particular, the marginal variance and

kurtosis of y_t are given by Eqs. (5) and (6) with

$$P(1, \phi) = \prod_{i=1}^{\infty} \exp(-0.5\phi^{i-1}\alpha) \exp\left(-\frac{\phi^{i-1}\gamma_2\Gamma(2/\nu)}{\sqrt{\Gamma(3/\nu)\Gamma(1/\nu)}}\right) \left\{ \sum_{k=0}^{\infty} \left(\left(\frac{\Gamma(1/\nu)}{\Gamma(3/\nu)}\right)^{k/2} \frac{\Gamma((k+1)/\nu)}{2\Gamma(1/\nu)k!} \phi^{k(i-1)} \right. \right.$$
$$\left. \left. \times \left[(\gamma_1 + \gamma_2)^k + \exp(\alpha\phi^{i-1})(\gamma_2 - \gamma_1)^k \right] \right) \right\}, \tag{10}$$

and

$$P(2, \phi) = \prod_{i=1}^{\infty} \exp(-\phi^{i-1}\alpha) \exp\left(-\frac{2\phi^{i-1}\gamma_2\Gamma(2/\nu)}{\sqrt{\Gamma(3/\nu)\Gamma(1/\nu)}}\right) \left\{ \sum_{k=0}^{\infty} \left(\left(\frac{\Gamma(1/\nu)}{\Gamma(3/\nu)}\right)^{k/2} \frac{\Gamma((k+1)/\nu)}{\Gamma(1/\nu)k!} 2^{k-1} \phi^{k(i-1)} \right. \right.$$
$$\left. \left. \times \left[(\gamma_1 + \gamma_2)^k + \exp(2\alpha\phi^{i-1})(\gamma_2 - \gamma_1)^k \right] \right) \right\}, \tag{11}$$

where $\Gamma(\cdot)$ is the Gamma function. Note that, in order to compute $P(\cdot, \cdot)$, one needs to truncate the corresponding infinite product and summation. Our experience is that truncating the product at $i = 500$ and the summation at $k = 1000$ gives very stable results.

Given that the Gaussian distribution is a particular case of the GED distribution when $\nu = 2$, closed-form expressions of $E(|y_t|^c)$ can also be obtained in this case; see Appendix B.2 for the corresponding expectations. Note that these expectations are derived by Demos (2002) when $c = 0$ or 1 and $\alpha = 0$. In particular, the marginal variance is given by expression (5) while the kurtosis is given by expression (6) with

$$P(1, \phi) = \prod_{i=1}^{\infty} \exp(-0.5\phi^{i-1}\alpha) \exp\left(-\phi^{i-1}\gamma_2\sqrt{2/\pi}\right) \left\{ \exp\left(\alpha\phi^{i-1} + \frac{\phi^{2i-2}(\gamma_1 - \gamma_2)^2}{2}\right) \Phi(\phi^{i-1}(\gamma_2 - \gamma_1)) \right.$$
$$\left. + \exp\left(\frac{\phi^{2i-2}(\gamma_1 + \gamma_2)^2}{2}\right) \Phi(\phi^{i-1}(\gamma_2 + \gamma_1)) \right\}, \tag{12}$$

and

$$P(2, \phi) = \prod_{i=1}^{\infty} \exp(-\phi^{i-1}\alpha) \exp\left(-2\phi^{i-1}\gamma_2\sqrt{2/\pi}\right) \left\{ \exp\left(2\alpha\phi^{i-1} + 2\phi^{2i-2}(\gamma_1 - \gamma_2)^2\right) \Phi(2\phi^{i-1}(\gamma_2 - \gamma_1)) \right.$$
$$\left. + \exp\left(2\phi^{2i-2}(\gamma_1 + \gamma_2)^2\right) \Phi(2\phi^{i-1}(\gamma_2 + \gamma_1)) \right\}, \tag{13}$$

where $\Phi(\cdot)$ is the Gaussian cumulative distribution function. The expressions of the marginal variance derived by Taylor (1994) for the classical asymmetric SV model of Harvey and Shephard (1996), by Ruiz and Veiga (2008) for the asymmetric long-memory SV model, by Asai and McAleer (2011) for their general specification and by Mao et al. (2017) for a Threshold SV model, can be obtained as particular cases of expression (5) with $P(1, \phi)$ defined as in (12).

When $\nu < 1$, we cannot obtain analytical expressions of $E(|y_t|^c)$. However, in Appendix B.1, we show that, in this case, $E(|y_t|^c)$ in Eq. (4) is finite if $\gamma_2 + \gamma_1 \leq 0$ and $\gamma_2 - \gamma_1 \leq 0$. It is important to note that the same conditions have to be satisfied for the finiteness of $E(|y_t|^c)$ when ϵ_t follows a Student-t distribution with $d > 2$ degrees of freedom. Finally, if $\nu = 1$, the conditions for the existence of $E(|y_t|^c)$ are $\gamma_2 + \gamma_1 < 2\sqrt{2}/c$ and $\gamma_2 - \gamma_1 < 2\sqrt{2}/c$.

The expectations needed to obtain closed-form expressions of the autocorrelations in expression (7) and cross-correlations in (8) have been derived in Appendix B.1 for the TGA-SV model with a GED distribution with parameter $\nu > 1$ and in Appendix B.2 for the particular case of Gaussian errors, i.e. $\nu = 2$. As above, when $\nu \leq 1$, we can only obtain conditions for the existence of the autocorrelations and cross-correlations. The codes to evaluate the moments, autocorrelations and cross-correlations are available online.

We now illustrate the influence of the leverage parameters on these moments. We consider a TGA-SV model with Gaussian errors and plot, in the left panel of Fig. 1, the first order autocorrelations of squared returns, namely, $\rho_2(1)$, as a function of the leverage parameters, γ_1 and α, when $\phi = 0.98$, $\sigma_\eta^2 = 0.05$ and $\gamma_2 = 0$ and 0.1. We observe that the autocorrelations of squared returns are larger for larger values of γ_2. However, both surfaces are rather flat and, consequently, the leverage parameters do not have large effects on the first order autocorrelations of squares.

In the right panel of Fig. 1, we illustrate the effect of the parameters on the cross-correlations between y_t and y_{t+1}^2, denoted by $\rho_{21}(1)$. First order cross-correlations between returns and future squared returns are indistinguishable for the two values of γ_2 considered. Furthermore, Fig. 1 shows that $|\gamma_1|$ drags $\rho_{21}(1)$ in an approximately linear way while the effect of α is nonlinear. Also note that the effect of the threshold, α, on $\rho_{21}(1)$ is relevant when $\gamma_1 = 0$. However, as γ_1 increases, the role of α to explain the leverage effect decreases.

The first column of Fig. 2 illustrates the autocorrelation function (acf) of squared returns (top panel) and the cross-correlation function (ccf) between returns and future squared returns (bottom panel) in a TGA-SV model with parameters $\phi = 0.98$, $\sigma_\eta^2 = 0.05$, $\alpha = 0.07$, $\gamma_1 = -0.08$, $\gamma_2 = 0.1$ and $\nu = 1.5, 1.7, 2$ and 2.5. As expected, the acf of y_t^2 has an exponential decay. Furthermore, fatter tails of ϵ_t imply smaller autocorrelations of squared returns; see Carnero et al. (2004) for similar conclusions in the context of symmetric SV models. The ccf plotted in the bottom panel of the first column shows that the parameter ν of the GED distribution has a very mild influence on the cross-correlations.
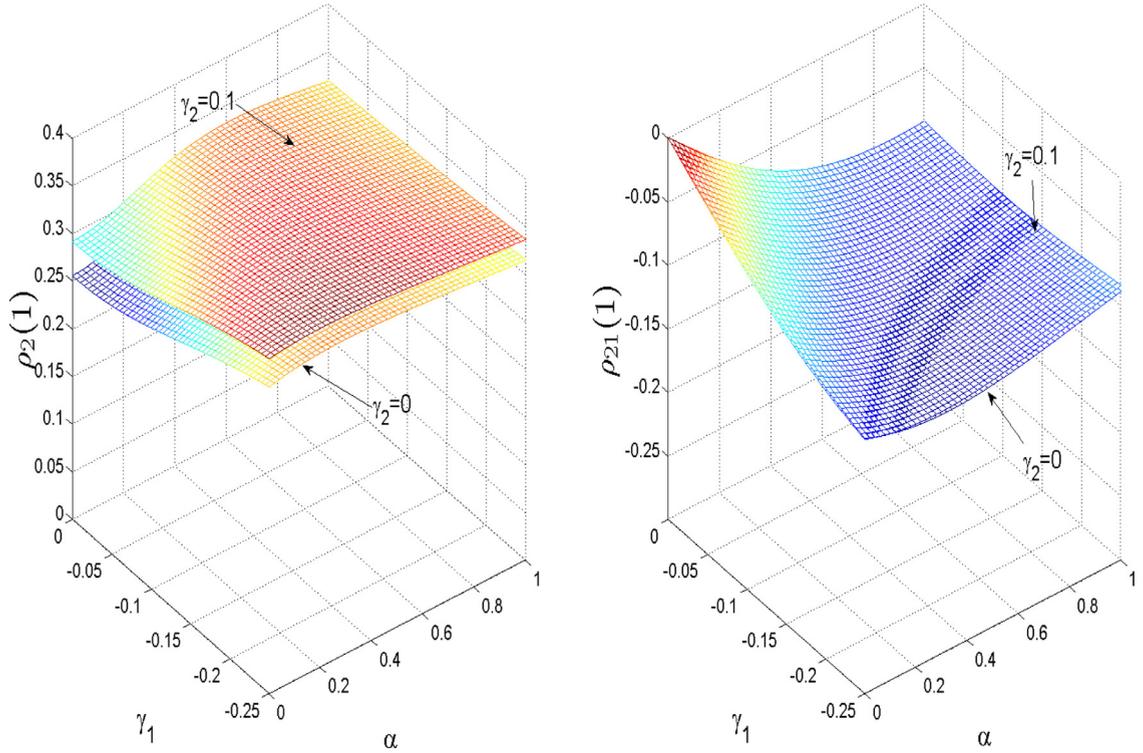
Fig. 1. First order autocorrelations of squares (left panel) and first order cross-correlations between returns and future squared returns (right panel) of different Gaussian TGA-SV models with parameters $\phi = 0.98$ and $\sigma_\eta^2 = 0.05$.

The second column of Fig. 2 plots the acf and ccf of an A-ARSV model with $\alpha = \gamma_2 = 0$ and otherwise the same parameters as in the TGA-SV model above. We can observe that the autocorrelations of squared returns and the absolute cross-correlations are slightly smaller than those of the corresponding TGA-SV model. Therefore, including γ_2 and α in the TGA-SV model allows for stronger volatility clustering and leverage effect.

The third column of Fig. 2 plots the autocorrelations and cross-correlations for an E-SV model with $\alpha = 0$ and otherwise, the same parameter values as in the TGA-SV model considered above. Comparing the plots of the A-ARSV and E-SV models in Fig. 2, we can observe that adding $|\epsilon_{t-1}|$ into the A-ARSV model generates larger autocorrelations of squared returns. However, as expected, the cross-correlations are almost identical. Therefore, the E-SV model represents wider dependencies of volatility than the A-ARSV model but not more volatility leverage.

Fig. 2 shows that, given a particular E-SV model, we may find an A-ARSV model with almost the same autocorrelations and cross-correlations. For instance, the autocorrelations in the E-SV model with $\nu = 2$ are very similar to those in the A-ARSV model with $\nu = 2.5$. Furthermore, the cross-correlations are indistinguishable in any case. Nevertheless, these two models generate returns with different kurtoses. Therefore, if the parameter ν is free, it is difficult to identify the parameters γ_2 and σ_η^2 using the information of the autocorrelations and cross-correlations only. However, the distribution of returns and, in particular the kurtosis, implied by both models is different and this information can be used to identify these parameters.

The autocorrelations in the TGA-SV and E-SV models are almost identical. Including α only has a paltry effect on volatility clustering. However, the cross-correlations are different and show that α allows for more leverage effect.

The last specification of $f(\cdot; \theta)$ considered in this paper specifies the log-volatility with different constant levels depending on the sign of past returns as $f(\epsilon_t; \theta) = \alpha[I(\epsilon_t < 0) - 0.5]$. This specification has been previously considered by Asai and McAleer (2006) and it is a restricted version of the Threshold SV model proposed by Breidt (1996) and So et al. (2002); see Mao et al. (2017) for the relevance of this specification when compared to the more general Threshold SV model. Hereafter, we refer to this model as restricted Threshold SV (RT-SV). The last column of Fig. 2 illustrates the shape of the autocorrelations of squared returns and the cross-correlations between returns and future squared returns for a RT-SV model with the same values of the parameters as in the TGA-SV above and $\gamma_1 = \gamma_2 = 0$. Comparing the autocorrelations of squared returns in the TGA-SV and the RT-SV models represented in the top panel, we observe that the latter are slightly smaller than the former. However, the absolute cross-correlations in the RT-SV model are the smallest among all the models considered. The presence of α in the TGA-SV model seems to reinforce the role of the leverage parameter γ_1.
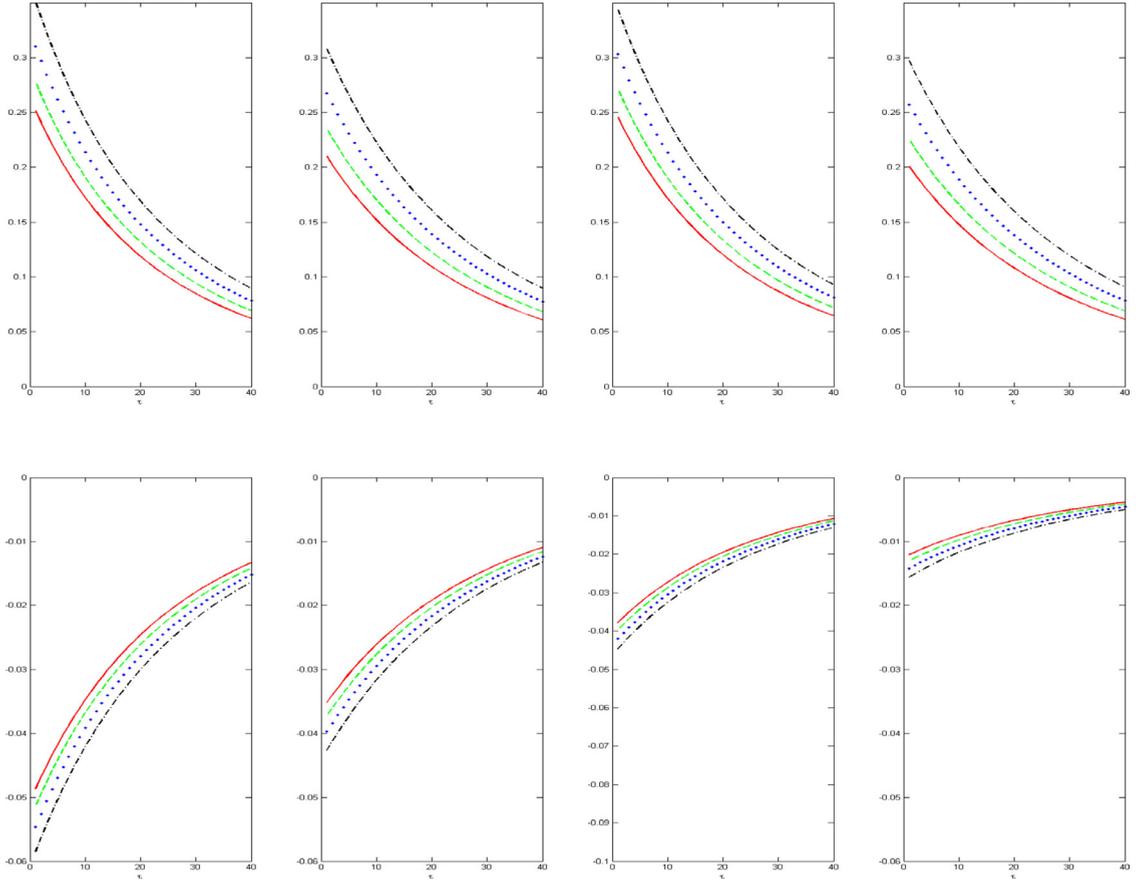
Fig. 2. Autocorrelation function of squares (first row) and cross-correlations between returns and future squared returns (second row) for different specifications of asymmetric SV models. The first column corresponds to a TGA-SV model with $\alpha = 0.07, \phi = 0.98, \sigma_\eta^2 = 0.05, \gamma_1 = -0.08, \gamma_2 = 0.1$ and $\nu = 1.5$ (solid lines), $\nu = 1.7$ (dashed lines), $\nu = 2$ (dotted lines) and $\nu = 2.5$ (dashdot lines). The second column corresponds to the A-ARSV model with $\alpha = \gamma_2 = 0$. The third column corresponds to an E-SV model with $\alpha = 0$. Finally, the last column corresponds to the RT-SV model with $\gamma_1 = \gamma_2 = 0$.

In order to illustrate the effect of past shocks, ϵ_{t-1} and η_{t-1}, on the volatility at time t, in the TGA-SV model, Fig. 3 plots the following volatility surface

$$V_t = \exp((1 - \phi)\mu)\sigma_y^{2\phi} \exp\left(f(\epsilon_{t-1}; \theta) + \eta_{t-1}\right), \tag{14}$$

when $\phi = 0.98$, $\sigma_\eta^2 = 0.05$, $\exp(\mu) = 0.01$ and $\nu = 2$. We consider the specification of $f(\epsilon_t; \theta)$ in Eq. (9) with a) $(\alpha, \gamma_1, \gamma_2) = (0.07, -0.08, 0.1)$ (top panel), b) $(\alpha, \gamma_1, \gamma_2) = (0, -0.08, 0)$ (middle panel) and c) $(\alpha, \gamma_1, \gamma_2) = (0, -0.08, 0.1)$ (bottom panel). Note that the shape of V_t does not depend on the distribution of ϵ_t as it is a function of $f(\epsilon_{t-1}; \theta)$ and η_{t-1}. Note that the only dependence is through the constant $E(|\epsilon_t|)$ which affects the marginal variance of returns. The response of volatility in (14) is obtained from (2) by evaluating the lagged volatilities at the marginal variance and assuming that at time $t-1$ the volatility is equal to an "average" volatility. It is important to point out that, given that ϵ_t and η_t are mutually independent, the relation between the log-volatility, h_t, and ϵ_{t-1} is measured just by the covariance between $f(\epsilon_{t-1}; \theta)$ and ϵ_{t-1}. However, when looking at the more interesting relation between the volatility and ϵ_{t-1}, we observe that the leverage effect depends on η_{t-1}. The volatility surface in Eq. (14) is different from the News Impact Curve proposed by Yu (2012a) and Takahashi et al. (2013) in that they use numerical methods to compute the conditional variance. In (14), we are more interested in describing how the volatility, instead of the conditional variance, responds to the level and volatility shocks.

The top panel of Fig. 3 plots V_t for the TGA-SV model when the parameters are given by the values in (a). We can observe that V_t shows a discontinuity due to the presence of the indicator function in $f(\epsilon_t; \theta)$. Furthermore, the leverage effect of the TGA-SV model is different depending on the values of the log-volatility shock, η_{t-1}. It is visible when η_{t-1} is positive. However, when η_{t-1} is negative, the leverage effect is weaker. The second panel of Fig. 3 plots the volatility response when the parameters are given by (b), corresponding to an A-ARSV model. In this case, we observe that the leverage effect is very weak for negative log-volatility shocks. Finally, the bottom panel of Fig. 3 plots the volatility response when the parameter
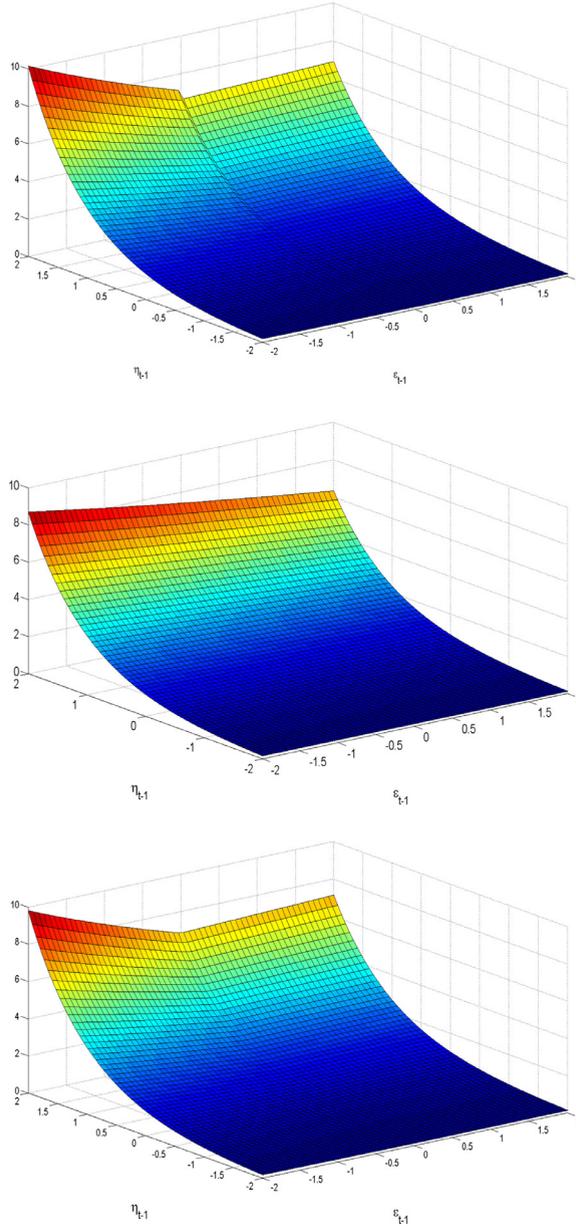
Fig. 3. Volatility responses of GA-SV models with $\phi = 0.98$, $\sigma_\eta^2 = 0.05$ $\exp(\mu) = 0.01$ and $\nu = 2$. Top panel $f(\epsilon_t; \theta) = \alpha(I(\epsilon_t < 0) - 0.5) + \gamma_1\epsilon_t + \gamma_2(|\epsilon_t| - E(|\epsilon_t|))$; middle panel $f(\epsilon_t; \theta) = \gamma_1\epsilon_t$ and bottom panel $f(\epsilon_t; \theta) = \gamma_1\epsilon_t + \gamma_2(|\epsilon_t| - E(|\epsilon_t|))$. The parameter values are $\{\alpha, \gamma_1, \gamma_2\} = \{0.07, -0.08, 0.1\}$.

values are given by the values in (c), corresponding to the E-SV model. In this case, there is no discontinuity but the effect of ϵ_{t-1} on V_t still depends on η_{t-1}.

4. Model estimation

As shown in the previous section, GA-SV models are attractive because of their flexibility to mimic the dynamic properties of financial returns and because they incorporate several popular specifications of the leverage effect often considered before in the literature. The downside to adopting an appealing GA-SV model is the difficulty involved with parameter estimation and volatility forecasts. In this section, we briefly describe alternative estimators of SV models and propose estimating the parameters of GA-SV models by combining a particle filter-based SML procedure with continuous resampling introduced by Malik and Pitt (2011). We analyze the finite sample properties of the proposed estimator and compared them with those of a MCMC estimator implemented using a friendly and accesible software.

4.1. Estimators of stochastic volatility models

The likelihood function of SV models is expressed by a high dimension integral which cannot be solved analytically due to the presence of the unobserved volatility. As a result, maximum likelihood estimation of the parameters of SV models is not trivial. In the context of asymmetric SV models, the intractability of the likelihood is inherited from standard SV models; see Broto and Ruiz (2004) and Yu (2012b) for surveys on estimation of SV models. Asymmetric SV models have been estimated via a variety of methods including EIS, MCL and MCMC methods. The EIS estimator, originally proposed by Liesenfeld and Richard (2003) and Richard and Zang (2007), evaluates the likelihood function numerically by integrating out the latent volatility process via importance sampling techniques, followed by numerical maximization of the approximated likelihood function. Consequently, it requires finding an important sampler such that the variance of the importance weights exists; see Koopman et al. (2009) for conditions for the existence of the variance. Scharth and Kohn (2016) propose an alternative procedure to avoid the exponential growth of the variance as the dimension of the integration increases. Even if the existence conditions are satisfied, one has to select a class of auxiliary samplers such that they are close to the original model and it is easy to sample from them. Finding the auxiliary samplers in the case of GA-SV models is a challenging task. Very recently, Koopman et al. (2019) propose a computational simplification of the EIS estimator based on using standard Kalman filter and smoothing methods. This modification avoids computing (possibly) complicated density kernels that need to be tracked for every time period. However, as pointed out by Huang and Yu (2008), a drawback of EIS-based estimators is that they may not be easily generalized to asymmetric SV models due to the nonlinear specification of the unobserved log-volatility process.

Bayesian imputation via MCMC is usually more complicated. One exception is the MCMC estimator implemented in Meyer and Yu (2000), Yu (2005), Yu (2012a) and Wang et al. (2013) that is friendly and easy to implement. However, it is based on a single-move Gibbs sampling algorithm that it is known to have problems to converge to the conditional posterior, severely impacting the statistical properties of the resulting parameter estimators; see Kim et al. (1998). New Bayesian estimators have recently been proposed in the literature. They estimate in a fairly precise way the parameters of several SV models. For instance, Andrieu et al. (2010) propose a particle Markov chain Monte Carlo (PMCMC) that combines MCMC with sequential Monte Carlo methods that are able to cope with complex patterns of dependence. More recently, Kastner and Fruhwirth-Schnatter (2014) show that an interweaving strategy can greatly improve sampling efficiency for all parameters of some SV models. Shestopaloff and Neal (2014) show that ensemble MCMC sampling methods for Bayesian computations are much more efficient that the interweaving strategy since the later relies on the latent process being linear and Gaussian. Alternatively, McCausland et al. (2011) propose using the efficient Cholesky Factor Algorithm to draw states and compare several estimators concluding that there is not a dominant method.

Finally, to our knowledge, estimation of a general specification of the GA-SV family with GED standardized returns has not been considered in the literature yet. Consequently, in this paper, we implement a relatively simple and fast estimator of the parameters of GA-SV models based on a particle filtering SML procedure with continuous resampling as proposed by Malik and Pitt (2011). Although MCMC methods are very powerful for estimating nonlinear and non Gaussian state space models, the choice of a prior distribution can often influence the performance of the sampler. For instance, requiring the sampler to take many iterations to converge from the starting value to the target distribution; see van Ravenzwaaij et al. (2018). It is also true that particle filtering methods can be inefficient in the presence of highly informative observations. However, in the context of SV models, the observations are not very informative individually and, consequently, this is not a concern; see Pitt et al. (2014). According to our Monte Carlo experiments, the SML estimator can be very accurate even with sample sizes as small as $T = 500$ observations.

4.2. Simulated ML estimation of GA-SV models via particle filtering

Consider a set of observations $y_{1:T} = (y_1, \ldots, y_T)$ generated from the TGA-SV model in Eqs. (1) and (2) with $|\phi| < 1$, ϵ_t having a centered and standardized $GED(\nu)$ distribution and $f(\epsilon_{t-1}; \theta)$ given by (9). At date $t = 0$, the initial log-volatility distribution is assumed to be given by $h_0 \sim \mathcal{N}(\mu, \frac{\sigma_\eta^2}{1-\phi^2})$. Note that, in the TGA-SV model, the marginal variance of h_0 is given by $Var(h_0) = \frac{1}{1-\phi^2}(\sigma_\eta^2 + 0.25\alpha^2 + \gamma_1^2 + \gamma_2^2 - \frac{\Gamma(2/\nu)}{\Gamma(1/\nu)\Gamma(3/\nu)}(\alpha\gamma_1\sqrt{\Gamma(1/\nu)\Gamma(3/\nu)} + \gamma_2\Gamma(2/\nu)))$. However, using $\sigma_\eta^2/(1 - \phi^2)$ as variance of the initial distribution does not have effects on the final estimates. Since $\epsilon_{t-1} = y_{t-1}\exp(-h_{t-1}/2)$, Eq. (2) can be rewritten as:

$$h_t - \mu = \phi(h_{t-1} - \mu) + f[y_{t-1}\exp(-h_{t-1}/2); \theta] + \eta_{t-1}. \tag{15}$$

Denote by $\theta^* = (\theta, \mu, \phi, \sigma_\eta^2)$ the vector of all parameters in the log-volatility equation, (2) except the parameter of the distribution of ε_t, ν, and by $g_y(y_t|h_t; \nu)$ the density of returns, y_t, conditional on contemporaneous unobserved log-volatilities, h_t, which is given by the GED density with parameter ν. Finally, $g_h(h_t|h_{t-1}, y_{t-1}; \theta^*)$ is the conditional density of log-volatility given past log-volatilities and observations, which is given by $\mathcal{N}(\mu(1 - \phi) + \phi h_{t-1} + f[y_{t-1}\exp(-h_{t-1}/2); \theta], \sigma_\eta^2)$.

We estimate $\theta = (\theta^*, \nu)$ by SML where the likelihood is estimated by the particle filtering iterative algorithm described below; see Gordon et al. (1993). Given θ and y_0, at date $t = 0$ particles of the initial log-volatilities, $\{h_0^{(n)}\}_{n=1}^N$, are generated from the initial log-volatility distribution. Then, for $t > 0$, iterate the following three steps:

Step 1 (Sampling): For every $n = 1, \ldots, N$, simulate a new state $\tilde{h}_t^{(n)}$ from $g_h(\cdot \mid h_{t-1}^{(n)}, y_{t-1}; \theta^*)$.

Step 2 (Correction): Given the new data point y_t, compute

$$w_t^{(n)}(\theta) = g_y(y_t \mid \tilde{h}_t^{(n)}; \nu), \qquad n \in \{1, \ldots, N\}.$$

Step 3 (Selection): For every $n = 1, \ldots, N$, draw $h_t^{(n)}$ from $\tilde{h}_t^{(1)}, \ldots, \tilde{h}_t^{(N)}$ with importance weights $p_t^{(1)}(\theta), \ldots, p_t^{(N)}(\theta)$ where $p_t^{(n)}(\theta) = w_t^{(n)}(\theta)/\sum_{i=1}^{N} w_t^{(i)}(\theta)$.

Given θ, the true but intractable log-likelihood function can be estimated by

$$\log \hat{L}(\theta) = \sum_{t=1}^{T} \log \left(\frac{1}{N} \sum_{n=1}^{N} w_t^{(n)}(\theta) \right). \tag{16}$$

Since the simulated estimator $\hat{L}(\theta)$ is a consistent and unbiased estimator of $L(\theta^*)$ (Del Moral, 2004, Chapter 9), θ^* can be estimated by maximizing the estimated log-likelihood as follows:

$$\hat{\theta}_{SML}^* = \arg \max_{\theta^*} \log \hat{L}(\theta). \tag{17}$$

The estimator in (17) is a SML estimator; see Gouriéroux and Monfort (1996, Chapter 3). When T and $N \to \infty$ and $\sqrt{T}/N \to 0$, it is asymptotically equivalent to the maximum likelihood estimator; see Malik and Pitt (2011).

In order to facilitate maximization of the log-likelihood in (16), it is important that its estimator is continuous as a function of the parameters. With this purpose, in step 1, we use the continuous resampling proposed by Malik and Pitt (2011) for univariate hidden state processes as follows

Step (i) (Sorting): Sort the particles $\{\tilde{h}_t^{(n)}\}$ in ascending order

$$\hat{h}_t^{(1)} \leq \hat{h}_t^{(2)} \leq \cdots \leq \hat{h}_t^{(N)}.$$

Denote the importance weight associated with $\hat{h}_t^{(n)}$ by $\hat{p}_t^{(n)}$.

Step (ii) (Sorted uniforms): Draw sorted $\{u_i\}$ using Malmquist's (1950) sorted uniforms. That is draw $\{v_i\}$ from a uniform distribution on the interval (0,1). Set $u_N = v_N^{1/N}$ and for $n = N - 1, \ldots, 1$:

$$u_n = u_{n+1} v_n^{1/n}.$$

The resulting $u_1 < u_2 < \cdots < u_N$ are ordered uniform variates.

Step (iii) (Index sampling): Define $\pi_t^{(0)} = \hat{p}_t^{(1)}/2$, $\pi_t^{(N)} = \hat{p}_t^{(N)}/2$ and $\pi_t^{(k)} = (\hat{p}_t^{(k+1)} + \hat{p}_t^{(k)})/2$ for $k = 1, \ldots, N - 1$, and set $s = 0$, $j = 1$. For $i = 0, \ldots, N$, sample indexes $r_1, \ldots, r_N$ and a new set of uniforms $u_1^*, \ldots, u_N^*$:

$$s = s + \pi_t^{(i)}$$
$$\text{while } (u_j \leq s \text{ and } j \leq N) \, \{$$
$$\quad r_j = i$$
$$\quad u_j^* = (u_j - (s - \pi_t^{(i)}))/\pi_t^{(i)}$$
$$\quad j = j + 1\}.$$

Step (iv) (Particle drawing): For the selected regions where $r_j = 0$, we set $h_t^{(j)} = \hat{h}_t^{(1)}$ and where $r_j = N$, we set $h_t^{(j)} = \hat{h}_t^{(N)}$. Otherwise, set

$$h_t^{(j)} = (\hat{h}_t^{(r_j+1)} - \hat{h}_t^{(r_j)}) \times u_j^* + \hat{h}_t^{(r_j)}.$$

Note that it is important to guarantee that, for a fixed seed, the random number generator provides continuously changing numbers in the parameters. For instance, the inverse cdf method to generate continuous random numbers satisfies this property.

Finally, also note that there is not weight degeneracy as step 3 of the particle filtering is carried out at each moment of time t.

4.3. Finite sample performance of the particle filter-based simulated ML estimator

In this section, we carry out Monte Carlo experiments to analyze the finite sample performance of the particle filter-based SML estimator of the parameters of the TGA-SV model. We consider two designs for the Monte Carlo experiments. First, ϵ_t has a centered normal distribution and ν is treated as known while all other parameters in the model are estimated. Second, we consider the TGA-SV model with GED standardized returns with $\nu = 1.5$. In this case, the parameter ν is also estimated. The Monte Carlo results are based on $R = 500$ replicates of sizes $T = 500$, 1000 and 2000 generated by each of these TGA-SV models with parameters $(\mu, \phi, \alpha, \gamma_1, \gamma_2, \sigma_\eta^2) = (0, 0.98, 0.07, -0.08, 0.1, 0.05)$ and $N = 10^4$ particles. For computational efficiency, we choose to estimate $\mu \cdot (1 - \phi)$ instead of μ as it leads to a more stable numerical optimization

Table 1

Monte Carlo results of the particle filter-based SML estimator of the TGA-SV parameters based on $N = 10^4$ particles. The table reports Monte Carlo averages and standard deviations (in parenthesis) of 500 estimates of the TGA-SV-N (left panel) and TGA-SV-G (right panel) models, and the mean and standard deviation of computing times. True parameters are in bold.

	TGA-SV-N							TGA-SV-G							
True	$\mu(1-\phi)$ **0**	ϕ **0.98**	α **0.07**	γ_1 **−0.08**	γ_2 **0.1**	σ_η^2 **0.05**	**Time**	$\mu(1-\phi)$ **0**	ϕ **0.98**	α **0.07**	γ_1 **−0.08**	γ_2 **0.1**	σ_η^2 **0.05**	ν **1.5**	**Time**
$T = 500$															
Mean	0.012	0.969	0.082	−0.078	0.106	0.042	1194.3	0.009	0.971	0.078	−0.078	0.037	0.026	1.623	1738.1
Std. dev.	(0.043)	(0.016)	(0.080)	(0.051)	(0.144)	(0.025)	(356.5)	(0.054)	(0.018)	(0.088)	(0.059)	(0.287)	(0.034)	(0.416)	(566.8)
$T = 1000$															
Mean	0.007	0.975	0.075	−0.078	0.101	0.047	2341.3	0.006	0.976	0.071	−0.079	0.085	0.035	1.550	3630.6
Std. dev.	(0.030)	(0.009)	(0.060)	(0.038)	(0.093)	(0.016)	(728.7)	(0.030)	(0.009)	(0.065)	(0.041)	(0.198)	(0.028)	(0.275)	(1045.7)
$T = 2000$															
Mean	0.006	0.978	0.066	−0.081	0.103	0.049	4822.0	0.004	0.978	0.072	−0.080	0.099	0.042	1.513	7522.1
Std. dev.	(0.023)	(0.006)	(0.044)	(0.025)	(0.059)	(0.011)	(1436.3)	(0.024)	(0.006)	(0.048)	(0.029)	(0.135)	(0.021)	(0.178)	(2086.2)

Table 2

Monte Carlo results comparing SML and MCMC estimators. $T = 2000$ and normal errors.

	Parameter estimates						Average time
	$\mu(1-\phi)$	ϕ	α	γ_1	γ_2	σ_η^2	
True values	0.006	0.94	0.1	−0.12	0	0.05	
SML	0.059 (0.013)	0.937 (0.012)	0.103 (0.068)	−0.122 (0.036)	0.006 (0.062)	0.049 (0.013)	4301.18 s
MCMC	0.011 (0.038)	0.931 (0.014)	0.106 (0.070)	−0.123 (0.037)	0.014 (0.064)	0.054 (0.014)	260.60 s

when ϕ is close to one, as it is the case in our Monte Carlo designs. It is important to point out that the average computing time involved in the estimation of the TGA-SV model using the particle filter-based SML estimator is below half an hour. Programs were implemented in C and computation was performed on an Intel(R) Xeon(R) CPU E7-4830 v3 @ 2.10 GHz. The C codes are available from the second author. To maximize the log-likelihood, 5 maximizations are run with the initial parameter values chosen randomly in the parameter space. The maximization is carried out using the Downhill simplex procedure and the global maximum of the 5 maximization results is taken as the global maximum.

Table 1 reports the Monte Carlo averages and standard deviations of the particle filter-based SML estimates. We can observe that, regardless of whether the distribution of standardized returns is assumed to be known or not, all the parameters are accurately estimated, even when the sample size is as small as $T = 500$. Note that when ν is estimated as a further parameter, the standard errors of the SML estimator of γ_2 and σ_η^2 increase with respect to those obtained when the distribution of ϵ_t is assumed to be known and normal. These Monte Carlo experiments show that the particle filter-based SML estimator does a remarkable job in estimating the parameters of the general asymmetric SV model considered in this paper.

Finally, Table 1 also reports the means and standard deviations of computing time (in seconds) of the 500 Monte Carlo replicates for the two models and three sample sizes considered. Computing times reported in this table are reasonably small for the SML to be used in practice.

In order to compare the particle filter-based SML estimator with alternative estimators based on Gibbs samplers, we carry out an additional simulation experiment. In this second experiment, the parameters of the DGP are choosen to reproduce those estimated in the empirical application. In particular, 500 replicates of size $T = 2000$ are generated by the TGA-SV model with Normal standardized returns and parameters $\mu = 0.1$, $\phi = 0.94$, $\alpha = 0.1$, $\gamma_1 = -0.12$, $\gamma_2 = 0$ and $\sigma_\eta^2 = 0.05$. Table 2 reports the Monte Carlo means and standard deviations when the parameters are estimated using the SML estimator and the Just Another Gibbs Sampler (JAGS) proposed by Plummer (2003) which is a recent alternative to OpenBugs. The JAGS estimator is implemented with 3 chains each of them having 10000 iterations and a burning of 5000. Table 2 shows that, if the sample size is as large as $T = 2000$, the particle filter SML estimates are slightly closer to the true values and their standard errors are also slightly smaller than those provided by JAGS. Furthermore, the Gibbs estimator requires on average 16.5 times less computer time. The differences between the estimates of the two estimators are greater for smaller sample sizes and benefit the SML estimator.

5. Empirical application

In this section, the results are illustrated by fitting the TGA-SV model and several of its particular specifications to S&P 500 returns. We consider first weekly and then daily returns.
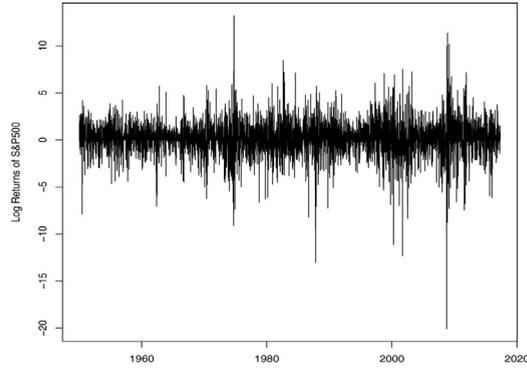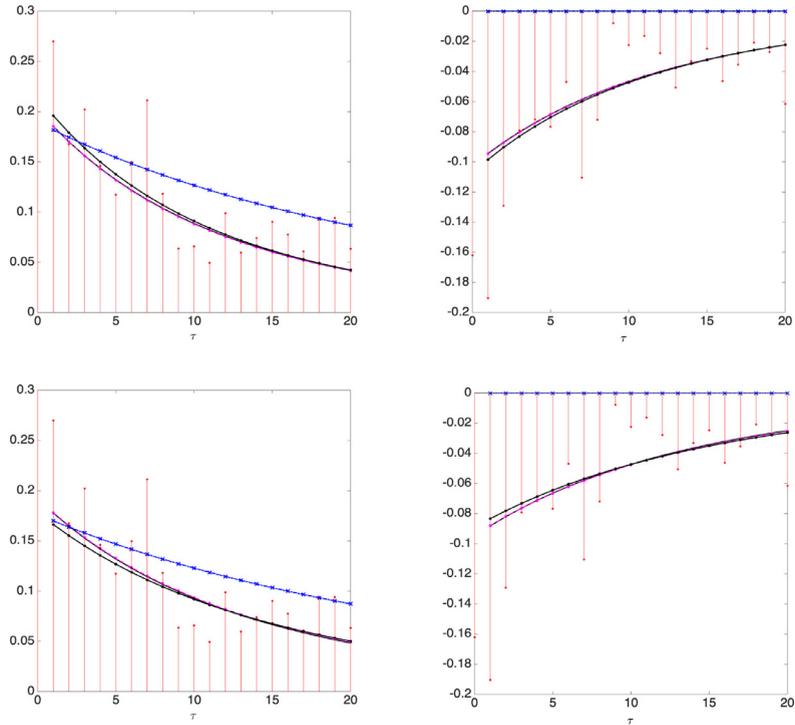
Fig. 4. (a) Weekly S&P 500 returns observed from January 9, 1950 to May 18, 2017. (b) Sample and plug-in autocorrelations (left column) and cross-correlations (right column) for the TGA-SV (solid lines), E-SV (dash lines), A-ARSV (asterisk lines) and ARSV (dash-dotted lines) models when the errors follow GED (first row) and Gaussian distributions (second row).

5.1. Weekly returns

S&P 500 prices are observed weekly between January 9, 1950 and May 8, 2017, with a total of $T = 3513$ observations. The returns, computed as $y_t = 100 \times \triangle \log P_t$, where P_t is the adjusted close price at time t obtained from finance.yahoo.com, are plotted in Fig. 4 which also plots the sample autocorrelations of squares and the cross-correlations between returns and future squared returns. We can observe the usual pattern of volatility clustering and negative and significative cross-correlations between returns and future squared returns.

We fit the TGA-SV, E-SV, A-ARSV and ARSV models with GED (denoted by -G) and Normal (denoted by -N) errors. The parameters are estimated by the particle filter-based procedure described in Section 4 with $N = 10^4$ particles. Table 3 reports the parameter estimates together with their corresponding asymptotic standard errors in parentheses. The particle-filter estimated log-likelihoods are also reported together with the p-values of the likelihood-ratio test which compares the likelihood of each model to that of the TGA-SV-G specification. First of all, we can observe that, regardless of the

Table 3

Empirical parameter estimates and log-likelihood values for weekly S&P 500 returns. Standard errors of parameter estimates are reported in parentheses. *P*-values of the likelihood tests comparing restricted models to the TGA-SV-G model are reported in parentheses below the log-likelihood values.

Model	Parameter estimates							Log $\hat{L}$
	$\mu(1-\phi)$	ϕ	α	γ_1	γ_2	σ_η^2	ν	
TGA-SV-G	0.085 (0.009)	0.936 (0.007)	0.099 (0.052)	−0.116 (0.025)	−0.070 (0.058)	0.051 (0.008)	2.521 (0.186)	−7068.13
E-SV-G	0.083 (0.009)	0.935 (0.007)	–	−0.157 (0.014)	−0.070 (0.056)	0.051 (0.008)	2.523 (0.185)	−7070.06 (0.049)
A-ARSV-G	0.085 (0.009)	0.934 (0.007)	–	−0.158 (0.014)	–	0.048 (0.007)	2.371 (0.125)	−7071.07 (0.053)
ARSV-G	0.035 (0.008)	0.969 (0.007)	–	–	–	0.036 (0.007)	2.119 (0.103)	−7134.63 (0.0000)
TGA-SV-N	0.076 (0.008)	0.945 (0.007)	0.093 (0.050)	−0.106 (0.023)	0.040 (0.041)	0.033 (0.005)	–	−7073.78 (0.001)
E-SV-N	0.073 (0.008)	0.945 (0.007)	–	−0.143 (0.013)	0.038 (0.040)	0.033 (0.005)	–	−7075.60 (0.001)
A-ARSV-N	0.069 (0.007)	0.949 (0.006)	–	−0.140 (0.013)	–	0.033 (0.005)	–	−7076.04 (0.001)
ARSV-N	0.032 (0.008)	0.972 (0.006)	–	–	–	0.031 (0.006)	–	−7135.24 (0.000)

particular log-volatility specification and error distribution, the persistence parameter, ϕ, ranges always between 0.934 and 0.972. However, the estimation of the variance of the log-volatility noise depends on the distribution assumed for standardized returns and on whether the leverage effect is modeled. When standardized returns are normal or the volatility is assumed to be symmetric, the estimated σ_η^2 is always smaller than in models with leverage effect and non-Gaussian standardized returns. With respect to the asymmetric parameters, it is interesting to observe that the threshold parameter is always significant while the estimates of γ_1 are slightly smaller for Gaussian errors and for the TGA-SV specification. The estimates of the EGARCH parameter, γ_2, are not statistically significant. Finally, the estimates of the GED parameter, ν, are significantly larger than 2, when the volatility is asymmetric. However, when there is no leverage effect, the estimated ARSV model is almost the same regardless of whether the errors are normal or not. When looking at the *p*-values reported in Table 3, we can observe that models with normal errors or symmetric volatilities are rejected when compared with the TGA-SV-G model. Therefore, models with GED errors and leverage effect are closer to the true Data Generating Process (DGP) than models with Gaussian errors or symmetric volatilities.

In order to analyze the effect of these estimates on the moments that each of the models can explain, we compute the plug-in autocorrelations of squares and cross-correlations, obtained when the estimated parameters are substituted in the corresponding expressions (7) and (8). These plug-in autocorrelations and cross-correlations are plotted in Fig. 4 together with the corresponding sample moments. In concordance with the analytical results on the properties of TGA-SV models, we can observe that the plug-in cross-correlations are nearly the same regardless of whether the errors are Gaussian or GED. Furthermore, all asymmetric specifications of the log-volatility imply very similar cross-correlations. We can observe that these plug-in cross-correlations are smaller than the corresponding sample cross-correlations for the first two and the seventh lags while they explain rather well the sample cross-correlations for all other lags. The sample cross-correlations plotted in Fig. 4 suggest that a specification of the leverage effect with further lags may be appropriate to represent the dynamic dependence of S&P 500 volatilities. With respect to the particular specification of the leverage effect, it seems that the E-SV specification provides the best fit between plug-in and sample moments while introducing the threshold parameter does not have a clear benefit with respect to the moments that the models can explain.

Fig. 5 plots the volatility responses obtained after substituting in Eq. (14) the unknown parameters by their corresponding estimates obtained assuming GED errors. We can observe that the response to positive shocks implied by the three asymmetric specifications is very similar. However, the response of volatility to negative shocks can be rather different depending on the particular specification of the function $f(\varepsilon_t; \theta)$ considered.

We assess the volatility forecast accuracy of the various alternative models considered by forecasting the tail distribution of returns, namely, the Value-at-risk (VaR) and by constructing volatility forecast intervals. With this purpose, we run a particle filter with a large number of particles, namely, $N = 10^6$, and compute the 1%, 5% and 10% one-step-ahead VaRs and construct one-step-ahead forecast intervals for future volatilities with nominal coverages of 80%, 90%, 95% and 99%. Table 4 reports the failure rates for the VaR and the volatility intervals constructed following negative returns while Table 5 reports the same quantities constructed following positive returns. We also provide the mean deviation between the squared returns and forecasted squared returns computed as $T^{-1} \sum_{t=1}^{T} (y_t^2 - \mathbb{E}[y_t^2|y_{1:t-1}])$. Asymptotic standard errors are reported in parenthesis with bold values indicating significant differences from nominal values at the 5% significance level. Note that Andersen and Bollerslev (1998) report that volatility models have poor predictive power when volatility forecasts are evaluated against squared returns, as a consequence of the noise existing in the return generating process. An alternative would be to evaluate the volatility forecasts against the realised volatility obtained from high frequency data. In total 1520 forecasts follow negative returns while 1988 follow positive returns. Consider first the results corresponding to VaR and volatility

Table 4

Failure rates of VaR and of volatility forecast intervals following past negative returns. Average volatility forecast errors. Asymptotic standard errors are in parentheses. Bold values indicate significant differences from theoretical values at the 5% significance level.

Model	VaR			Volatility forecast intervals				Volatility forecast error
	1%	5%	10%	80%	90%	95%	99%	
TGA-SV-G	1.25 (0.29)	4.41 (0.53)	9.28 (0.74)	20.86 (1.04)	9.01 (0.73)	4.61 (0.54)	1.05 (0.26)	0.69 (0.37)
E-SV-G	1.38 (0.30)	4.61 (0.54)	9.28 (0.74)	21.25 (1.05)	9.54 (0.75)	4.80 (0.55)	1.12 (0.27)	**0.76** (0.37)
A-ARSV-G	1.51 (0.31)	4.67 (0.54)	9.01 (0.73)	20.99 (1.04)	9.41 (0.75)	4.80 (0.55)	1.18 (0.28)	**0.72** (0.37)
ARSV-G	**1.78** (0.34)	5.98 (0.61)	10.79 (0.80)	**23.95** (1.09)	**12.04** (0.83)	**6.25** (0.62)	**1.78** (0.34)	**1.32** (0.38)
TGA-SV-N	1.18 (0.28)	4.61 (0.54)	9.41 (0.75)	21.38 (1.05)	9.28 (0.74)	4.34 (0.52)	0.99 (0.25)	0.71 (0.32)
E-SV-N	1.25 (0.29)	4.87 (0.55)	9.54 (0.75)	21.51 (1.05)	9.54 (0.75)	4.54 (0.53)	1.05 (0.26)	**0.78** (0.37)
A-ARSV-N	1.18 (0.28)	4.87 (0.55)	9.47 (0.75)	21.78 (1.06)	9.61 (0.76)	4.47 (0.53)	1.05 (0.26)	**0.82** (0.37)
ARSV-N	**1.71** (0.33)	5.99 (0.61)	10.79 (0.80)	**24.01** (1.10)	**12.11** (0.84)	5.92 (0.61)	**1.64** (0.33)	**1.34** (0.38)
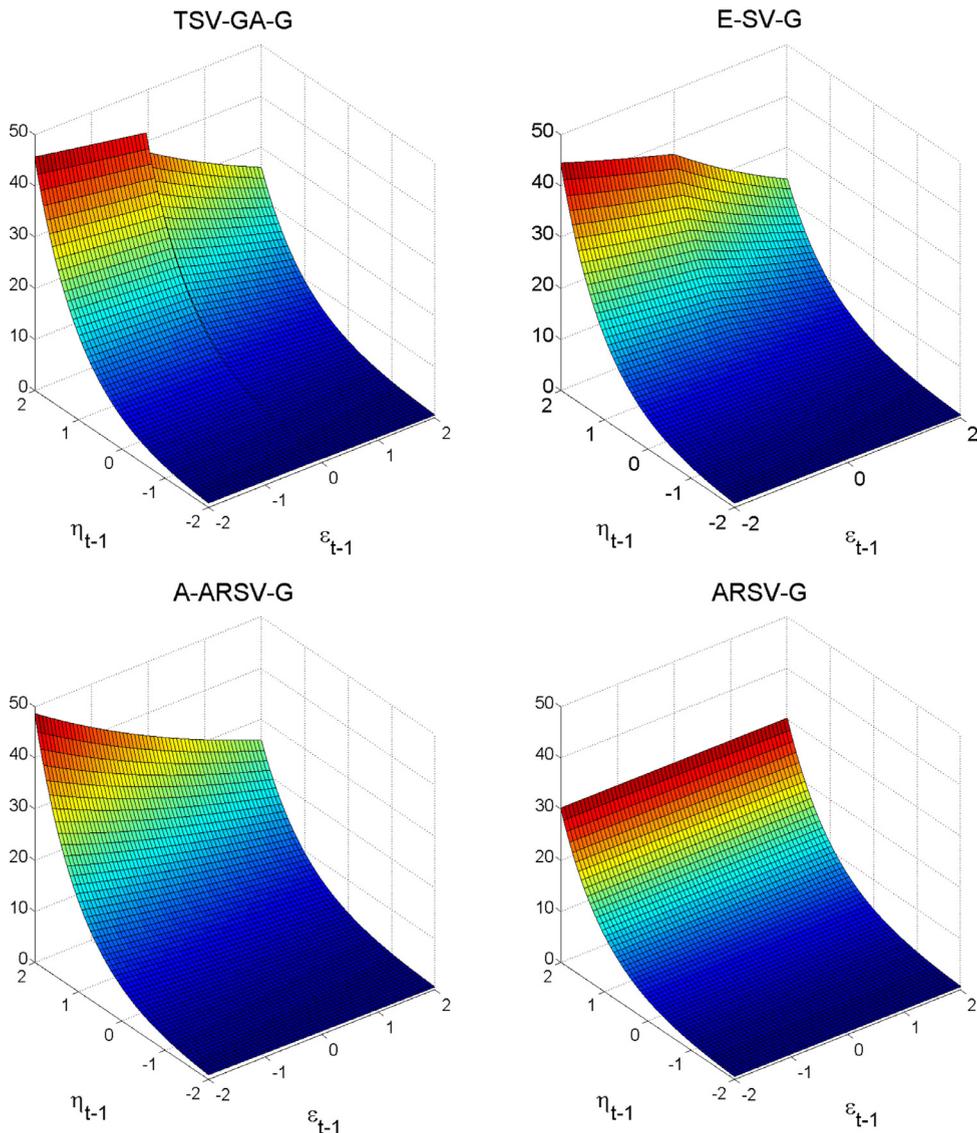
Fig. 5. Plug-in volatility responses of models with GED errors fitted to S&P 500 returns.

Table 5
Failure rates of VaR forecasts, of volatility forecast intervals and mean volatility forecast errors following past positive returns. Asymptotic standard errors are in parentheses. Bold values indicate significant differences from theoretical values at the 5% significance level.

Model	VaR			Forecast volatility intervals				Forecast volatility error
	1%	5%	10%	80%	90%	95%	99%	$T^{-1}\sum_{t=1}^{T}(y_t^2 - \mathbb{E}[y_t^2|y_{1:t-1}])$
TGA-SV-G	1.21 (0.25)	5.13 (0.50)	9.71 (0.66)	19.93 (0.90)	9.16 (0.65)	4.43 (0.46)	1.06 (0.23)	−0.17 (0.19)
E-SV-G	1.11 (0.23)	5.03 (0.49)	9.66 (0.66)	19.73 (0.89)	9.26 (0.65)	4.48 (0.46)	0.91 (0.21)	−0.21 (0.16)
A-ARSV-G	1.16 (0.24)	5.03 (0.49)	9.46 (0.66)	19.48 (1.04)	9.11 (0.75)	4.58 (0.55)	0.91 (0.28)	−0.22 (0.16)
ARSV-G	0.75 (0.19)	4.53 (0.47)	**7.85** (0.60)	**16.00** (0.82)	**7.85** (0.60)	**3.37** (0.40)	0.70 (0.19)	**−0.76** (0.17)
TGA-SV-N	1.06 (0.23)	4.93 (0.49)	9.51 (0.66)	19.63 (0.89)	8.81 (0.64)	4.23 (0.45)	0.96 (0.22)	−0.22 (0.16)
E-SV-N	0.96 (0.29)	4.78 (0.55)	9.51 (0.75)	19.58 (0.89)	**8.66** (0.63)	4.23 (0.45)	0.91 (0.21)	−0.27 (0.16)
A-ARSV-N	0.96 (0.22)	4.88 (0.48)	9.56 (0.66)	19.58 (0.89)	8.76 (0.63)	4.28 (0.45)	0.91 (0.21)	−0.27 (0.16)
ARSV-N	0.70 (0.19)	4.48 (0.46)	**7.95** (0.61)	**16.15** (0.83)	**7.70** (0.60)	**3.32** (0.40)	0.70 (0.19)	**−0.76** (0.17)

Table 6
Estimation results of TGA-SV-G model for daily returns observed from January 4 1950 to December 31 1971 (fist column), January 3 1972 to December 31 1993 (second column) and January 3 1994 to May 8 2017 (third column). Standard errors of SML estimates are in parenthesis. Log-likelihood together with p-values for alternative specifications of the log-volatility.

Parameter	1950–1971	1972–1993	1994–2017
$\mu(1-\phi)$	−0.040 (0.005)	−0.003 (0.001)	−0.000 (0.002)
ϕ	0.956 (0.005)	0.990 (0.006)	0.977 (0.003)
α	0.026 (0.033)	−0.008 (0.021)	0.024 (0.026)
γ_1	**−0.109** (0.017)	−0.049 (0.010)	−0.146 (0.014)
γ_2	−0.134 (0.041)	0.001 (0.025)	−0.038 (0.038)
σ_η^2	0.035 (0.004)	0.008 (0.001)	0.0166 (0.002)
ν	1.911 (0.090)	1.578 (0.058)	1.712 (0.071)
Log likelihood	−5039.75	−7830.71	−7777.76
E-SV-G	**−5040.04** (0.445)	−6830.78 (0.697)	−7778.13 (0.388)
A-ARSV-G	**−5045.99** (0.002)	−6830.79 (0.926)	−7778.67 (0.405)
ARSV-G	**−5128.25** (0.3000)	−6852.17 (0.000)	−7904.30 (0.000)
TGA-SV-N	**−5040.07** (0.422)	−6849.91 (0.000)	−7784.44 (0.000)
E-SV-N	**−5040.36** (0.540)	−6849.91 (0.000)	−7785.00 (0.001)
A-ARSV-N	**−5051.96** (0.000)	−6858.73 (0.000)	−7794.95 (0.000)
ARSV-N	**−5603.09** (0.000)	−7411.97 (0.000)	−8321.53 (0.000)

forecast intervals obtained following negative returns. In this case, we can see that, if the asymmetry is not modeled, the VaR is underestimated with the failures being larger than the nominal. The symmetric models are clearly rejected. Furthermore, the coverage of symmetric volatility intervals following negative returns is below the nominal. Therefore, it seems that after negative returns, not only the volatility based on symmetric models is underestimated but also its variability is too small. When comparing different specifications of the asymmetry, we can observe that none of them are rejected when looking at VaR failures or volatility forecast intervals. However, the mean volatility forecast errors are positive for all specifications but for the general TGA-SV model. In this latter case, regardless of whether the errors are normal or GED, the mean errors are not significantly different from zero. Finally, when looking at one-step-ahead VaR failures and volatility forecast intervals obtained after positive returns, we can observed that, as before, the symmetric specification of the volatility is rejected. In this case, all symmetric specifications are similar with zero mean forecast errors.

5.2. Daily returns

In this subsection, we analyze daily returns observed from January 4 1950 to May 8 2017. In order to investigate the dynamic evolution of the specification of log-volatilities, the sample has been splitted into three subperiods: (i) January 4 1950 to December 31 1971 ($T = 5506$); (ii) January 3 1972 to December 31 1993 ($T = 5560$); (iii) January 3 1994 to May 8 2017 ($T = 5878$). Using data in each of these three subperiods, we estimate the parameters of the general TGA-SV-G model using the particle filter SML estimator. The estimated parameters together with their standard errors are reported in Table 6. We can observe that, in all subperiods, the persistence is rather large with ϕ being estimated over 0.95. The large estimate of ϕ together with the small estimate of σ_η^2 imply a rather smooth evolution of the volatility in the decades of the seventies and eighties. The threshold parameter, α, is not significant. The best specification of the asymmetric response of the volatility to past positive and negative returns within each of the three subperiods considered is rather different. In the second subperiod, decades of the seventies and eighties, the only asymmetric parameter that is significative is γ_1 that is estimated as -0.049. Therefore, the basic asymmetric SV model with correlated errors seems to be preferred. Also note that, in these two decades, normality of standardized returns is clearly rejected. The A-ARSV model with GED errors seems to be an adequate representation for the evolution of volatilities. However, in the first subperiod, decades of the fifties and sixties, normality of standarized returns is not rejected and the asymmetric behavior of the volatility is much stronger with the estimates of both γ_1 and γ_2 being significant. The threshold parameter, α, although not significant, takes a clearly larger value. In the more recent period, we can observe that normality of standardized returns is rejected and that all asymmetric parameters have t-statistics with absolute values larger than 1. Results on the accuracy of the alternative specifications to capture the tail distributions of forecasted returns after positive and negative past returns are available upon request.

6. Conclusions and further research

In this paper we provide a powerful instrument to decide among alternative specifications of symmetric responses of volatility in the context of stochastic volatility models. First, we derive the statistical properties of a general family of asymmetric SV models, denoted as GA-SV, which includes some of the most popular models usually implemented when modeling heteroscedastic series with leverage effect. In particular, the A-ARSV model which incorporates the leverage effect through the correlation between the standardized returns and log-volatility equations, the E-SV model which adds a noise to the log-volatility equation specified as an EGARCH model and a restricted T-SV model, in which the constant of the volatility equation is different depending on whether one-lagged returns are positive or negative, are included in the GA-SV family. Closed-form expressions of the statistical properties of these models are obtained when the errors are GED. We also show that the asymmetric response of volatility to positive and negative return shocks is different depending on the size and sign of the volatility shocks.

Second, we analyze the finite sample properties of the particle filter-based simulated ML estimator and show that it leads to accurate estimates of the parameters of a general specification belonging to the GA-SV family even when the sample size is relatively small.

Third, the results are illustrated by fitting the general specification belonging to the GA-SV family and several restricted models, to estimate and forecast the volatility of S&P 500 weekly returns. We show that symmetric volatility specifications are clearly rejected when looking at log-likelihood ratios and one-step-ahead VaR estimates and volatility forecast intervals. Furthermore, although the alternative specifications of the leverage effect considered generate similar plug-in autocorrelations of squared returns and cross-correlations between returns and future squared returns, the properties of VaR forecasts and volatility forecast intervals are clearly in favor of the general specification. Finally, we also consider daily S&P 500 returns and show that the most appropriate specification of the asymmetric response of volatility can be different in different periods of time.

Appendix A. Proof of Theorems

A.1. Proof of Theorem 2.1

Consider y_t, which, according to Eq. (1), is given by $y_t = \epsilon_t \exp(h_t/2)$. From Eq. (2), h_t can be written as

$$h_t - \mu = \sum_{i=1}^{\infty} \phi^{i-1}(f(\epsilon_{t-i}; \theta) + \eta_{t-i}). \tag{A.1}$$

First, note that if $|\phi| < 1$, then $\Psi(x) = \sum_{i=1}^{\infty} \phi^{i-1} x_i$ is a measurable function for $x = (x_1, x_2, \ldots) \in \mathbb{R}_{\infty}$. Second, given that ϵ_t and η_t are both IID and mutually independent for any lag and lead, then $\{f(\epsilon_t; \theta) + \eta_t\}$ is also an IID sequence. Lemma 3.5.8 of Stout (1974) states that an IID sequence is always strictly stationary. Therefore, in (A.1), if $|\phi| < 1$, h_t is expressed as a measurable function of a strictly stationary process and, consequently, according to Theorem 3.5.8 of Stout (1974), h_t is strictly stationary. As σ_t is a continuous function of h_t, σ_t is also strictly stationary. The level noise ϵ_t is independent of σ_t and strictly stationary by definition. Therefore, it is easy to show that $y_t = \sigma_t \epsilon_t$ is strictly stationary.

When $|\phi| < 1$ and $Var(f(\epsilon_t; \theta)) < \infty$, y_t and σ_t^2 are strictly stationary and, consequently, any existing moments are time invariant. Next we show that σ_t has finite moments of arbitrary positive order c when ϵ_t follows a distribution such that $E(\exp(0.5cf(\epsilon_t; \theta))) < \infty$.

From expression (A.1), the power-transformed volatility can be written as follows

$$\sigma_t^c = \exp(0.5c\mu) \exp\left(0.5c \sum_{i=1}^{\infty} \phi^{i-1}(f(\epsilon_{t-i}; \theta) + \eta_{t-i})\right). \tag{A.2}$$

Given that ϵ_t and η_t are mutually independent for all lags and leads, the following expression is obtained after taking expectations on both sides of Eq. (A.2)

$$E(\sigma_t^c) = \exp(0.5c\mu) E\left[\exp\left(0.5c \sum_{i=1}^{\infty} \phi^{i-1} f(\epsilon_{t-i}; \theta)\right)\right] E\left[\exp\left(0.5c \sum_{i=1}^{\infty} \phi^{i-1} \eta_{t-i}\right)\right]. \tag{A.3}$$

As η_t is Gaussian, the last expectation in (A.3) can be evaluated using the expression of the moments of the Log-Normal. Furthermore, given that η_t and ϵ_t are both IID sequences, it is easy to show that (A.3) becomes

$$E(\sigma_t^c) = \exp(0.5c\mu) \exp\left(\frac{c^2 \sigma_\eta^2}{8(1 - \phi^2)}\right) \prod_{i=1}^{\infty} E\left[\exp\left(0.5c\phi^{i-1} f(\epsilon_{t-i}; \theta)\right)\right]. \tag{A.4}$$

We need to show that $P(0.5c, \phi) \equiv \prod_{i=1}^{\infty} E\left[\exp\left(0.5c\phi^{i-1} f(\epsilon_{t-i}; \theta)\right)\right]$ is finite when $E(\exp(0.5cf(\epsilon_{t-i}; \theta))]) < \infty$. In general, we are going to prove that when $\sum_{i=1}^{\infty} |b_i| < \infty$ and $E(\exp(b_i f(\epsilon_{t-i}; \theta))) < \infty$, then $\prod_{i=1}^{\infty} E[\exp(b_i f(\epsilon_{t-i}; \theta))]$ is always finite.

Define $a_i = E(\exp(b_i f(\epsilon_{t-i}; \theta)))$. As $0 < a_i < \infty$, according to Section 0.25 of Ryzhik et al. (2007), the sufficient and necessary condition for the infinite product $\prod_{i=1}^{\infty} a_i$ to converge to a finite, nonzero number is that the series $\sum_{i=1}^{\infty} (a_i - 1)$ converge. Expanding a_i in Taylor series around $b_i = 0$, we have

$$a_i - 1 = O(b_i) \text{ as } b_i \to 0.$$

Consequently, for some $\varsigma > 0$, there exist a finite M independent of i such that

$$\sup_{|b_i| < \varsigma, b_i \neq 0} |O(b_i)| < M|b_i|.$$

$\sum_{i=1}^{\infty} |b_i| < \infty$ implies $\sum_{i=1}^{\infty} |a_i - 1| < \infty$, therefore $\sum_{i=1}^{\infty} (a_i - 1) < \infty$. Thus $\prod_{i=1}^{\infty} a_i < \infty$.

Here $b_i = 0.5c\phi^{i-1}$. Therefore, if $|\phi| < 1$, then $\sum_{i=1}^{\infty} |b_i| = \frac{0.5c}{1-\phi} < \infty$. Thus, the product $\prod_{i=1}^{\infty} E(\exp(0.5c\phi^{i-1} f(\epsilon_{t-i}; \theta)))$ and, consequently, $E(\sigma_t^c)$ are finite when $E(\exp(0.5c\phi^{i-1} f(\epsilon_{t-i}; \theta))) < \infty$. Note that when $|\phi| < 1$, $E(\exp(0.5cf(\epsilon_t; \theta))) < \infty$ guarantees that $E(\exp(0.5c\phi^{i-1} f(\epsilon_{t-i}; \theta))) < \infty$ for any positive integer i. Therefore, if $|\phi| < 1$ and $E(\exp(0.5cf(\epsilon_t; \theta))) < \infty$, $E(\sigma_t^c)$ is finite.

Finally, consider y_t, which, according to Eq. (1), is given by $y_t = \sigma_t \epsilon_t$. Given that σ_t and ϵ_t are contemporaneously independent, the following expressions are obtained

$$E(|y_t|^c) = E(\sigma_t^c) E(|\epsilon_t|^c), \tag{A.5}$$

$$E(y_t^c) = E(\sigma_t^c) E(\epsilon_t^c). \tag{A.6}$$

Replacing formula (A.4) into (A.5) yields the following required expression

$$E(|y_t|^c) = \exp(0.5c\mu) E(|\epsilon_t|^c) \exp\left(\frac{c^2 \sigma_\eta^2}{8(1 - \phi^2)}\right) P(0.5c, \phi), \tag{A.7}$$

where $P(x, y) \equiv \prod_{i=1}^{\infty} E(\exp(xy^{i-1}f(\epsilon_{t-i}; \theta)))$. Therefore, if further ϵ_t follows a distribution such that $E(\epsilon_t^c) < \infty$, which is equivalent to $E(|\epsilon_t|^c) < \infty$, then $|y_t|$ has finite moments of arbitrary order c. On the other hand, following the same steps, we obtain

$$E(y_t^c) = \exp(0.5c\mu)E(\epsilon_t^c)\exp\left(\frac{c^2\sigma_\eta^2}{8(1-\phi^2)}\right)P(0.5c, \phi). \tag{A.8}$$

Thus, $E(y_t^c) < \infty$ if $|\phi| < 1$, $E(\epsilon_t^c) < \infty$ and $E(\exp(0.5cf(\epsilon_t; \theta))) < \infty$.

A.2. Proof of Theorem 2.2

Consider y_t as given in Eqs. (1) and (2). We first compute the τ-th order auto-covariance of $|y_t|^c$ which is given by

$$E(|\epsilon_t|^c\sigma_t^c|\epsilon_{t-\tau}|^c\sigma_{t-\tau}^c) - [E(|y_t|^c)]^2. \tag{A.9}$$

Note that from Eq. (2), $\sigma_t^c = \exp\{0.5ch_t\}$ can be written as follows

$$\sigma_t^c = \exp\{0.5c\mu(1-\phi^\tau)\}\exp\left\{0.5c\sum_{i=1}^{\tau}\phi^{i-1}(f(\epsilon_{t-i}; \theta) + \eta_{t-i})\right\}\sigma_{t-\tau}^{c\phi^\tau}. \tag{A.10}$$

The following expression of the auto-covariance is obtained after substituting (A.7) and (A.10) into (A.9)

$$cov(|y_t|^c, |y_{t-\tau}|^c) = E\left(|\epsilon_t|^c|\epsilon_{t-\tau}|^c\exp(0.5c\mu(1-\phi^\tau))\exp\left(\sum_{i=1}^{\tau}0.5c\phi^{i-1}(f(\epsilon_{t-i}; \theta)+\eta_{t-i})\right)\sigma_{t-\tau}^{c(\phi^\tau+1)}\right)$$
$$-\left\{\exp(0.5c\mu)E(|\epsilon_t|^c)\exp\left(\frac{c^2\sigma_\eta^2}{8(1-\phi^2)}\right)P(0.5c, \phi))\right\}^2. \tag{A.11}$$

Given that ϵ_t and η_t are IID sequences mutually independent for any lag and lead and that $\sigma_{t-\tau}$ only depends on lagged disturbances, substituting the time-invariant moment of σ_t in (A.4), Eq. (A.11) can be written as follows

$$cov(|y_t|^c, |y_{t-\tau}|^c) = \exp(0.5c(1-\phi^\tau)\mu)E(|\epsilon_t|^c)E\left(\sigma_t^{c(1+\phi^\tau)}\right)E\left(|\epsilon_{t-\tau}|^c\exp\left(0.5c\phi^{\tau-1}f(\epsilon_{t-\tau})\right)\right)$$
$$\times E\left(\exp\left(\sum_{i=1}^{\tau-1}0.5c\phi^{i-1}f(\epsilon_{t-i})\right)\right)E\left(\exp\left(\sum_{i=1}^{\tau}0.5c\phi^{i-1}\eta_{t-i}\right)\right) - [E(|\epsilon_t|^c)E(\sigma_t^c)]^2. \tag{A.12}$$

Replacing $E(\sigma_t^c)$ as given in (A.4) into Eq. (A.12), we obtain the following expression of the autocorrelations

$$cov(|y_t|^c, |y_{t-\tau}|^c) = \exp(c\mu)E(|\epsilon_t|^c)\exp\left(\frac{1+\phi^\tau}{4(1-\phi^2)}c^2\sigma_\eta^2\right)E\left(|\epsilon_t|^c\exp\left(0.5c\phi^{\tau-1}f(\epsilon_t; \theta)\right)\right)$$
$$\times \prod_{i=1}^{\tau-1}E\left(\exp\left(0.5c\phi^{i-1}f(\epsilon_{t-i}; \theta)\right)\right)\prod_{i=1}^{\infty}E\left(\exp\left(0.5c(1+\phi^\tau)\phi^{i-1}f(\epsilon_{t-i}; \theta)\right)\right)$$
$$- \exp(c\mu)(E(|\epsilon_t|^c))^2\exp\left(\frac{c^2\sigma_\eta^2}{4(1-\phi^2)}\right)[P(0.5c, \phi)]^2.$$

The required expression of $\rho_c(\tau)$ follows directly from $\rho_c(\tau) = \frac{cov(|y_t|^c, |y_{t-\tau}|^c)}{E(|y_t|^{2c}) - [E(|y_t|^c)]^2}$, where the denominator can be obtained from (A.7).

A.3. Proof of Theorem 2.3

The calculation of the cross-covariance between $|y_t|^c$ and $y_{t-\tau}$ is obtained following the same steps as in Appendix A.2. That is

$$cov(|y_t|^c, y_{t-\tau}) = \exp(0.5(c+1)\mu)E(|\epsilon_t|^c)\exp\left(\frac{1+c^2+2c\phi^\tau}{8(1-\phi^2)}\sigma_\eta^2\right)E\left(\epsilon_t\exp\left(0.5c\phi^{\tau-1}f(\epsilon_t; \theta)\right)\right)$$
$$\times \prod_{i=1}^{\infty}E\left(\exp\left(0.5(1+c\phi^\tau)\phi^{i-1}f(\epsilon_{t-i}; \theta)\right)\right)\prod_{i=1}^{\tau-1}E\left(\exp\left(0.5c\phi^{i-1}f(\epsilon_{t-i}; \theta)\right)\right). \tag{A.13}$$

Finally, $\rho_{c1}(\tau) = \frac{cov(|y_t|^c, y_{t-\tau})}{\sqrt{E(|y_t|^{2c}) - E^2(|y_t|^c)}\sqrt{E(y_t^2)}}$ together with (A.7) and (A.13) yields the required Eq. (8).

Appendix B. Expectations

B.1. Expectations for computing $E(|y_t|^c)$, $corr(|y_t|^c, |y_{t+\tau}|^c)$ *and* $corr(y_t, |y_{t+\tau}|^c)$ *when* $\epsilon \sim GED(\nu)$ *and* $f(\epsilon; \theta) = \alpha[I(\epsilon < 0) - 0.5] + \gamma_1\epsilon + \gamma_2(|\epsilon| - E|\epsilon|)$

If ϵ has a centered and standardized GED distribution, with parameter $0 < \nu \leq \infty$, then, the density function of ϵ is given by $\psi(\epsilon) = C_0 \exp(-\frac{|\epsilon|^\nu}{2\lambda^\nu})$, where $C_0 \equiv \frac{\nu}{\lambda 2^{1+1/\nu}\Gamma(1/\nu)}$ and $\lambda \equiv (2^{-2/\nu}\Gamma(1/\nu)/\Gamma(3/\nu))^{1/2}$, with $\Gamma(\cdot)$ being the Gamma function. Thus, given that the distribution of ϵ is symmetric with support $(-\infty, \infty)$, if p is a nonnegative finite integer, then

$$E(|\epsilon|^p) = C_0 \int_{-\infty}^{+\infty} |\epsilon|^p \exp\left(-\frac{|\epsilon|^\nu}{2\lambda^\nu}\right) d\epsilon = 2C_0 \int_0^{+\infty} \epsilon^p \exp\left(-\frac{\epsilon^\nu}{2\lambda^\nu}\right) d\epsilon.$$

Substituting $s = \frac{\epsilon^\nu}{2\lambda^\nu}$ and solving the integral yields

$$E(|\epsilon|^p) = 2^{\frac{p}{\nu}} \lambda^p \Gamma((p+1)/\nu)/\Gamma(1/\nu). \tag{B.1}$$

On the other hand,

$$E(|\epsilon|^p \exp(bf(\epsilon; \theta))) = \int_{-\infty}^{+\infty} |\epsilon|^p \exp(b\alpha[I(\epsilon < 0) - 0.5] + b\gamma_1\epsilon + b\gamma_2(|\epsilon| - E|\epsilon|))C_0 \exp\left(-\frac{|\epsilon|^\nu}{2\lambda^\nu}\right) d\epsilon$$

$$= C_0 \exp(-0.5b\alpha - b\gamma_2 E|\epsilon|)\left[\int_{-\infty}^0 (-\epsilon)^p \exp(b\alpha) \exp(b(\gamma_1 - \gamma_2)\epsilon) \exp\left(-\frac{(-\epsilon)^\nu}{2\lambda^\nu}\right) d\epsilon\right.$$

$$\left. + \int_0^{+\infty} \epsilon^p \exp(b(\gamma_1 + \gamma_2)\epsilon) \exp\left(-\frac{\epsilon^\nu}{2\lambda^\nu}\right) d\epsilon\right].$$

Integrating by substitution with $s = -\epsilon$ in the first integral, we obtain

$$E(|\epsilon|^p \exp(bf(\epsilon; \theta))) = C_0 \exp(-0.5b\alpha - b\gamma_2 E|\epsilon|)\left[\int_0^{+\infty} s^p \exp(b\alpha) \exp(b(\gamma_2 - \gamma_1)s) \exp\left(-\frac{s^\nu}{2\lambda^\nu}\right) ds\right.$$

$$\left. + \int_0^{+\infty} \epsilon^p \exp(b(\gamma_1 + \gamma_2)\epsilon) \exp\left(-\frac{\epsilon^\nu}{2\lambda^\nu}\right) d\epsilon\right] \tag{B.2}$$

$$= C_0 \exp(-0.5b\alpha - b\gamma_2 E|\epsilon|) \int_0^{+\infty} \epsilon^p \exp\left(-\frac{\epsilon^\nu}{2\lambda^\nu}\right)[\exp(b\alpha) \exp(b(\gamma_2 - \gamma_1)\epsilon)$$

$$+ \exp(b(\gamma_1 + \gamma_2)\epsilon)] d\epsilon.$$

We can rewrite the previous equation by replacing ϵ with $\lambda(2y)^{1/\nu}$ as follows

$$E(|\epsilon|^p \exp(bf(\epsilon; \theta))) = C_0 \exp(-0.5b\alpha - b\gamma_2 E|\epsilon|)\frac{\lambda^{p+1}2^{\frac{1+p}{\nu}}}{\nu}$$

$$\times \int_0^{+\infty} y^{-1+\frac{1+p}{\nu}} \exp(-y)\left[\exp(b\alpha) \exp(b(\gamma_2 - \gamma_1)\lambda 2^{\frac{1}{\nu}} y^{\frac{1}{\nu}}) + \exp(b(\gamma_1 + \gamma_2)\lambda 2^{\frac{1}{\nu}} y^{\frac{1}{\nu}})\right] dy.$$

Expanding the expression within the square brackets in a Taylor series and substituting C_0, the following expression is obtained

$$E(|\epsilon|^p \exp(bf(\epsilon; \theta))) = \exp\left(-0.5b\alpha - b\gamma_2 2^{\frac{1}{\nu}} \lambda \Gamma(2/\nu)/\Gamma(1/\nu)\right)\frac{\lambda^p 2^{\frac{p}{\nu}-1}}{\Gamma\left(\frac{1}{\nu}\right)}$$

$$\times \int_0^{+\infty} \sum_{k=0}^{+\infty}\left[\exp(b\alpha)\left(b\lambda 2^{\frac{1}{\nu}}(\gamma_2 - \gamma_1)\right)^k + \left(b\lambda 2^{\frac{1}{\nu}}(\gamma_1 + \gamma_2)\right)^k\right]\frac{y^{-1+\frac{1+p+k}{\nu}} \exp(-y)}{k!} dy. \tag{B.3}$$

$$E(|\epsilon|^p \exp(bf(\epsilon; \theta))) = \exp\left[-0.5b\alpha - b\gamma_2 2^{1/\nu} \lambda \Gamma(2/\nu)/\Gamma(1/\nu)\right]2^{p/\nu}\lambda^p \sum_{k=0}^{\infty}(2^{1/\nu}\lambda b)^k$$

$$\times \left[(\gamma_1 + \gamma_2)^k + \exp(b\alpha)(\gamma_2 - \gamma_1)^k\right]\frac{\Gamma((p+k+1)/\nu)}{2\Gamma(1/\nu)k!} < \infty. \tag{B.4}$$

Following the same steps, the following expression is obtained when $\nu > 1$,

$$
\begin{aligned}
E(\epsilon^p \exp(bf(\epsilon;\theta))) =\ & \exp\left[-0.5b\alpha - b\gamma_2 2^{1/\nu}\lambda\Gamma(2/\nu)/\Gamma(1/\nu)\right]2^{p/\nu}\lambda^p \\
& \times \sum_{k=0}^{\infty}(2^{1/\nu}\lambda b)^k\left[(\gamma_1+\gamma_2)^k + (-1)^p\exp(b\alpha)(\gamma_2-\gamma_1)^k\right]\frac{\Gamma((p+k+1)/\nu)}{2\Gamma(1/\nu)k!} < \infty.
\end{aligned} \tag{B.5}
$$

Note that the expectations (B.4) and (B.5) are only valid when $\nu > 1$. When $0 < \nu \le 1$, it is not possible to obtain closed-form expression of the required expectations. In this case, we can only obtain the conditions for the expectations to be finite. When $0 < \nu < 1$, it is very easy to verify that $E(|\epsilon|^p\exp(bf(\epsilon;\theta))) < \infty$ if and only if the both integrals in (B.2) are finite, which holds if and only if $b(\gamma_2 - \gamma_1) \le 0$ and $b(\gamma_2 + \gamma_1) \le 0$. When $\nu = 1$, similarly, the sufficient and necessary conditions for the infinity of $E(|\epsilon|^p\exp(bf(\epsilon;\theta)))$ are $b(\gamma_2 - \gamma_1) < \frac{1}{2\lambda}$ and $b(\gamma_2 + \gamma_1) < \frac{1}{2\lambda}$. That is $b(\gamma_2 - \gamma_1) < \sqrt{2}$ and $b(\gamma_2 + \gamma_1) < \sqrt{2}$. The conditions for the infinity of $E(\epsilon^p\exp(bf(\epsilon;\theta)))$ are the same as those for $E(|\epsilon|^p\exp(bf(\epsilon;\theta)))$ $0 < \nu \le 1$.

Finally, when $\epsilon \sim$ Student-t with d degrees of freedom $(d > 2)$ and is normalized to satisfy $E(\epsilon) = 0$, $var(\epsilon) = 1$, then

$$
\begin{aligned}
E(|\epsilon|^p \exp(bf(\epsilon;\theta))) =\ & C_1\exp(-0.5b\alpha)\left[\int_0^{+\infty}\epsilon^p\exp(b\alpha)\exp(b(\gamma_2-\gamma_1)\epsilon)\left(1+\frac{\epsilon^2}{d-2}\right)^{-\frac{d+1}{2}}d\epsilon \right.\\
& \left. +\int_0^{+\infty}\epsilon^p\exp(b(\gamma_1+\gamma_2)\epsilon)\left(1+\frac{\epsilon^2}{d-2}\right)^{-\frac{d+1}{2}}d\epsilon\right],
\end{aligned} \tag{B.6}
$$

where $C_1 = \frac{\Gamma\left(\frac{d+1}{2}\right)}{\sqrt{(d-2)\pi}\Gamma\left(\frac{d}{2}\right)}\exp\left[-b\gamma_2 2^{1/\nu}\lambda\Gamma(2/\nu)/\Gamma(1/\nu)\right]$. We can verify that $E(|\epsilon|^p\exp(bf(\epsilon;\theta))) = \infty$ unless $b(\gamma_2 - \gamma_1) \le 0$ and $b(\gamma_2 + \gamma_1) \le 0$.

B.2. Expectations for computing $E(|y_t|^c)$, $corr(|y_t|^c, |y_{t+\tau}|^c)$ and $corr(y_t, |y_{t+\tau}|^c)$ when $\epsilon \sim N(0, 1)$

Assume that all the parameters are defined as in Eqs. (1) and (2). When $\epsilon \sim N(0, 1)$, using the expression (B.2) and the formula 3.462–1 of Ryzhik et al. (2007), the following expressions for any positive integer p and any integer b are derived

$$
\begin{aligned}
E(|\epsilon|^p\exp(bf(\epsilon;\theta))) =\ & \frac{1}{\sqrt{2\pi}}\exp\left(-0.5b\alpha - b\gamma_2\sqrt{\frac{2}{\pi}}\right)\left\{\exp(b\alpha)\Gamma(p+1)\exp\left(\frac{b^2(\gamma_1-\gamma_2)^2}{4}\right)D_{-p-1}(b(\gamma_1-\gamma_2))\right.\\
& \left. +\Gamma(p+1)\exp\left(\frac{b^2(\gamma_1+\gamma_2)^2}{4}\right)D_{-p-1}(-b(\gamma_1+\gamma_2))\right\}
\end{aligned} \tag{B.7}
$$

and

$$
\begin{aligned}
E(\epsilon^p\exp(bf(\epsilon;\theta))) =\ & \frac{1}{\sqrt{2\pi}}\exp\left(-0.5b\alpha - b\gamma_2\sqrt{\frac{2}{\pi}}\right)\left\{(-1)^p\exp(b\alpha)\Gamma(p+1)\exp\left(\frac{b^2(\gamma_1-\gamma_2)^2}{4}\right)\right.\\
& \left. \times D_{-p-1}(b(\gamma_1-\gamma_2)) + \Gamma(p+1)\exp\left(\frac{b^2(\gamma_1+\gamma_2)^2}{4}\right)D_{-p-1}(-b(\gamma_1+\gamma_2))\right\},
\end{aligned} \tag{B.8}
$$

where $D_{-a}(\cdot)$ is the parabolic cylinder function. Particularly, when $p = 0, 1$ or 2, the expressions are reduced to

$$
E(\exp(bf(\epsilon;\theta))) = \exp\left(-0.5b\alpha - b\gamma_2\sqrt{\frac{2}{\pi}}\right)\left(\exp(b\alpha)\exp(\bar{A})\Phi(\bar{C}) + \exp(\bar{B})\Phi(\bar{D})\right),
$$

$$
\begin{aligned}
E(\epsilon\exp(bf(\epsilon;\theta))) =\ & \frac{1}{\sqrt{2\pi}}\exp\left(-0.5b\alpha - b\gamma_2\sqrt{\frac{2}{\pi}}\right)\left\{-\exp(b\alpha)\left[1 + \sqrt{2\pi}\bar{C}\exp(\bar{A})\Phi(\bar{C})\right]\right.\\
& \left. +\left[1 + \sqrt{2\pi}\bar{D}\exp(\bar{B})\Phi(\bar{D})\right]\right\},
\end{aligned}
$$

$$
\begin{aligned}
E(|\epsilon|\exp(bf(\epsilon;\theta))) =\ & \frac{1}{\sqrt{2\pi}}\exp\left(-0.5b\alpha - b\gamma_2\sqrt{\frac{2}{\pi}}\right)\left\{\exp(b\alpha)\left[1 + \sqrt{2\pi}\bar{C}\exp(\bar{A})\Phi(\bar{C})\right]\right.\\
& \left. +\left[1 + \sqrt{2\pi}\bar{D}\exp(\bar{B})\Phi(\bar{D})\right]\right\}
\end{aligned}
$$

and

$$E(|\epsilon|^2 \exp(bf(\epsilon;\theta))) = \frac{1}{\sqrt{2\pi}} \exp\left(-0.5b\alpha - b\gamma_2\sqrt{\frac{2}{\pi}}\right)$$
$$\times \left\{ \exp(b\alpha)\left[\bar{C} + \sqrt{2\pi}(\bar{C}^2 + 1)\exp(\bar{A})\Phi(\bar{C})\right] + \left[\bar{D} + \sqrt{2\pi}(\bar{D}^2 + 1)\exp(\bar{B})\Phi(\bar{D})\right] \right\},$$

where $\Phi(\cdot)$ is the Normal distribution function, $\bar{A} = \frac{b^2(\gamma_1-\gamma_2)^2}{2}$, $\bar{B} = \frac{b^2(\gamma_1+\gamma_2)^2}{2}$, $\bar{C} = -b(\gamma_1-\gamma_2)$ and $\bar{D} = b(\gamma_1+\gamma_2)$.

References

Andersen, T.G., 1996. Return volatility and trading volume: an information flow interpretation of stochastic volatility. J. Financ. 51 (1), 169–204.

Andersen, T.G., Bollerslev, T., 1998. Answering the skeptics: yes, standard volatility models do provide accurate forecasts. Int. Econ. Rev. 39 (4), 885–905.

Andersen, T.G., Bollerslev, T., Diebold, F., Labys, P., 2001a. The distribution of realized exchange rate volatility. J. Am. Stat. Assoc. 96 (453), 42–55.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Ebens, H., 2001b. The distribution of realized stock return volatility. J. Financ. Econ. 61 (1), 43–76.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71 (2), 579–625.

Andrieu, C., Doucet, A., Holenstein, R., 2010. Particle Markov chain Monte Carlo methods. J. R. Stat. Soc.: Ser. B 72 (3), 269–342.

Asai, M., 2008. Autoregressive stochastic volatility models with heavy-tailed distributions: a comparison with multifactor volatility models. J. Empir. Financ. 15 (2), 332–341.

Asai, M., McAleer, M., 2005. Dynamic asymmetric leverage in stochastic volatility models. Econom. Rev. 24 (3), 317–332.

Asai, M., McAleer, M., 2006. Asymmetric multivariate stochastic volatility. Econom. Rev. 25 (2–3), 453–473.

Asai, M., McAleer, M., 2011. Alternative asymmetric stochastic volatility models. Econom. Rev. 30, 548–564.

Asai, M., McAleer, M., Medeiros, M.C., 2012. Asymmetry and long memory in volatility modeling. J. Financ. Econom. 10 (3), 495–512.

Breidt, F.J., 1996. A threshold autoregressive stochastic volatility model. In: Proceedings of the VI Latin American Congress of Probability and Mathematical Statistics (CLAPEM), Valparaiso, Chile. Citeseer.

Broto, C., Ruiz, E., 2004. Estimation methods for stochastic volatility models: a survey. J. Econ. Surv. 18 (5), 613–649.

Cappuccio, N., Lubian, D., Raggi, D., 2004. MCMC Bayesian estimation of a skew-GED stochastic volatility model. Stud. Nonlinear Dyn. Econom. 8 (2).

Cappuccio, N., Lubian, D., Raggi, D., 2006. Investigating in US stock market indexes: evidence from a stochastic volatility model. Appl. Financ. Econ. 16, 479–490.

Carnero, M.A., Peña, D., Ruiz, E., 2004. Persistence and kurtosis in GARCH and stochastic volatility models. J. Financ. Econom. 2 (2), 319–342.

Choy, S.T.B., Walker, S.G., 2003. The extended exponential power distribution and Bayesian robustness. Stat. Probab. Lett. 65, 227–232.

Del Moral, P., 2004. FeynmanKac Formulae: Genealogical and Interacting Particle Systems with Applications. Springer Verlag.

Demos, A., 2002. Moments and dynamic structure of a time-varying parameter stochastic volatility in mean model. Econom. J. 5 (2), 345–357.

Feunou, B., Tédongap, R., 2012. A stochastic volatility model with conditional skewness. J. Bus. Econ. Stat. 30 (4), 576–591.

Gordon, N., Salmon, D., Smith, A.F., 1993. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. IEE Proc. F 140, 107–113.

Gouriéroux, C., Monfort, A., 1996. Simulation Based Econometric Methods. Oxford University Press, Oxford.

Harvey, A., Ruiz, E., Shephard, N., 1994. Multivariate stochastic variance models. Rev. Econ. Stud. 61 (2), 247–264.

Harvey, A.C., 1990. The Econometric Analysis of Time Series, Second ed. MIT Press, Cambridge.

Harvey, A.C., Shephard, N., 1996. Estimation of an asymmetric stochastic volatility model for asset returns. J. Bus. Econ. Stat. 14 (4), 429–434.

Horváth, L., Kokoszka, P., Zitikis, R., 2006. Sample and implied volatility in GARCH models. J. Financ. Econom. 4 (4), 617–635.

Huang, S.J., Yu, J., 2008. An efficient method for maximum likelihood estimation of a stochastic volatility model. Stat. Interface 1, 289–296.

Jensen, M.J., Maheu, J.M., 2014. Estimating a semiparametric asymmetric stochastic volatility model with a Dirichlet process mixture. J. Econom. 178, 523–538.

Kastner, G., Fruhwirth-Schnatter, S., 2014. Ancillarity–sufficiency interweaving strategy (ASIS) for boosting MCMC estimation of stochastic volatility models. Comput. Stat. Data Anal. 76, 408–423.

Kim, S., Shephard, N., Chib, S., 1998. Stochastic volatility: likelihood inference and comparison with ARCH models. Rev. Econ. Stud. 65 (3), 361–393.

Koopman, S.J., Lit, R., Nguyen, T., 2019. Modified efficient importance sampling for partially non-Gaussian state space models. Stat. Neerl. 73, 44–62.

Koopman, S.J., Shephard, N., Creal, D., 2009. Testing the assumptions behind importance sampling. J. Econom. 149, 2–11.

Liesenfeld, R., Jung, R.C., 2000. Stochastic volatility models: conditional normality versus heavy-tailed distributions. J. Appl. Econom. 15 (2), 137–160.

Liesenfeld, R., Richard, J.F., 2003. Univariate and multivariate stochastic volatility models: estimation and diagnostics. J. Empir. Financ. 10 (4), 505–531.

Malik, S., Pitt, M.K., 2011. Particle filters for continuous likelihood evaluation and maximisation. J. Econom. 165, 190–209.

Malmquist, S., 1950. On a property of order statistics from a rectangular distribution. Skand. Aktuarietidskr. 33, 214–222.

Mao, X., Ruiz, E., Veiga, H., 2017. Threshold stochastic volatility: properties and forecasting. Int. J. Forecast. 33 (4), 1105–1123.

McAleer, M., 2005. Automated inference and learning in modeling financial volatility. Econom. Theory 21 (1), 232–261.

McCausland, W., Miller, S., Pelletire, D., 2011. Simulation smoothing for state-space models: a computational efficiency analysis. Comput. Stat. Data Anal. 55 (1), 199–212.

Meyer, R., Yu, J., 2000. BUGS for a Bayesian analysis of stochastic volatility models. Econom. J. 3 (2), 198–215.

Nakajima, J., Omori, Y., 2012. Stochastic volatility model with leverage and asymmetrically heavy-tailed error using GH skew Student's t-distribution. Comput. Stat. Data Anal. 56 (11), 3690–3704.

Nelson, D., 1991. Conditional heteroscedasticity in asset pricing: a new approach. Econometrica 59, 347–370.

Pérez, A., Ruiz, E., Veiga, H., 2009. A note on the properties of power-transformed returns in long-memory stochastic volatility models with leverage effect. Comput. Stat. Data Anal. 53 (10), 3593–3600.

Pitt, M.K., Malik, S., Doucet, A., 2014. Simulated likelihood inference for stochastic volatility models using particle filtering. Ann. Inst. Stat. Math. 66 (3), 527–552.

Plummer, M., 2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In: Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), pp. 1–10.

van Ravenzwaaij, D., Cassey, P., Brown, S.D., 2018. A simple introduction to Markov chain Monte Carlo sampling. Psychon. Bull. Rev. 25 (1), 143–154.

Richard, J.-F., Zang, W., 2007. Efficient high-dimensional importance sampling. J. Econom. 141 (2), 1385–1411.

Ruiz, E., Veiga, H., 2008. Modelling long-memory volatilities with leverage effect: A-LMSV versus FIEGARCH. Comput. Stat. Data Anal. 52 (6), 2846–2862.

Ryzhik, I.M., Jeffrey, A., Zwillinger, D., 2007. Table of Integrals, Series and Products, Seventh ed. Academic Press, London.

Scharth, M., Kohn, R., 2016. Particle efficient importance sampling. J. Econom. 190, 133–147.

Shestopaloff, A., Neal, R., 2014. Efficient Bayesian Inference for Stochastic Volatility Models with Ensamble MCMC Methods. Technical Report, URL http://arxiv.org/abs/1412.3013.

So, M.K.P., Li, W.K., Lam, K., 2002. A threshold stochastic volatility model. J. Forecast. 21 (7), 473–500.

Stout, W.F., 1974. Almost Sure Convergence. Academic Press, London.

Takahashi, M., Omori, Y., Watanabe, T., 2013. News impact curve for stochastic volatility models. Econ. Lett. 120 (1), 130–134.

Taylor, S.J., 1982. Financial returns modelled by the product of two stochastic processes-a study of the daily sugar prices 1961-75. In: Anderson, O.D. (Ed.), Time Series Analysis: Theory and Practice 1 Amsterdam. North-Holland. chapter 1, pp. 203–226.

Taylor, S.J., 1986. Modelling Financial Time Series. Wiley, Chichester.

Taylor, S.J., 1994. Modelling stochastic volatility: a review and comparative study. Math. Financ. 4, 183–204.

Taylor, S.J., 2007. Asset Price Dynamics, Volatility, and Prediction. Princeton University Press, Princeton, NJ.

Tsiotas, G., 2012. On generalised asymmetric stochastic volatility models. Comput. Stat. Data Anal. 56 (1), 151–172.

Wang, J.J.J., Chan, J.S.K., Choy, S.T.B., 2011. Stochastic volatility models with leverage and heavy-tailed distributions: a Bayesian approach using scale mixtures. Comput. Stat. Data Anal. 55 (1), 852–862.

Wang, J.J.J., Chan, J.S.K., Choy, S.T.B., 2013. Modelling stochastic volatility using generalized t distribution. J. Stat. Comput. Simul. 83 (2), 340–354.

Yu, J., 2005. On leverage in a stochastic volatility model. J. Econom. 127 (2), 165–178.

Yu, J., 2012a. A semiparametric stochastic volatility model. J. Econom. 167 (2), 473–482.

Yu, J., 2012b. Simulation-Based Estimation Methods for Financial Time Series Models. Springer, Berlin, Heidelberg, pp. 401–435. Springer Handbooks of Computational Statistics

Yu, J., Yang, Z., Zhang, X., 2006. A class of nonlinear stochastic volatility models and its implications for pricing currency options. Comput. Stat. Data Anal. 51 (4), 2218–2231.