Statistics and Probability Letters 165 (2020) 108817

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

GARCH quasi-likelihood ratios for SV model and the diffusion limit

Xinyu Song [a], Yazhen Wang [b,*]

[a] School of Statistics and Management, Shanghai University of Finance and Economics, China
[b] Department of Statistics, University of Wisconsin-Madison, United States of America

ARTICLE INFO

Article history:
Received 10 February 2020
Accepted 11 May 2020
Available online 12 June 2020

MSC:
62M07
62M86

Keywords:
Approximating process
Financial modeling
Likelihood process
Stochastic differential equation
Stochastic volatility

ABSTRACT

There is a widely known intriguing phenomenon that discrete-time GARCH and stochastic volatility (SV) models share the same continuous-time diffusion model as their weak convergence limit, but statistically, the GARCH model is not asymptotically equivalent to the SV or diffusion model. This paper investigates GARCH-type quasi-likelihood ratios for the SV and diffusion models whose own likelihoods are analytically intractable. We show that the two quasi-likelihood ratios for the SV and diffusion models asymptotically have the same closed-form expression that is different from the limiting likelihood ratio of the GARCH model.

1. Introduction

Asset volatilities play a central role in finance, and it is important to develop sound statistical inferences for volatility models in financial econometrics. Empirical and theoretical financial studies often employ discrete-time models such as the GARCH and stochastic volatility (SV) models, as well as continuous-time models like the diffusions. It is well known that the GARCH and SV models may have the same diffusion model as their weak-convergence limit. However, statistically, they are not asymptotically equivalent. The statistical non-equivalence is established by showing that their respective likelihoods present different asymptotic behaviors. See Brown et al. (2003), Nelson (1990) and Wang (2002) for details. We further note that the likelihoods for the SV and diffusion models are intractable even asymptotically. Thus, in this paper, we develop a quasi-likelihood approach for the SV and diffusion models and examine the corresponding quasi-likelihood ratios given statistical testing problems. In specific, we plug observations from the SV model and discretely sampled observations from the diffusion model into the GARCH likelihood function to obtain their respective quasi-likelihood ratios. We show that their quasi-likelihood ratios share the same weak limit with a closed-form expression that is different from the limit of the GARCH likelihood ratio.

The rest of this paper is organized as follows. Section 2 reviews the GARCH and SV models and shows that their approximating processes share the same diffusion limit. Section 3 develops quasi-likelihoods for the SV and diffusion models and derives the same weak convergence limit for their quasi-likelihood ratios. Additional technical proofs are collected in an Appendix as the online Supplementary Material.

* Correspondence to: 1175 Medical Science Center, 1300 University Avenue, Madison, WI 53706, United States of America.
E-mail address: yzwang@stat.wisc.edu (Y. Wang).

https://doi.org/10.1016/j.spl.2020.108817

2. Volatility models and their diffusion limits

2.1. GARCH models

The well-known GARCH models (Bollerslev, 1986; Engle, 1982) are defined by setting the conditional variances of a series of prediction errors equal to some functions of lagged errors. Given observed time series x_i, $i = 1, \ldots, n$, we consider the following multiplicative GARCH (p, q) model defined by

$$x_i = \mu_i + z_i, \qquad z_i = \tau_i \varepsilon_i, \tag{1}$$

$$\log \tau_i^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j \log \tau_{i-j}^2 + \sum_{j=1}^{q} \alpha_{p+j} \log \varepsilon_{i-j}^2, \tag{2}$$

where ε_i is a sequence of i.i.d. standard normal random errors, τ_i^2 is the conditional variance of x_i given all information up to time $i-1$, and α_j's are model parameters. In particular, the GARCH(1,1) specification has been found to be sufficient for modeling the dynamics of most financial time series. Thus, we confine our analysis to the GARCH(1,1) model in this paper and employ a common financial parameterization of the drift term μ_i such that

$$\mu_i = c_0 + c_1 \tau_i^2. \tag{3}$$

By taking advantage of the conditional structure in the GARCH(1,1) model, we easily derive its likelihood function in the following

$$\prod_{i=1}^{n} \left[\frac{1}{\sqrt{2\pi}\,\tau_i} \exp\left(-\frac{(x_i - \mu_i)^2}{2\tau_i^2} \right) \right] \propto \left[\prod_{i=1}^{n} \tau_i^{-1} \right] \exp\left(-\frac{\sum_{i=1}^{n} \varepsilon_i^2}{2} \right). \tag{4}$$

Since the above likelihood function is in a relatively simple form, statistical inferences for the GARCH model can be carried out easily.

2.2. Stochastic volatility models

In contrast to the GARCH models defined in (2) where the conditional variance is a deterministic function of past data and model parameters, the stochastic volatility (SV) models generate the conditional variance, known as volatility, by a probability mechanism. As a result, the volatilities are unobservable random variables and the density function is then a mixture over the volatility distribution that is rather complicated. In specific, given the observed time series x_i, $i = 1, \ldots, n$, we consider the following widely used SV models that assume the conditional variance of each x_i obeys a log-AR(p) process such that

$$x_i = \mu_i + y_i, \qquad y_i = \varsigma_i \varepsilon_i, \tag{5}$$

$$\log \varsigma_i^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j \log \varsigma_{i-j}^2 + \alpha_{p+1}\, \delta_i, \tag{6}$$

where ε_i's and δ_i's are independent standard normal random variables, and ς_i's are the conditional variances. As in the GARCH model case, this paper focuses on SV models with a log-AR(1) specification.

Due to the latent random component in the volatility process, the likelihood function for the SV model involves a n-dimensional integration with respect to the unobservable latent volatilities ς_i's and thus does not have a closed form. Statistical inferences for the SV models can be carried out with Markov chain Monte Carlo (MCMC) simulation methods (Jacquier et al., 2002), which are computationally much harder than those for the GARCH models.

2.3. Diffusion processes

Besides discrete-time models such as the GARCH and SV, continuous-time models are widely employed in modern finance such as option pricing and high-frequency finance. In specific, given security price S_t, $t \in [0, T]$, we consider the following stochastic differential equation

$$dS_t = \mu_t\, S_t\, dt + \sigma_t\, S_t\, dW_t, \tag{7}$$

where W_t is a standard Brownian motion process, μ_t and σ_t are the instantaneous drift and volatility, respectively. In particular, the well-known Black–Scholes model (Black and Scholes, 1973) is constructed on (7) with constant μ_t and σ_t. However, empirical financial series tend to be highly heteroskedastic. To capture this important feature, we allow σ_t to be random and assume it obeys some stochastic differential equation. Such σ_t is called stochastic volatility. Further let $X_t = \log S_t$ be the log price process, from (7) and by Itô's lemma, we can show that

$$dX_t = (\gamma_0 + \gamma_1 \sigma_t^2)\, dt + \sigma_t\, dW_t$$

where the drift of X_t includes the volatility term σ_t^2, which shares the same structure as (3). Since the GARCH models are designed to model the increments of log prices, parametrizing the GARCH drift μ_i by (3) is very natural from a diffusion point of view.

Given discretely sampled data from a diffusion model, the likelihood function is only available in a closed form for a handful of simple cases. Thus, approximations of the likelihood function are typically necessary, and statistical inferences are much harder than those of the GARCH models (Aït-Sahalia, 2002, 2008; Beskos et al., 2005, 2006; Durham and Gallant, 2002).

2.4. Diffusion limits of the discrete-time models

In this section, we define the approximating processes of the GARCH and SV models and show that they share the same diffusion limit. Divide the interval $[0, T]$ into n subintervals of length $s_n = T/n$ and set $t_i = is_n$, $i = 1, \ldots, n$. Given i.i.d. standard normal random variables ε_i's, let

$$\xi_i = \kappa_1(\log \varepsilon_i^2 - \kappa_0), \qquad \zeta_i = 2^{-1/2}(\varepsilon_i^2 - 1), \tag{8}$$

where κ_0 and κ_1 are generic constants such that

$$\kappa_0 = \mathrm{E} \log \varepsilon_1^2 \approx -1.27, \qquad \kappa_1 = [\mathrm{Var}(\log \varepsilon_1^2)]^{-1/2} \approx 0.45. \tag{9}$$

First, we define the approximating process for the continuous multiplicative GARCH(1,1) model and will show in Section 2.5 that there is a one-to-one correspondence between the discrete multiplicative GARCH (1,1) model defined by (1)–(2) and the approximating process specified in (10)–(11). Let

$$X_{n,i}^G - X_{n,i-1}^G = (\gamma_0 + \gamma_1 \tau_{n,i}^2)s_n + z_{n,i}, \qquad z_{n,i} = Z_{n,i} - Z_{n,i-1} = \tau_{n,i} s_n^{1/2} \varepsilon_i, \tag{10}$$

$$\log \tau_{n,i+1}^2 = \beta_0 s_n + (1 + \beta_1 s_n) \log \tau_{n,i}^2 + \beta_2 s_n^{1/2} \xi_i. \tag{11}$$

For any $t \in [t_i, t_{i+1})$, denote $X_{n,t}^G = X_{n,i}^G, Z_{n,t} = Z_{n,i}$, and $\tau_{n,t}^2 = \tau_{n,i}^2$. As $n \to \infty$, Nelson (1990) demonstrated that the normalized partial sum process of (ε_i, ξ_i) converges in distribution to a planar Wiener process and the process $(X_{n,t}^G, Z_{n,t}, \tau_{n,t}^2)$ converges weakly to a diffusion process (X_t, Z_t, σ_t^2) satisfying the following stochastic differential equation system

$$dX_t = (\gamma_0 + \gamma_1 \sigma_t^2)dt + dZ_t, \qquad dZ_t = \sigma_t dW_{1,t}, \tag{12}$$

$$d \log \sigma_t^2 = (\beta_0 + \beta_1 \log \sigma_t^2)\, dt + \beta_2 dW_{2,t}, \tag{13}$$

where $W_{1,t}$ and $W_{2,t}$ stand for two independent standard Wiener processes. We refer the process (X_t, Z_t, σ_t^2) the diffusion limit of the approximating process $(X_{n,t}^G, Z_{n,t}, \tau_{n,t}^2)$.

We now define the approximating process for the continuous SV model and will show in Section 2.5 that there is a one-to-one correspondence between the discrete SV model defined by (5)–(6) and the approximating process specified in (14)–(15). Let

$$X_{n,i}^S - X_{n,i-1}^S = (\gamma_0 + \gamma_1 \varsigma_{n,i}^2)s_n + y_{n,i}, \qquad y_{n,i} = Y_{n,i} - Y_{n,i-1} = \varsigma_{n,i} s_n^{1/2} \varepsilon_i, \tag{14}$$

$$\log \varsigma_{n,i+1}^2 = \beta_0 s_n + (1 + \beta_1 s_n) \log \varsigma_{n,i}^2 + \beta_2 s_n^{1/2} \delta_i, \tag{15}$$

where ε_i's and δ_i's are independent standard normal random variables. For any $t \in [t_i, t_{i+1})$, denote $X_{n,t}^S = X_{n,i}^S$, $Y_{n,t} = Y_{n,i}$, and $\varsigma_{n,t}^2 = \varsigma_{n,i}^2$. As $n \to \infty$, since the normalized partial sum process of the i.i.d. sequence $(\varepsilon_i, \delta_i)$ converges in distribution to a planar Wiener process, the process $(X_{n,t}^S, Y_{n,t}, \varsigma_{n,t}^2)$ converges weakly to the same diffusion process (X_t, Y_t, σ_t^2) defined by (12)–(13) with Z_t replaced by Y_t. Thus, we show that the approximating processes specified by (10)–(11) and (14)–(15) share the same diffusion limit described by (12)–(13).

2.5. Approximating processes of the discrete-time models

In this section, we connect the approximating processes of the GARCH and SV models with their respective specifications provided in Sections 2.1 and 2.2.

For the GARCH model, using the relationship between ξ_i and ε_i^2 given by (8)–(9) such that $\xi_i = \kappa_1(\log \varepsilon_i^2 - \kappa_0)$, we may rewrite Eq. (11) by

$$\log \tau_{n,i+1}^2 = \beta_0 s_n - \beta_2 s_n^{1/2} \kappa_0 \kappa_1 + (1 + \beta_1 s_n) \log \tau_{n,i}^2 + \beta_2 s_n^{1/2} \kappa_1 \log \varepsilon_i^2. \tag{16}$$

Comparing the volatility specification for the GARCH(1,1) model by (2) and for the approximating process by (16), we demonstrate that they share the same structure with

$$\alpha_0 = \beta_0 s_n - \beta_2 s_n^{1/2} \kappa_0 \kappa_1, \qquad \alpha_1 = 1 + \beta_1 s_n, \qquad \alpha_2 = \beta_2 s_n^{1/2} \kappa_1.$$

Further comparing the drift specification for the GARCH(1,1) model by (1) and (3) and for the approximating process by (10), we show their relative relationship in the following

$$c_0 = \gamma_0 s_n, \qquad c_1 = \gamma_1 s_n.$$

Therefore, the parameters γ's and β's are, respectively, the rescaled versions of drift parameters c's and local reparameterization of volatility parameters α's. It follows that the diffusion limit in (12)–(13) can be established for the GARCH model in (1)–(2).

For the SV model, comparing the volatility specification for the SV model by (6) and for the approximating process by (15), we have

$$\alpha_0 = \beta_0 s_n, \qquad \alpha_1 = 1 + \beta_1 s_n, \qquad \alpha_2 = \beta_2 s_n^{1/2}.$$

There is also a one-to-one correspondance between the drift terms in (5) and in (14). Thus, the diffusion limit in (12)–(13) can be obtained for the SV model in (5)–(6). We can further conclude that the GARCH and SV models share the same diffusion limit.

In the rest of the paper, we assume $T = 1$ and thus $s_n = 1/n$, moreover, the initial values $e^{\beta_3} \equiv \tau_{n,0}^2 = \varsigma_{n,0}^2 = \sigma_0^2$ and $X_{t,0}^G = X_{t,0}^S = X_0$ are known constants.

3. Asymptotic analysis of likelihood ratios

In this section, we examine the likelihood ratios for making statistical inferences based on the approximating processes of the GARCH (10)–(11) and SV (14)–(15) models as well as the discretized process of their diffusion limit (12)–(13). We show that although the GARCH and SV models share the same diffusion limit, their respective likelihood ratios differ even in the asymptotic sense. Moreover, since the likelihood ratio based on the approximating SV model is intractable, we further adopt the idea of quasi-likelihood by plugging samples obtained from the approximating SV model into the GARCH likelihood. A similar quasi-likelihood function can be established for the diffusion model by plugging discrete samples obtained from the diffusion model into the GARCH likelihood. We show in Theorem 1 that the quasi-likelihood ratios for the SV and diffusion models share the same asymptotic limit, which differs from the limit of GARCH likelihood ratio.

Let us specify the testing problem and define the likelihood ratios. Set $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3)$, $\boldsymbol{\gamma} = (\gamma_0, \gamma_1)$ and $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\gamma})$. The asymptotic study in statistics often needs to investigate the behavior of distributions and likelihoods in a shrinking neighborhood around certain parameters. Thus, we consider parameters in a $n^{-1/2}$-shrinking neighborhood of $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\gamma})$. For a given $\boldsymbol{\theta} = (\boldsymbol{\beta} + n^{-1/2}\boldsymbol{\varphi}, \boldsymbol{\gamma})$, denote $\boldsymbol{\vartheta} = (\boldsymbol{\varphi}, \boldsymbol{\gamma})$. In particular, for a fixed $\boldsymbol{\beta}^*$, consider $\boldsymbol{\theta}^* \equiv (\boldsymbol{\beta}^*, \mathbf{0})$, or equivalently, $\boldsymbol{\vartheta}^* \equiv (\mathbf{0}, \mathbf{0})$, we are interested in testing the null $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}^*$, or equivalently, $\boldsymbol{\vartheta} = \boldsymbol{\vartheta}^*$, against the alternative $H_a : \boldsymbol{\theta} = (\boldsymbol{\beta}^* + n^{-1/2}\boldsymbol{\varphi}, \boldsymbol{\gamma})$ for $\boldsymbol{\vartheta} = (\boldsymbol{\varphi}, \boldsymbol{\gamma}) \neq \mathbf{0}$. Since there is a one-to-one correspondence between $\boldsymbol{\theta}$ and $\boldsymbol{\vartheta}$, we use $\boldsymbol{\vartheta}$ instead of $\boldsymbol{\theta}$, and write $\boldsymbol{\vartheta}^*$ for $\boldsymbol{\theta}^*$ when there is no confusion. We note that the likelihood processes have non-degenerate limiting distributions over the entire $\boldsymbol{\gamma}$ and over only a $n^{-1/2}$-shrinking neighborhood of $\boldsymbol{\beta}$. Thus, only a shrinking neighborhood of the volatility parameter $\boldsymbol{\beta}$ is studied.

We now examine the asymptotic behaviors of the likelihood ratios for the testing problem. Denote by $L_{n,1}(\boldsymbol{\vartheta})$ the likelihood function of the GARCH approximating process X_{n,t_i}^G. Let $\Lambda_{n,1}(\boldsymbol{\vartheta}) = L_{n,1}(\boldsymbol{\vartheta})/L_{n,1}(\boldsymbol{\vartheta}^*)$ be the corresponding likelihood ratio given respective parameters of $\boldsymbol{\vartheta}$ in H_a and $\boldsymbol{\vartheta}^*$ in H_0. We can show that as $n \to \infty$, $\Lambda_{n,1}(\boldsymbol{\vartheta})$ converges in distribution to $\Lambda_1(\boldsymbol{\vartheta})$ defined by

$$\Lambda_1(\boldsymbol{\vartheta}) = \exp\left[\frac{1}{\sqrt{2}}\int_0^1 V_t\, dW_{3,t} - \frac{1}{4}\int_0^1 V_t^2\, dt + \int_0^1 \sigma_{t,0}^{-1}\left(\gamma_0 + \gamma_1\sigma_{t,0}^2\right) dW_{1,t} - \frac{1}{2}\int_0^1 \sigma_{t,0}^{-2}\left(\gamma_0 + \gamma_1\sigma_{t,0}^2\right)^2 dt\right], \quad (17)$$

where $\sigma_{t,0}^2$ is the diffusion volatility σ_t^2 in (13) given the null parameter $\boldsymbol{\vartheta} = \boldsymbol{\vartheta}^*$,

$$V_t = \sum_{i=0}^{3} \varphi_i \frac{\partial \log \sigma_{t,0}^2}{\partial \beta_i^*}, \quad (18)$$

and $W_{1,t}, W_{2,t}, W_{3,t}$ are standard Brownian motions. We note that $W_{1,t}$, $W_{2,t}$, and $W_{3,t}$ are associated with the empirical processes of ε_i in (10), standardized $\log \varepsilon_i^2$ in (16) [see its standardized version ξ_i in (8)] and ε_i^2 appeared in the GARCH likelihood (4) [see its standardized version ζ_i in (8)]. Note that $\log \varepsilon_i^2$ and ε_i^2 are correlated, but each is uncorrelated with ε_i, thus, $W_{1,t}$ is independent of $W_{2,t}$ and $W_{3,t}$ with $\operatorname{corr}(W_{2,t}, W_{3,t}) = \operatorname{corr}(\log \varepsilon_1^2, \varepsilon_1^2) \approx 0.64$.

Denote by $L_{n,2}(\boldsymbol{\vartheta})$ the likelihood function of the SV approximating process X_{n,t_i}^S. Let $\Lambda_{n,2}(\boldsymbol{\vartheta}) = L_{n,2}(\boldsymbol{\vartheta})/L_{n,2}(\boldsymbol{\vartheta}^*)$ be the corresponding likelihood ratio given respective parameters of $\boldsymbol{\vartheta}$ in H_a and $\boldsymbol{\vartheta}^*$ in H_0. We will demonstrate that $\Lambda_{n,2}(\boldsymbol{\vartheta})$

relates to $\Lambda_2(\boldsymbol{\vartheta})$ defined by

$$\Lambda_2(\boldsymbol{\vartheta}) = \mathrm{E}_{W_2}\left\{\exp\left[\frac{1}{\sqrt{2}}\int_0^1 V_t\, dW_{4,t} - \frac{1}{4}\int_0^1 V_t^2\, dt + \int_0^1 \sigma_{t,0}^{-1}\left(\gamma_0 + \gamma_1\,\sigma_{t,0}^2\right) dW_{1,t} \right.\right.$$
$$\left.\left. -\frac{1}{2}\int_0^1 \sigma_{t,0}^{-2}\left(\gamma_0 + \gamma_1\,\sigma_{t,0}^2\right)^2 dt\right]\right\}, \tag{19}$$

where $W_{4,t}$ is a standard Brownian motion process independent of $W_{1,t}$, $W_{2,t}$, $W_{3,t}$, and E_{W_2} denotes the expectation taken with respect to $W_{2,t}$ in V_t and $\sigma_{t,0}^2$. We note that $W_{1,t}$, $W_{2,t}$, and $W_{4,t}$ are associated with the empirical processes of ε_i in (14), δ_i in (15), and ε_i^2 appeared in the conditional normal likelihood given δ_i. Since ε_i, δ_i, and ε_i^2 are uncorrelated, $W_{1,t}$, $W_{2,t}$, and $W_{4,t}$ are independent. Moreover, given the fact that ε_i and δ_i are independent while the SV volatility $\varsigma_{n,i}^2$'s in (15) are latent, the joint density of $y_{n,i}$'s is equal to the product of the conditional normal density of each $y_{n,i}$ given $\delta_1, \ldots, \delta_n$ with expectation taken with respect to all δ_i's whose empirical process converges to $W_{2,t}$. Thus, $\Lambda_2(\boldsymbol{\vartheta})$ involves the expectation E_{W_2}. As a matter of fact, neither $\Lambda_{n,2}(\boldsymbol{\vartheta})$ nor $\Lambda_2(\boldsymbol{\vartheta})$ are tractable even asymptotically. It is only under some very special case such as given zero drift and deterministic volatility, $\Lambda_{n,2}(\boldsymbol{\vartheta})$ converges in distribution to $\Lambda_2(\boldsymbol{\vartheta})$ as $n \to \infty$. In particular, we note that the limiting likelihood ratios in (17) and (19) are different.

Note that the likelihood $\Lambda_1(\boldsymbol{\vartheta})$ in (17) has an explicit expression but $\Lambda_2(\boldsymbol{\vartheta})$ in (19) has no closed form. We thus develop quasi-likelihood ratios for the study of SV and diffusion models. In specific, consider a quasi-likelihood $L_{n,0}(\boldsymbol{\vartheta})$ that is obtained by plugging samples from the SV approximating process X_{n,t_i}^S into the GARCH likelihood $L_{n,1}(\boldsymbol{\vartheta})$. We define the corresponding quasi-likelihood ratio $\Lambda_{n,0}(\boldsymbol{\vartheta}) = L_{n,0}(\boldsymbol{\vartheta})/L_{n,0}(\boldsymbol{\vartheta}^*)$ given respective parameters of $\boldsymbol{\vartheta}$ in H_a and $\boldsymbol{\vartheta}^*$ in H_0. Similarly, denote by $\Lambda_n(\boldsymbol{\vartheta})$ the quasi-likelihood ratio that is obtained by plugging discrete samples X_{t_i} obtained from the diffusion model (12)–(13) into the GARCH likelihood ratio. The theorem below derives an explicit asymptotic expression for the proposed quasi-likelihood ratios.

Theorem 1. *As $n \to \infty$, both the quasi-likelihood ratios for the SV model, $\Lambda_{n,0}(\boldsymbol{\vartheta})$, and for the diffusion model, $\Lambda_n(\boldsymbol{\vartheta})$, converge in distribution to $\Lambda(\boldsymbol{\vartheta})$, where*

$$\Lambda(\boldsymbol{\vartheta}) = \exp\left[\frac{1}{\sqrt{2}}\int_0^1 V_t\, dW_{4,t} - \frac{1}{4}\int_0^1 V_t^2\, dt + \int_0^1 \sigma_{t,0}^{-1}\left(\gamma_0 + \gamma_1\,\sigma_{t,0}^2\right) dW_{1,t}\right.$$
$$\left. -\frac{1}{2}\int_0^1 \sigma_{t,0}^{-2}\left(\gamma_0 + \gamma_1\,\sigma_{t,0}^2\right)^2 dt\right]. \tag{20}$$

where V_t is defined in (18), and $\sigma_{t,0}^2$ denotes the diffusion volatility σ_t^2 in (13) with $\boldsymbol{\vartheta} = \boldsymbol{\vartheta}^*$.

Proof. From the GARCH likelihood (3) we easily derive the following expression for the log GARCH likelihood ratio,

$$\log \Lambda_{n,1}(\boldsymbol{\vartheta}) = \frac{1}{2}\sum_{i=1}^n\left(1 - \frac{\tau_{n,t_i,0}^2}{\tau_{n,t_i}^2}\right)(\varepsilon_i^2 - 1) + \frac{1}{2}\sum_{i=1}^n\left(1 + \log\frac{\tau_{n,t_i,0}^2}{\tau_{n,t_i}^2} - \frac{\tau_{n,t_i,0}^2}{\tau_{n,t_i}^2}\right)$$
$$+ n^{-1/2}\sum_{i=1}^n\left(\gamma_0 + \gamma_1\tau_{n,t_i}^2\right)\tau_{n,t_i,0}\,\tau_{n,t_i}^{-2}\,\varepsilon_i - \frac{1}{2n}\sum_{i=1}^n\left(\gamma_0 + \gamma_1\tau_{n,t_i}^2\right)^2\tau_{n,t_i}^{-2}, \tag{21}$$

where $\tau_{n,t_i,0}$ denotes the GARCH volatility τ_{n,t_i} in (11) with $\boldsymbol{\vartheta} = \boldsymbol{\vartheta}^*$ (see more details in Appendix A.2). To obtain the quasi-likelihood $\Lambda_n(\boldsymbol{\vartheta})$, we plug observations X_{n,t_i}^S from the SV approximating process (14) or discrete observations X_{t_i} from the diffusion process (12) into the log GARCH likelihood ratio $\log \Lambda_{n,1}(\boldsymbol{\vartheta})$ in (21). Since Appendix A.3 shows that the SV and diffusion samples behave the same asymptotically, we adopt the diffusion model here for demonstration and replace τ_{n,t_i}^2 and ε_i in the above log GARCH likelihood ratio (21) by the corresponding $\bar{\sigma}_{n,t_i}^2$ and $\check{\varepsilon}_i$, where $\check{\varepsilon}_i$ are i.i.d. standard normal random variables. In specific, $\bar{\sigma}_{n,t_i}^2$ and $\check{\varepsilon}_i$ are given by

$$\bar{\sigma}_{n,t}^2 = n\int_{t-1/n}^t \sigma_u^2\, du, \qquad \check{\varepsilon}_i = \left(\bar{\sigma}_{n,t_i}^2/n\right)^{-1/2}\int_{t_{i-1}}^{t_i}\sigma_u\, dW_{1,u}, \tag{22}$$

where σ_t^2 is defined in (13) [see (A.7) and (A.8) in Appendix A.3 for more details about their properties]. Then we obtain the log quasi-likelihood ratio for the diffusion model,

$$\log \Lambda_n(\boldsymbol{\vartheta}) = \frac{1}{2}\sum_{i=1}^n\left(1 - \frac{\bar{\sigma}_{n,t_i,0}^2}{\bar{\sigma}_{n,t_i}^2}\right)(\check{\varepsilon}_i^2 - 1) + \frac{1}{2}\sum_{i=1}^n\left(1 + \log\frac{\bar{\sigma}_{n,t_i,0}^2}{\bar{\sigma}_{n,t_i}^2} - \frac{\bar{\sigma}_{n,t_i,0}^2}{\bar{\sigma}_{n,t_i}^2}\right)$$
$$+ n^{-1/2}\sum_{i=1}^n\left(\gamma_0 + \gamma_1\bar{\sigma}_{n,t_i}^2\right)\bar{\sigma}_{n,t_i,0}\,\bar{\sigma}_{n,t_i}^{-2}\check{\varepsilon}_i - \frac{1}{2n}\sum_{i=1}^n\left(\gamma_0 + \gamma_1\bar{\sigma}_{n,t_i}^2\right)^2\bar{\sigma}_{n,t_i}^{-2}$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{2}} \int_0^1 \bar{V}_{n,t}\, dW_{4,t}^{(n)} + \frac{1}{2n} \sum_{i=1}^{n} \bar{H}_{n,t_i} + \int_0^1 \left(\gamma_0 + \gamma_1 \bar{\sigma}_{n,t}^2\right) \bar{\sigma}_{n,t,0}\, \bar{\sigma}_{n,t}^{-2}\, dW_{1,t}^{(n)} \\
&\quad - \frac{1}{2n} \sum_{i=1}^{n} \left(\gamma_0 + \gamma_1 \bar{\sigma}_{n,t_i}^2\right)^2 \bar{\sigma}_{n,t_i}^{-2},
\end{aligned} \tag{23}
$$

where

$$
\bar{V}_{n,t} = n^{1/2} \left(1 - \frac{\bar{\sigma}_{n,t,0}^2}{\bar{\sigma}_{n,t}^2}\right), \qquad \bar{H}_{n,t} = n\left[n^{-1/2}\bar{V}_{n,t} + \log\left(1 - n^{-1/2}\bar{V}_{n,t}\right)\right], \tag{24}
$$

and

$$
W_{1,t}^{(n)} = n^{-1/2} \sum_{i=1}^{[nt]} \check{\varepsilon}_i, \qquad W_{4,t}^{(n)} = (2n)^{-1/2} \sum_{i=1}^{[nt]} (\check{\varepsilon}_i^2 - 1).
$$

By strong approximation (Komlós et al., 1975, 1976), we can realize the processes $W_{1,t}^{(n)}$, $W_{4,t}^{(n)}$, $W_{1,t}$, $W_{2,t}$, and $W_{4,t}$ on some common probability spaces, and approximate $(W_{1,t}^{(n)}, W_{4,t}^{(n)})$ by $(W_{1,t}, W_{4,t})$ with order $O_p(n^{-1/2}\log^2 n)$. Using the definition of $\bar{\sigma}_{n,t}^2$ in (22) and the explicit solution of σ_t^2 in (13), we have that $\bar{\sigma}_{n,t}^2$ approaches $\sigma_{t,0}^2$, with $\bar{\sigma}_{n,t}^2/\bar{\sigma}_{n,t,0}^2$ approximated by $\sigma_t^2/\sigma_{t,0}^2$. Thus, an application of the Taylor expansion together with the definition of V_t in (18) leads to

$$
\log\left(\frac{\bar{\sigma}_{n,t}^2}{\bar{\sigma}_{n,t,0}^2}\right) \sim \log\left(\frac{\sigma_t^2}{\sigma_{t,0}^2}\right) \sim n^{-1/2} V_t.
$$

Therefore, from (24) we arrive at

$$
\bar{V}_{n,t} \sim n^{1/2}\left[1 - \exp\left(-n^{-1/2} V_t\right)\right] \sim V_t, \qquad \bar{H}_{n,t} \sim -V_t^2/2. \tag{25}
$$

(See Appendix A.3 for more details about above derivations.) Combining the strong approximation with (25) and the expression (23) for $\log \Lambda_n(\boldsymbol{\vartheta})$, we conclude that $\Lambda_n(\boldsymbol{\vartheta})$ converges in distribution to $\Lambda(\boldsymbol{\vartheta})$ and hence prove Theorem 1. □

As indicated in Wang (2002), the GARCH and SV models share the same diffusion limit, but they employ different noise propagation systems in their conditional variances to yield different behaviors in likelihood and thus the statistical non-equivalence. It may be further demonstrated by the fact that the limiting quasi-likelihood ratio $\Lambda(\boldsymbol{\vartheta})$ in (20) is different from the GARCH limiting likelihood ratio $\Lambda_1(\boldsymbol{\vartheta})$ in (17). The explicit expression for the asymptotic quasi-likelihood ratios in Theorem 1 is very handy in developing statistical inferences for these models. For example, the result can be naturally employed to design and carry out hypothesis tests for the SV and diffusion models. Furthermore, we may exploit the difference between $\Lambda_1(\boldsymbol{\vartheta})$ in (17) and $\Lambda(\boldsymbol{\vartheta})$ in (20) to statistically distinguish the GARCH and SV models based on observed data.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The research of Xinyu Song was supported by the Fundamental Research Funds for the Central Universities, China (2018110128), China Scholarship Council (201806485017) and National Natural Science Foundation of China (Grant No. 11871323). The research of Yazhen Wang was supported in part by National Science Foundation (NSF), United States of America Grants DMS-1528735, DMS-1707605, and DMS-1913149.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.spl.2020.108817.

References

Aït-Sahalia, Y., 2002. Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach. Econometrica 70 (1), 223–262.

Aït-Sahalia, Y., 2008. Closed-form likelihood expansions for multivariate diffusions. Ann. Statist. 36 (2), 906–937.

Beskos, A., Papaspiliopoulos, O., Roberts, G.O., 2006. Retrospective exact simulation of diffusion sample paths with applications. Bernoulli 12 (6), 1077–1098.

Beskos, A., Roberts, G.O., et al., 2005. Exact simulation of diffusions. Ann. Appl. Probab. 15 (4), 2422–2444.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. J. Political Econ. 81 (3), 637–654.
Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. J. Econometrics 31 (3), 307–327.
Brown, L.D., Wang, Y., Zhao, L.H., 2003. On the statistical equivalence at suitable frequencies of GARCH and stochastic volatility models with the corresponding diffusion model. Statist. Sinica 13 (4), 993–1013.
Durham, G.B., Gallant, A., 2002. Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. J. Bus. Econom. Statist. 20 (3), 297–338.
Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50 (4), 987–1007.
Jacquier, E., Polson, N.G., Rossi, P.E., 2002. Bayesian analysis of stochastic volatility models. J. Bus. Econom. Statist. 20 (1), 69–87.
Komlós, J., Major, P., Tusnády, G., 1975. An approximation of partial sums of independent RV'-s, and the sample DF. I. Z. Wahrscheinlichkeitstheor. Verwandte Geb. 32 (1–2), 111–131.
Komlós, J., Major, P., Tusnády, G., 1976. An approximation of partial sums of independent RV's, and the sample DF. II. Probab. Theory Related Fields 34 (1), 33–58.
Nelson, D.B., 1990. ARCH models as diffusion approximations. J. Econometrics 45 (1–2), 7–38.
Wang, Y., 2002. Asymptotic nonequivalence of GARCH models and diffusions. Ann. Statist. 30 (3), 754–783.