Convergence Speed of GARCH Option Price to Diffusion Option Price*

Jin-Chuan Duan

National University of Singapore and University of Toronto

Yazhen Wang

University of Connecticut and National Science Foundation

Jian Zou

University of Connecticut

To appear in International Journal of Theoretical and Applied Finance

May 9, 2009

Duan is Director of Risk Management Institute and Cycle & Carriage Professor of Finance at National University of Singapore and Manulife Chair in Financial Services at Rotman School of Management, University of Toronto. Wang is Professor of Statistics at University of Connecticut and Program Director at National Science Foundation. His research was partially supported by NSF grant DMS-0504323, and this material was based on work supported by the National Science Foundation, while working at the Foundation. Any opinion, finding, and conclusions or recommendations expressed in this material; are those of the authors and do not necessarily reflect the views of the National Science Foundation. Zou is a Ph. D student at University of Connecticut. The authors thank the editor, associate editor, and an anonymous referee for comments and suggestions, which led to significant improvements of the paper.

Abstract

It is well known that as the time interval between two consecutive observations shrinks to zero, a properly constructed GARCH model will weakly converge to a bivariate diffusion. Naturally the European option price under the GARCH model will also converge to its bivariate diffusion counterpart. This paper investigates the convergence speed of the GARCH option price. We show that the European option prices under the two corresponding models are equal up to an order near the square root of the length of discrete time interval.

Key words: convergence rate, European option, and stochastic volatility.

The Journal of Economic Literature index number: G12.

The 2000 Mathematics Subject Classification: 60H30, 62P05, 62M10.

Running Head: **GARCH Option Price and Diffusion Limit**

1 Introduction

Stochastic volatility models are commonly employed in financial analysis. Popular models include a class of continuous-time bivariate diffusion models such as Hull and White (1987), Wiggins (1987), Scott (1987), Melino and Turnbull (1990), Stein and Stein (1991), and Heston (1993). The discrete-time ARCH/GARCH model of Engle (1982) and Bollerslev (1986) and their variants such as Nelson (1991) and Engle and Ng (1993) constitute another important class of models with stochastic volatility. These two classes of stochastic volatility models are often employed in modeling financial time series and for derivative pricing. There is an intriguing link between these two classes of models. As the length of the discrete time interval in the GARCH model shrinks to zero, the GARCH model weakly converges to a bivariate diffusion (Nelson 1990; Duan, 1997). Such weak convergence result combined with uniform integrability that is satisfied by typical option payoff functions leads to the convergence of the GARCH option price to the option price under the corresponding bivariate diffusion model. In fact, such convergence has been demonstrated numerically in Ritchken and Trevor (1999) and Lyuu and Wu (2005).

In this paper we study the convergence speed of the GARCH European option price to its bivariate diffusion limit price, particularly under the exponential GARCH(1,1) specification, EGARCH(1,1) for short. Although the convergence result based on weak convergence and uniform integrability is of great value, it is silent as to how fast such convergence takes place. Our derivation relies on tackling the option price formulas directly and employs more accurate probability approximations. We show that the EGARCH(1,1) option price and the option price under the corresponding bivariate diffusion are equal up to an order near the square root of the length of discrete time interval. To the best of our knowledge, our paper is the first to derive a convergence result for the GARCH option price with a known convergence rate.

The remainder of this paper is organized as follows. Section 2 reviews the weak convergence result and then presents the main theorem on the convergence rate of the EGARCH(1,1) European option price to its diffusion limit price. Section 3 offers the conclusion. All proofs are given in Section 4. In Section 4.1, we discretize the continuous-time bivariate diffusion model to introduce a discrete-time stochastic volatility model and use it as an intermediate model to bridge the gap between the GARCH and continuous-time bivariate diffusion models. The proofs in various steps are provided in Sections 4.2-4.6. The Appendix collects all probability inequalities needed for establishing the main result.

2 Option valuation models

Denote by $\mathcal{C}(t, x; T, K; y)$ the price of a European call option with stock price x at valuation time t, strike price K and expiration time T, where y is the volatility at the time t. We use subscript "BS", "D", and "G" to label the prices associated with the Black-Scholes, diffusion and GARCH models. For example, $\mathcal{C}_{BS}(t, x; T, K; \sigma)$ denotes the price of a European call option based on the Black-Scholes model. Although the main issue of this paper centers around the price/volatility system under the risk-neutral pricing measure P, we will start with a brief description of the GARCH option pricing theory that bridges the system under the physical probability measure P_0 and that under the risk-neutral pricing measure P.

2.1 The GARCH and GARCH option pricing models

Our theoretical analysis only addresses the EGARCH(1,1) model. The idea can be applied to other GARCH(1,1) specifications. Define a discrete-time EGARCH(1,1) model over time interval $[0, T]$ under the physical probability measure P_0. The interval is divided into nT subintervals of length $1/n$ and set $t_k = k/n$, $k = 0, 1, \cdots, nT$. This EGARCH(1,1) model assumes asset price process S_{G,n,t_k}, $k = 1, \cdots, nT$, to follow

$$\begin{aligned}
\log \frac{S_{G,n,t_k}}{S_{G,n,t_{k-1}}} &= \left(r + \lambda \sigma_{G,n,t_k} - \sigma^2_{G,n,t_k}/2\right) n^{-1} + \sigma_{G,n,t_k} \epsilon_k n^{-1/2}, &(1)\\
\log \sigma^2_{G,n,t_{k+1}} &= \alpha n^{-1} + (1 + \vartheta_1 n^{-1}) \log \sigma^2_{G,n,t_k} + \vartheta_2 \epsilon_k n^{-1/2} \\
&\quad + (1 - 2/\pi)^{-1/2} \vartheta_3 \left(|\epsilon_k| - (2/\pi)^{1/2}\right) n^{-1/2}, &(2)
\end{aligned}$$

where r is constant risk-free interest rate, λ is the risk premium per unit of standard deviation, ϵ_k is a sequence of *i.i.d.* standard normal random variables under measure P_0, and the conditional variance (or volatility), σ^2_{G,n,t_k}, of $\log \frac{S_{G,n,t_k}}{S_{G,n,t_{k-1}}}$ depends on lagged errors ϵ's. When $n = 1$, the above model becomes the standard EGARCH(1,1) process.

For option pricing, it is customary to change the probability measure to the one with respect to which economic agents could behave as if they were risk-neutral. This new measure is typically referred to as the risk-neutral pricing measure. The measure change is accomplished by essentially absorbing the stochastic discount factor arising from agents' intertemporal marginal rate of substitution into the probability measure. One then attempts to characterize the dynamic asset price system under this new measure. For the GARCH models, the fundamental pricing theory known as the local risk-neutral valuation principle was established in Duan (1995). By which, we can express the asset price system with respect to the risk-neutral pricing measure P as follows:

$$\begin{aligned}
\log \frac{S_{G,n,t_k}}{S_{G,n,t_{k-1}}} &= \left(r - \sigma^2_{G,n,t_k}/2\right) n^{-1} + \sigma_{G,n,t_k} n^{-1/2}\, \varepsilon_k, &(3)\\
\log \sigma^2_{G,n,t_{k+1}} &= \vartheta_0 n^{-1} + (1+\vartheta_1 n^{-1}) \log \sigma^2_{G,n,t_k} + \vartheta_2 \varepsilon_k n^{-1/2} \\
&\quad + (1-2/\pi)^{-1/2} \vartheta_3 \left(|\varepsilon_k - \lambda n^{-1/2}| - (2/\pi)^{1/2}\right) n^{-1/2}, &(4)
\end{aligned}$$

where $\varepsilon_k \equiv \epsilon_k + \lambda n^{-1/2}$ has been shown in Duan (1995) to be a standard normal random variable under measure P and $\vartheta_0 = \alpha + \vartheta_2 \lambda$.

To associate the GARCH model (3)-(4) with a continuous-time model, we extend $(S_{G,n,t_k}, \sigma^2_{G,n,t_{k+1}})$ to $[0,T]$ by letting

$$(S_{G,n,s}, \sigma^2_{G,n,s}) = (S_{G,n,t_k}, \sigma^2_{G,n,t_{k+1}}), \qquad \text{for } s \in [t_k, t_{k+1}), \quad k = 0, \cdots, nT-1.$$

Base on the risk-neutral system in (3) and (4), and the European call option can be computed as

$$\mathcal{C}_G(t,x;T,K;y) = E^P\left[e^{-r\,(T-t)}\,(S_{G,n,T} - K)^+ | (S_{G,n,t}, \sigma_{G,n,t}) = (x,y)\right], \qquad (5)$$

where the expectation is taken under the risk-neutral pricing measure P. For the remainder of the paper, we only deal with the system with respect to P. Thus, we will drop the superscript P for the ease of exposition.

2.2 The diffusion model

As $n \to \infty$, the GARCH process $(S_{G,n,s}, \sigma^2_{G,n,s})$, $s \in [0,T]$, weakly converges to the bivariate diffusion process $(S_{D,s}, \sigma^2_{D,s})$, $s \in [0,T]$. Under the risk-neutral pricing measure P, the diffusion process $(S_{D,s}, \sigma^2_{D,s})$ is governed by the following stochastic differential equation system

$$\begin{aligned} d\log S_{D,s} &= (r - \sigma^2_{D,s}/2)\, ds + \sigma_{D,s}\, dW_{1,s}, & (6)\\ d\log \sigma^2_{D,s} &= (\vartheta_0 + \vartheta_1 \log \sigma^2_{D,s})\, ds + \vartheta_2\, dW_{1,s} + \vartheta_3\, dW_{2,s}, & (7)\end{aligned}$$

where $(W_{1,s}, W_{2,s})$ are two independent standard Brownian motions. The diffusion model (6)-(7) [or the process $(S_{D,s}, \sigma^2_{D,s})$] is referred to as the diffusion limit of the exponential GARCH model (3)-(4). (See Duan 1997, Nelson 1990.)

Under the limiting bivariate diffusion model (6)-(7), the price of the European call option has an expression

$$\mathcal{C}_D(t, x; T, K; y) = E\left[\mathcal{C}_{BS}\left(t, x\, M_D; T, K; \sqrt{\overline{\sigma^2_D}}\right)\right], \tag{8}$$

where $\mathcal{C}_{BS}$ is the Black-Scholes formula [given by (23) in Section 4.3],

$$\overline{\sigma^2_D} = \overline{\sigma^2_D}(T,t) = \frac{1-\rho^2}{T-t}\int_t^T \sigma^2_{D,s,t,y}\, ds, \tag{9}$$

$$M_D = M_D(T, t, \sigma_{D,s,t,y}) = \exp\left(\rho\int_t^T \sigma_{D,s,t,y}\, dB_{V,s} - \frac{\rho^2}{2}\int_t^T \sigma^2_{D,s,t,y}\, ds\right), \tag{10}$$

$$B_{V,s} = \rho\, W_{1,s} + \mathrm{sign}(\vartheta_3)\sqrt{1-\rho^2}\, W_{2,s}, \qquad \rho = \vartheta_2 \,/ \sqrt{\vartheta_2^2 + \vartheta_3^2},$$

$\sigma_{D,s,t,y}$ denotes $\sigma_{D,s}$ begining at $\sigma_{D,t} = y$, and the expectation on the right hand side of (8) is taken over the random source $B_{V,s}$ in $\overline{\sigma^2_D}$ and M_D. (See Fouque et. al. 2000, Heston 1993, Hull and White 1987, Scott 1987, Stein and Stein 1991, Wiggins 1987.)

Empirical studies and economic arguments show a negative correlation (or leverage effect) between asset price and volatility, that is, asset price tends to go down when volatility goes up, so it is nature to expect $\vartheta_2 < 0$ (or $\rho < 0$).

2.3 Main theorem

The GARCH model converges weakly to the diffusion model, and the GARCH European option prices converge to the diffusion European option prices. The theorem below establishes the convergence rate of the GARCH option price to the diffusion option price.

Theorem 2.1 *Assume that model parameters ϑ_i, $i = 0, 1, 2, 3$, risk premium λ, strike price K, and expiration time T are all bounded, time to maturity $T - t$ is bounded below from zero, and $\vartheta_2 \leq 0$. Then as $n \to \infty$,*

$$\mathcal{C}_G(t, x; T, K; y) = \mathcal{C}_D(t, x; T, K; y) + O\left(n^{-1/2} \exp(\Upsilon \log^{1/2} n)\right),$$

where Υ is a positive generic constant, and the convergence rate holds uniformly over ϑ_i, λ, t, T, and K.

In comparison with any positive power of n, the term $\exp(\Upsilon \log^{1/2} n)$ is negligible, so the convergence rate established in Theorem 2.1 is near $n^{-1/2}$. Such convergence rate is very useful in the study of statistical estimation of model parameters based on option data. As the Black-Scholes formula is a smooth function in its variables, Theorem 2.1 indicates that the difference between the diffusion and GARCH implied volatilities converges to zero at the same rate.

It is widely known that the GARCH process has a bivariate diffusion process as its weak limit. This fact is often used in the finance literature to justify a common belief that two models are more or less equivalent. Wang (2002) cast doubt on this common belief and showed that the two models are not statistically equivalent. The statistical non-equivalence is due to the difference in noise propagation in their conditional variances. That difference in turn results in two distinct likelihood processes. For European option pricing however, stochastic equivalence is not essential because it is about the first moment of some integrable function for which weak convergence suffices.

3 Conclusion

The continuous-time diffusion model and the discrete-time GARCH process are widely employed in financial modeling and data analysis. Their relationship is quite complex. Viewed as a discrete-time approximation, a properly constructed GARCH model weakly converges to a bivariate diffusion as the time interval between observations shrinks to zero. Thus, the difference in European option prices must also approach to zero in the limit. However, the rate of convergence is unknown.

This paper focuses on the issue of convergence rate. Suppose there are n observations equally spaced over a fixed time horizon. As n goes to infinity, the GARCH process converges weakly to a bivariate diffusion, and so do the European option prices. This paper is the first to show that the option price convergence is at a rate near $n^{-1/2}$.

4 Proofs

As mentioned in the text, the default probability measure in the analysis is the risk-neutral probability measure P. Thus, expectations E are taken under P, and price and volatility processes are considered under P. All reference materials such as sections and propositions in the Appendix are numbered with label "A".

4.1 Discrete-time stochastic volatility model

Discretizing the diffusion model (6)-(7) naturally yields a discrete-time stochastic volatility (SV) model where the corresponding SV process $(S_{V,n,s}, \sigma^2_{V,n,s})$, $s \in [0, T]$, is defined as follows. For $k = 1, \cdots, nT$, let

$$\log S_{V,n,t_k} - \log S_{V,n,t_{k-1}} = (r - \sigma^2_{V,n,t_k}/2)n^{-1} + \sigma_{V,n,t_k}\varepsilon_k n^{-1/2}, \tag{11}$$

$$\log\,\sigma^2_{V,n,t_{k+1}} = \vartheta_0 n^{-1} + (1+\vartheta_1\, n^{-1})\log\sigma^2_{V,n,t_k} + \vartheta_2\varepsilon_k n^{-1/2} + \vartheta_3\delta_k n^{-1/2}, \tag{12}$$

where ε_k and δ_k are *i.i.d.* N(0,1) random variables, and $(S_{V,n,s}, \sigma^2_{V,n,s}) = (S_{V,n,[ns]/n}, \sigma^2_{V,n,([ns]+1)/n})$. Of course, $(S_{V,n,s}, \sigma^2_{V,n,s})$ weakly converges to the bivariate diffusion $(S_{D,s}, \sigma^2_{D,s})$ described by (6)-(7).

4.2 Uniform expressions of price and volatility processes

We use "BS", "D", "G" and "V" to label the Black-Scholes, diffusion, GARCH and SV models, and their corresponding price and volatility processes and option price formulas, respectively, and denote their dummy label by Λ. The GARCH and SV price and volatility processes depend on n. For simplicity, we suppress the dependence by dropping off n and writing $(S_{\Lambda,n,s}, \sigma^2_{\Lambda,n,s})$ as $(S_{\Lambda,s}, \sigma^2_{\Lambda,s})$. We assume equal initial values $(S_0, \sigma_0) = (S_{G,0}, \sigma_{G,0}) = (S_{V,0}, \sigma_{V,0}) = (S_{D,0}, \sigma_{D,0})$. From equations (3)-(12) we adopt the following reparameterization and uniform expressions for price and volatility processes,

$$\begin{aligned}
S_{D,s} &= S_0 \exp\left\{ rs - \frac{1}{2}\int_0^s \sigma^2_{D,u}du + \int_0^s \sigma_{D,u} dW_{1,u} \right\}, && (13)\\
d\log\sigma^2_{D,s} &= (\beta_0 + \beta_1 \log\sigma^2_{D,s})ds + \beta_2 dB_{V,s}, && (14)\\
\text{for } \Lambda = V, G, \quad S_{\Lambda,s} &= S_0 \exp\left\{ rs - \frac{1}{2}\int_0^{[ns]/n} \sigma^2_{\Lambda,u}du + \int_0^s \sigma_{\Lambda,u} dW^{(n)}_{1,u} \right\}, && (15)\\
\log\sigma^2_{\Lambda,t_k} - \log\sigma^2_{\Lambda,t_{k-1}} &= (\beta_0 + \beta_1\log\sigma^2_{\Lambda,t_{k-1}})n^{-1} + \beta_2(B^{(n)}_{\Lambda,t_{k-1}} - B^{(n)}_{\Lambda,t_{k-2}}),\ \sigma^2_{\Lambda,s} = \sigma^2_{\Lambda,[ns]/n}, && (16)
\end{aligned}$$

where $B_{V,s}$ is a standard Brownian motion defined by

$$B_{V,s} = \rho\, W_{1,s} + \text{sign}(\vartheta_2)\sqrt{1-\rho^2}\, W_{2,s}, \qquad d < W_{1,s}, B_{V,s} >= \rho ds, \tag{17}$$

$$\beta_0 = \vartheta_0 - \lambda\vartheta_3, \qquad \beta_1 = \vartheta_1, \qquad \beta_2 = \sqrt{\vartheta_2^2 + \vartheta_3^2}, \qquad \rho = \vartheta_3/\beta_2, \tag{18}$$

$$W^{(n)}_{1,s} = n^{-1/2}\sum_{j=1}^{[ns]} \varepsilon_j, \qquad W^{(n)}_{2,s} = n^{-1/2}\sum_{j=1}^{[n\,s]} \delta_j, \tag{19}$$

$$B^{(n)}_{V,s} = \rho\, W^{(n)}_{1,s} + \text{sign}(\vartheta_2)\sqrt{1-\rho^2}\, W^{(n)}_{2,s}, \tag{20}$$

$$B_{G,s}^{(n)} = n^{-1/2} \sum_{j=1}^{[n\,s]} \left(\rho\, \varepsilon_j + \mathrm{sign}(\vartheta_2)\, \sqrt{1-\rho^2}\, \xi_j \right), \tag{21}$$

$$\xi_k = \left\{ |\varepsilon_k - \lambda\, n^{-1/2}| - (2/\pi)^{1/2} \right\} / (1 - 2/\pi)^{1/2}. \tag{22}$$

4.3 Conventions and Notations

To better track complex processes under different models, and manage technical arguments, we fix the following conventions and notations.

1. Convergence convention. Since weak convergence is invariant to the change of probability spaces so long as distributions remain the same, it is often necessary to put random variables and processes on some common probability spaces and to embed random variables in processes. At such occasions, we will automatically change probability spaces and consider versions of the random variables and the processes on new probability spaces, without altering notations. Because of the convention, when no confusion occurs, we use the same notation for random variables or processes with identical distribution.

2. Parameter convention. The assumptions in Theorem 2.1 guarantee that $(\beta_0, \beta_1, \beta_2)$ belong to a bounded rectangle in $I\!R^2 \times [0, \infty)$, $\rho \in [-1, 0]$, λ and K are bounded above by fixed positive constants, $T \le 1$, and $T - t$ is bounded below by a fixed positive constant. All O's and o's throughout the proofs hold uniformly over $(\beta_0, \beta_1, \beta_2, \rho, \lambda, t, T, K)$.

3. Reserved notations and symbols. Φ and ϕ are reserved for the standard normal distribution and density functions, respectively. We reserve Υ's for generic positive constants whose values are free from $(\beta_0, \beta_1, \beta_2, \rho, \lambda, t, T, K)$, and may change from appearance to appearance, and take $\zeta_n = n^{-1/2} \exp(\Upsilon \log^{1/2} n)$. $a \sim b$ means that their ratio has limit one. Denote by $[d]$ the integer part of d, and for $s \in [0, T]$, set $s_* = [ns]/n$.

4. Notations for the Black-Scholes formula. Notations about the Black-Scholes formula are needed to express their specific dependence on stock price, current time and volatility so that we can differentiate them with respect to these variables.

We express the Black-Scholes formula as follows,

$$\mathcal{C}_{BS}(t,x;T,K;\sigma) = x\,\Phi(H_1) - e^{-r(T-t)}\,K\,\Phi(H_2), \tag{23}$$

where

$$H_1 = H_1(T-t, x/K, \sigma^2) = \frac{\log(x/K) + (r+\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}, \tag{24}$$

$$H_2 = H_2(T-t, x/K, \sigma^2) = H_1(T-t, x/K, \sigma^2) - \sigma\sqrt{T-t}. \tag{25}$$

It is well known that

$$x\,\phi(H_1) = e^{-r(T-t)}\,K\,\phi(H_2), \tag{26}$$

$$\frac{\partial \mathcal{C}_{BS}(t,x;T,K;\sigma)}{\partial x} = \Phi(H_1), \qquad \frac{\partial \mathcal{C}_{BS}(t,x;T,K;\sigma)}{\partial \sigma} = e^{-r(T-t)}\,K\,\phi(H_2)\sqrt{T-t}, \tag{27}$$

and

$$\frac{\partial \mathcal{C}_{BS}(t,x;T,K;\sigma)}{\partial (T-t)} = r\,e^{-r\,(T-t)}\,\Phi(H_2) + e^{-r\,(T-t)}\,K\,\phi(H_2)\,\sigma\,(T-t)^{-1/2}/2. \tag{28}$$

4.4 The pricing formula for the diffusion model

For the bivariate diffusion model (13)-(14), its European option price given by (8) have the following property.

Theorem 4.1 *$\mathcal{C}_D(t,x;K,T;y)$ has derivative with respect to x between 0 and 1 and derivative with respect to y bounded by* $\exp\{\Upsilon_1 + \Upsilon_2|\log y|\}$.

Proof. From the Black-Scholes formula given by (23) and its derivatives in (27), we have

$$\frac{\partial \mathcal{C}_{BS}\left(t, x\,M_D; T, K; \sqrt{\overline{\sigma_D^2}}\right)}{\partial x} = M_D\,\Phi\left(H_1(T-t, x\,M_D/K, \overline{\sigma_D^2})\right), \tag{29}$$

$$\begin{aligned}\frac{\partial \mathcal{C}_{BS}\left(t, x\,M_D; T, K; \sqrt{\overline{\sigma_D^2}}\right)}{\partial y} &= x\,\Phi\left(H_1(T-t, x\,M_D/K, \overline{\sigma_D^2})\right)\,\frac{\partial M_D}{\partial y} \\ &+ \sqrt{T-t}\,e^{-r\,(T-t)}\,K\,\phi\left(H_2(T-t, x\,M_D/K, \overline{\sigma_D^2})\right)\,\frac{1}{2\sqrt{\overline{\sigma_D^2}}}\,\frac{\partial \overline{\sigma_D^2}}{\partial y}.\end{aligned} \tag{30}$$

On the other hand, from Proposition A3.1 [in Appendix] and (9)-(10) we obtain

$$\frac{1}{2\sqrt{\overline{\sigma_D^2}}}\frac{\partial\overline{\sigma_D^2}}{\partial y} = \frac{1}{\sqrt{\overline{\sigma_D^2}}}\frac{(1-\rho^2)}{y\,(T-t)}\int_t^T e^{\beta_1\,(s-t)}\,\sigma^2_{D,s,t,y}\,ds \le \frac{e^{|\beta_1|\,(T-t)}}{y}\sqrt{\overline{\sigma_D^2}},$$

$$\frac{\partial M_D}{\partial y} \;=\; 2\,\rho\,y^{-1}\,M_D\left(\int_t^T e^{\beta_1\,(s-t)}\,\sigma_{D,s,t,y}\,dB_{V,s} - \rho\int_t^T e^{\beta_1\,(s-t)}\,\sigma^2_{D,s,t,y}\,ds\right).$$

With $\Phi \le 1$ and $\phi \le (2\,\pi)^{-1/2}$, Lemma 4.1 below and Proposition A3.3 indicate that the partial derivatives of $\mathcal{C}_{BS}$ in (29)-(30) have bounded expectations. Hence we can exchange the expectation E and the partial derivatives of $\mathcal{C}_{BS}$ in (29)-(30), and use (8) to establish the bounds for the partial derivatives of $\mathcal{C}_D(t,x;T,K;y)$.

Lemma 4.1

$$E[M_D(T,t,\sigma_{D,s,t,y})] \le 1,$$

$$E\left[M_D(T,t,\sigma_{D,s,t,y})\left|\int_t^T e^{\beta_1\,(s-t)}\,\sigma_{D,s,t,y}\,dB_{V,s} - \rho\int_t^T e^{\beta_1\,(s-t)}\,\sigma^2_{D,s,t,y}\,ds\right|\right] \le \exp\left(\Upsilon_1 + \Upsilon_2\,|\log\,y|\right).$$

Proof. Define stopping time

$$\tau_\ell = T \wedge \inf\left\{s\in[t,T] : \int_t^s e^{\beta_1\,(s-u)}\,dB_{V,u} > \ell\right\}. \tag{31}$$

Then $\tau_\ell \to T$ as $\ell\to\infty$, and Proposition A3.1 and the continuity of $\sigma^2_{D,s,t,y}$ imply

$$\sigma^2_{D,s\wedge\tau_\ell,t,y} \le \exp\left\{e^{|\beta_1|}\,|\log\,y^2| + \beta_2\,\ell + |\beta_0|\,(e^{\beta_1}-1)/\beta_1\right\}.$$

Thus, $\{M_D(v,t,\sigma_{D,s\wedge\tau_\ell,t,y}), v\in[t,T]\}$ is an exponential martingale with

$$E[M_D(T,t,\sigma_{D,s\wedge\tau_\ell,t,y})] = 1.$$

By Fatou's lemma,

$$E[M_D(T,t,\sigma_{D,s,t,y})] = E\left[\lim_{\ell\to\infty} M_D(T,t,\sigma_{D,s\wedge\tau_\ell,t,y})\right] \le \liminf_{\ell\to\infty} E[M_D(T,t,\sigma_{D,s\wedge\tau_\ell,t,y})] = 1.$$

This proves the first inequality.

To prove the second one, we change probability measure by shifting $B_{V,s} - B_{V,t}$ to

$$\hat{B}_{V,s} = B_{V,s} - B_{V,t} - \int_t^s \rho\, \sigma_{D,u\wedge\tau_\ell,t,y}\, du. \tag{32}$$

As P and E stand for the original risk-neutral probability and its corresponding expectation, we denote by $\hat{P}$ the new probability and $\hat{E}$ the expectation taken under the new probability. Girsanov theorem shows that $M_D(T, t, \sigma_{D,s\wedge\tau_\ell,t,y})$ is the Radon-Nikodym derivative of $\hat{P}$ with respect to P, and under $\hat{P}$, $\hat{B}_{V,s}$, $s \in [t, T]$ is a standard Brownian motion.

Now we study the behavior of $\sigma_{D,s,t,y}$ under the new probability $\hat{P}$. Denote by $\hat{\sigma}^2_{D,s,t,y}$ the expression of $\sigma^2_{D,s,t,y}$ with $B_{V,s} - B_{V,t}$ replaced by $\hat{B}_{V,s}$. Then first, the probabilistic behavior of $\hat{\sigma}^2_{D,s,t,y}$ under $\hat{P}$ is the same as $\sigma^2_{D,s,t,y}$ under P; second, as $\rho \le 0$ and $\beta_2 \ge 0$, from Proposition A3.1 and the shifting relation between B_V and $\hat{B}_V$ we have

$$\sigma^2_{D,s,t,y} = \hat{\sigma}^2_{D,s,t,y} \exp\left(\rho\, \beta_2 \int_t^s e^{\beta_1 (u-t)}\, \sigma_{D,u\wedge\tau_\ell,t,y}\, du\right) \le \hat{\sigma}^2_{D,s,t,y}. \tag{33}$$

Then

$$\begin{aligned}
& E\left[M_D(T, t, \sigma_{D,s\wedge\tau_\ell,t,y}) \left|\int_t^T e^{\beta_1 (s-t)}\, \sigma_{D,s\wedge\tau_\ell,t,y}\, dB_{V,s} - \rho \int_t^T e^{\beta_1 (s-t)}\, \sigma^2_{D,s\wedge\tau_\ell,t,y}\, ds\right|\right] \\
&= \hat{E}\left[\left|\int_t^T e^{\beta_1 (s-t)}\, \sigma_{D,s\wedge\tau_\ell,t,y}\, d\hat{B}_{V,s}\right|\right] \\
&\le \left(\int_t^T e^{2\,\beta_1 (s-t)}\, \hat{E}[\sigma^2_{D,s\wedge\tau_\ell,t,y}]\, ds\right)^{1/2} \\
&\le \left(\int_t^T e^{2\,\beta_1 (s-t)}\, \hat{E}[\hat{\sigma}^2_{D,s\wedge\tau_\ell,t,y}]\, ds\right)^{1/2} \\
&\le \left(e^{2\,|\beta_1|}\, \hat{E}\left[\sup_{t\le s\le T} \hat{\sigma}^2_{D,s,t,y}\right]\right)^{1/2} \\
&\le \exp(\Upsilon_1 + \Upsilon_2\, |\log\, y|),
\end{aligned}$$

where, of above six lines of equation array, the equality in the second line of the equation array is due to the change of probability and relationship between B_V and $\hat{B}_V$, the inequality in the third line is using the quadratic variation of the stochastic integral with respect to Brownian motion $\hat{B}_V$ under $\hat{P}$, the inequality in the forth line is from (33), and the inequality in the last line is because of Proposition A3.3 and identical distribution of $\hat{\sigma}_D$ under $\hat{P}$ and σ_D under P. Finally, we prove the second inequality by letting $\ell \to \infty$ and applying Fatou's lemma.

4.5 Option pricing for the SV model

This section derives the order for the difference between $\mathcal{C}_V(t, x; T, K; y)$ and $\mathcal{C}_D(t, x; T, K; y)$.

Theorem 4.2

$$\mathcal{C}_V(t, x; T, K; y) = \mathcal{C}_D(t, x; T, K; y) + O(\zeta_n),$$

where $\zeta_n = n^{-1/2} \exp(\Upsilon \log^{1/2} n)$ *is the rate given by Theorem 2.1 in section 2.*

Proof. From (15)-(16) and embedding of $W_1^{(n)}$ in W_1, we have that for $s \in [t_k, t_{k+1})$, $\sigma^2_{V,s}$ and $S_{V,s}$ are constants, and

$$S_{V,s} = S_0 \exp\left\{ r\, t_k - \frac{1}{2} \int_0^{t_k} \sigma^2_{V,u}\, du + \int_0^{t_k} \sigma_{V,u}\, dW_{1,u} \right\}.$$

Note $t_* = [n\, t]/n$ and $T_* = [n\, T]/n$. Then

$$\begin{aligned} \mathcal{C}_V(t, x; T, K; y) &= E[e^{-r\,(T-t)}\, (S_{V,T} - K)_+ | S_{V,t} = x, \sigma_{V,t} = y] \\ &= e^{r\,(T_* - T - t_* + t)}\, E[e^{-r\,(T_* - t_*)}\, (S_{V,T_*} - K)_+ | S_{V,t_*} = x, \sigma_{V,t_*} = y]. \end{aligned}$$

Processes are evaluated on $[t_*, T_*]$, and $B_V^{(n)}$ is embedded in B_V, so on $[t_*, T_*]$, the SV model has the same structure as the diffusion model. Conditional on the path of $B_{V,s}$, we can manipulate the conditional distribution the same way as for the diffusion case and compute the conditional expectation to derive the following expression

$$\mathcal{C}_V(t, x; T, K; y) = e^{r\,(T_* - T - t_* + t)}\, E\left[\mathcal{C}_{BS}\left(t_*, x\, M_V; T_*, K; \sqrt{\sigma^2_V} \right) \right], \tag{34}$$

where

$$\overline{\sigma_V^2} = \overline{\sigma_V^2}(T_*, t_*) = \frac{1-\rho^2}{T_* - t_*} \int_{t_*}^{T_*} \sigma^2_{V,s,t_*,y}\, ds, \tag{35}$$

$$M_V = M_V(T_*, t_*, \sigma_{V,s,t_*,y}) = \exp\left(\rho \int_{t_*}^{T_*} \sigma_{V,s,t_*,y}\, dB_{V,s} - \frac{\rho^2}{2} \int_{t_*}^{T_*} \sigma^2_{V,s,t_*,y}\, ds \right), \tag{36}$$

and $\sigma^2_{V,s,t_*,y}$ is given by Proposition A3.1.

Applying the mean value theorem to $\mathcal{C}_{BS}(t_*, x; T_*, K; \sigma)$ as a function of (x, σ), and using (27) and boundness of Φ and ϕ, we obtain

$$\begin{aligned} &\Big| \mathcal{C}_{BS}\left(t_*, x\, M_V(T_*, t_*, \sigma_{V,s,t_*,y}); T_*, K; \sqrt{\overline{\sigma_V^2}(T_*, t_*)}\right) \\ &\qquad - \mathcal{C}_{BS}\left(t_*, x\, M_D(T_*, t_*, \sigma_{D,s,t_*,y}); T_*, K; \sqrt{\overline{\sigma_D^2}(T_*, t_*)}\right) \Big| \\ &\leq x\, |M_V(T_*, t_*, \sigma_{V,s,t_*,y}) - M_D(T_*, t_*, \sigma_{D,s,t_*,y})| + \Upsilon \left| \sqrt{\overline{\sigma_V^2}(T_*, t_*)} - \sqrt{\overline{\sigma_D^2}(T_*, t_*)} \right|, \end{aligned} \tag{37}$$

where $\overline{\sigma_D^2}$ and M_D are defined in (9)-(10), respectively. Taking expectation on the both side of (37) and using (34), $T - t - 1/n \leq T_* - t_* \leq T - t + 1/n$, and $C_{BS}(t, x; T, K; \sigma) \leq x$ we arrive at

$$\begin{aligned} &|\mathcal{C}_V(t, x; T, K; y) - \mathcal{C}_D(t_*, x; T_*, K; y)| \leq 3\, n^{-1}\, x\, E\left[M_D(T_*, t_*, \sigma_{D,s,t_*,y})\right] \\ &\quad + 3\, x\, E[|M_V(T_*, t_*, \sigma_{V,s,t_*,y}) - M_D(T_*, t_*, \sigma_{D,s,t_*,y})|] + \Upsilon\, E \left| \sqrt{\overline{\sigma_V^2}(T_*, t_*)} - \sqrt{\overline{\sigma_D^2}(T_*, t_*)} \right|. \end{aligned} \tag{38}$$

Again applying the mean value theorem to $\mathcal{C}_{BS}(t, x; T, K; \sigma)$ as a function of $(T - t, x, \sigma)$ and using its derivatives given by (27)-(28) and $T - t - 1/n \leq T_* - t_* \leq T - t + 1/n$, we get

$$\begin{aligned} &\Big| \mathcal{C}_{BS}\left(t_*, x\, M_D(T_*, t_*, \sigma_{D,s,t_*,y}); T_*, K; \sqrt{\overline{\sigma_D^2}(t_*, T_*)}\right) - \\ &\qquad\qquad \mathcal{C}_{BS}\left(t, x\, M_D(T, t, \sigma_{D,s,t,y}); T, K; \sqrt{\overline{\sigma_D^2}(T, t)}\right) \Big| \\ &\leq \Upsilon_1\, n^{-1} \left[\sqrt{\overline{\sigma_D^2}(T_*, t_*)} + \sqrt{\overline{\sigma_D^2}(T, t)} \right] + x\, |M_D(T_*, t_*, \sigma_{D,s,t_*,y}) - M_D(T, t, \sigma_{D,s,t,y})| \\ &\qquad + \Upsilon_2 \left| \sqrt{\overline{\sigma_D^2}(T_*, t_*)} - \sqrt{\overline{\sigma_D^2}(T, t)} \right|. \end{aligned} \tag{39}$$

Lemma 4.1 implies $E(M_D) \leq 1$ and Proposition A3.3 shows $\sqrt{\overline{\sigma_D^2}}$ has bounded expectation. Thus, taking expectation on both sides of (39) and then combining it with (38), we have

$$\begin{aligned}
&|\mathcal{C}_V(t,x;T,K;y) - \mathcal{C}_D(t,x;T,K;y)| \leq n^{-1}\ \{3\,x + \exp(\Upsilon_1 + \Upsilon_2\,|log\,y|)\} \\
&+ 3\,x\,E\,\{|M_V(T_*,t_*,\sigma_{V,s,t_*,y}) - M_D(T_*,t_*,\sigma_{D,s,t_*,y})| + |M_D(T_*,t_*,\sigma_{D,s,t_*,y}) - M_D(T,t,\sigma_{D,s,t,y})|\} \\
&+ \Upsilon_3\,E\left\{\left|\sqrt{\overline{\sigma_V^2}(T_*,t_*)} - \sqrt{\overline{\sigma_D^2}(T_*,t_*)}\right| + \left|\sqrt{\overline{\sigma_D^2}(T_*,t_*)} - \sqrt{\overline{\sigma_D^2}(T,t)}\right|\right\}.
\end{aligned}$$

Now the theorem is a consequence of Lemmas 4.2 and 4.3 below.

Lemma 4.2 *With $t_* = [n\,t]/n$ and $T_* = [n\,T]/n$, we have*

$$E\left|\sqrt{\overline{\sigma_D^2}(T_*,t_*)} - \sqrt{\overline{\sigma_D^2}(T,t)}\right| = O(n^{-1/2}).$$

If $t_* = t$ and $T = T_*$, then

$$E\left|\sqrt{\overline{\sigma_V^2}(T,t)} - \sqrt{\overline{\sigma_D^2}(T,t)}\right| = O(n^{-1/2}).$$

Proof. Because of similarity, we give the argument only for the second result. Set

$$a = \left|\overline{\sigma_V^2} - \overline{\sigma_D^2}\right|, \qquad b = \min\left\{\overline{\sigma_V^2}, \overline{\sigma_D^2}\right\}.$$

Using inequality $\sqrt{1+u} - 1 \leq u/2$ for $u \geq 0$, we have

$$\left|\sqrt{\overline{\sigma_V^2}} - \sqrt{\overline{\sigma_D^2}}\right| = \sqrt{b}\,\left|\sqrt{1 + a/b} - 1\right| \leq a\,b^{-1/2}/2.$$

An application of the Cauchy-Schwartz inequality yields

$$\left(E\left|\sqrt{\overline{\sigma_V^2}} - \sqrt{\overline{\sigma_D^2}}\right|\right)^2 \leq [E(a\,b^{-1/2})]^2 \leq E(a^2)\,E(b^{-1}). \tag{40}$$

The lemma is proved, if $E(b^{-1})$ is bounded and $E(a^2)$ is of order n^{-1}, which we will show below. In fact, first since

$$b^{-1} \leq \left(\overline{\sigma_V^2}\right)^{-1} + \left(\overline{\sigma_D^2}\right)^{-1},$$

Proposition A3.3 implies that $E(b^{-1})$ is bounded. Second,

$$E(a^2) \leq (T-t)^{-1} \int_t^T E\left[|\sigma^2_{V,s,t,y} - \sigma^2_{D,s,t,y}|^2\right] ds$$
$$\leq (T-t)^{-1} \int_t^T E\left[\sigma^4_{D,s,t,y}\right] E\left[|e^{A_{s,t}} - 1|^2\right] ds, \tag{41}$$

where the last inequality is due to the Cauchy-Schwartz inequality, and

$$A_{s,t} = \log \sigma^2_{V,s,t,y} - \log \sigma^2_{D,s,t,y}.$$

Proposition A3.3 shows $E(\sigma^4_{D,s,t,y})$ is bounded, and by Proposition A3.2 we have

$$E\left[|e^{A_{s,t}} - 1|^2\right] = e^{2\,E(A_{s,t})+2\,Var(A_{s,t})} - 2\,e^{E(A_{s,t})+Var(A_{s,t})/2} + 1 = O(n^{-1}).$$

Thus, from (41) we establish n^{-1} order for $E(a^2)$.

Lemma 4.3 *With $t_* = [n\,t]/n$ and $T_* = [n\,T]/n$, we have*

$$E\,|M_D(T_*, t_*, \sigma_{D,s,t_*,y}) - M_D(T, t, \sigma_{D,s,t,y})| = O(\zeta_n).$$

If $t = t_*$ and $T = T_*$, then

$$E\,|M_D(T, t, \sigma_{D,s,t,y}) - M_V(T, t, \sigma_{V,s,t,y})| = O(\zeta_n).$$

Proof. Because of similarity, we prove the second result only. For $t_k \geq t$, $\sigma^2_{V,t_k,t,y}$ relies on $\varepsilon_{n\,t+1}, \cdots, \varepsilon_{k-1}$ and $\delta_{n\,t+1}, \cdots, \delta_{k-1}$, thus $(\varepsilon_k, \delta_k)$ and $\sigma^2_{V,t_k,t,y}$ are independent. Using conditional argument and the normality of $\rho\,\varepsilon_k \pm \sqrt{1-\rho^2}\,\delta_k$ we get

$$\begin{aligned}
&E\left[M_V(t_k, t, \sigma_{V,s\wedge\tau_\ell,t,y})/M_V(t_{k-1}, t, \sigma_{V,s\wedge\tau_\ell,t,y})\right|\varepsilon_{n\,t+1}, \delta_{n\,t+1}, \cdots, \varepsilon_{k-1}, \delta_{k-1}] \\
&= E\left[\exp\left\{\sigma_{V,t_k\wedge\tau_\ell,t,y}\,(\rho\,\varepsilon_k \pm \sqrt{1-\rho^2}\,\delta_k) - \sigma^2_{V,t_k\wedge\tau_\ell,t,y}/2\right\}\middle|\,\varepsilon_{n\,t+1}, \delta_{n\,t+1}, \cdots, \varepsilon_{k-1}, \delta_{k-1}\right] \\
&= 1,
\end{aligned}$$

where stopping time τ_ℓ is defined in (31). This shows that $\{M_V(t_k, t, \sigma_{V,s\wedge\tau_\ell,t,y}), t \leq t_k \leq T\}$ is a martingale with $E[M_V(T, t, \sigma_{V,s\wedge\tau_\ell,t,y}] = 1$. The proof of Lemma 4.1 has shown

$EM_D(T,t,\sigma_{D,s\wedge\tau_\ell,t,y}) = 1$. Thus, $M_D(T,t,\sigma_{D,s\wedge\tau_\ell,t,y}) - M_V(T,t,\sigma_{V,s\wedge\tau_\ell,t,y})$ has zero expectation, and

$$\begin{aligned}
&0.5\, E\,|M_D(T,t,\sigma_{D,s\wedge\tau_\ell,t,y}) - M_V(T,t,\sigma_{V,s\wedge\tau_\ell,t,y})| \\
&\quad = E\{[M_D(T,t,\sigma_{D,s\wedge\tau_\ell,t,y}) - M_V(T,t,\sigma_{V,s\wedge\tau_\ell,t,y})]^+\} \\
&\quad = E\{M_D(T,t,\sigma_{D,s\wedge\tau_\ell,t,y})\,[1-\exp(U_T - Z_T/2)]^+\}, \qquad (42)
\end{aligned}$$

where

$$U_s \;=\; \rho\int_t^s (\sigma_{V,u\wedge\tau_\ell,t,y} - \sigma_{D,u\wedge\tau_\ell,t,y})\, d\hat{B}_{V,u}, \quad Z_s = \rho^2 \int_t^s (\sigma_{V,u\wedge\tau_\ell,t,y} - \sigma_{D,u\wedge\tau_\ell,t,y})^2\, du, \qquad (43)$$

and $\hat{B}_{V,s}$ is defined in (32). As in the proof of Lemma 4.1, we take probability transformation by shifting $B_{V,s} - B_{V,t}$, $s \in [t,T]$, to $\hat{B}_{V,s}$, $s \in [t,T]$ and use $\hat{P}$ and $\hat{E}$ to stand for the new probability and the expectation taken under the new probability. Then from (42) we yield

$$\begin{aligned}
&0.5\, E\,|M_D(T,t,\sigma_{D,s\wedge\tau_\ell,t,y}) - M_V(T,t,\sigma_{V,s\wedge\tau_\ell,t,y})| = \hat{E}\{[1-\exp(U_T - Z_T/2)]^+\} \\
&\le \hat{E}\{[1-\exp(U_T - Z_T/2)]^+ 1(\tau_n = T)\} + \hat{P}(\tau_n < T) \\
&\le \hat{E}\{|U_T - Z_T/2| 1(\tau_n = T)\} + \hat{P}(\tau_n < T) \\
&\le \hat{E}(|U_{\tau_n}|) + \hat{E}(Z_{\tau_n})/2 + \hat{P}(\tau_n < T) \\
&\le \left[\hat{E}\left(Z_{\tau_n}\right)\right]^{1/2} + \hat{E}\left(Z_{\tau_n}\right)/2 + \hat{P}\left(\tau_n < T\right), \qquad (44)
\end{aligned}$$

where stopping time τ_n is defined by

$$\tau_n = T \wedge \inf\left\{ s \in [t,T] : \log \sigma^2_{D,s,t,y} > \Upsilon_1 + \Upsilon_2\,|\log y| + \Upsilon_3 \log^{1/2} n \right\}, \qquad (45)$$

the second inequality is from inequality $[1-e^u]_+ \le 1 - e^{-|u|} \le |u|$, and the last inequality is due to the fact that (43) implies that under $\hat{P}$, Z_s is the variation process of U_s, and $\hat{E}\left(|U_{\tau_n}|^2\right) = \hat{E}\left(Z_{\tau_n}\right)$.

From (45) we obtain

$$
\begin{aligned}
\hat{P}\left(\tau_n < T\right) &= \hat{P}\left(\sup_{t\le s\le T}\log\sigma^2_{D,s,t,y} > \Upsilon_1 + \Upsilon_2\left|\log y\right| + \Upsilon_3\log^{1/2} n\right)\\
&\le \hat{P}\left(\sup_{t\le s\le T}\log\hat{\sigma}^2_{D,s,t,y} > \Upsilon_1 + \Upsilon_2\left|\log y\right| + \Upsilon_3\log^{1/2} n\right)\\
&\le \Upsilon n^{-1}\log^{-1/2} n. \qquad (46)
\end{aligned}
$$

where the first inequality is from (33), and the second inequality is due to Proposition A3.3 and identical distribution of $\hat{\sigma}_D$ under $\hat{P}$ and σ_D under P.

On the other hand, from (43) we get

$$
\begin{aligned}
\hat{E}\left(Z_{\tau_n}\right) &\le \hat{E}\int_t^{\tau_n\wedge\tau_\ell}(\sigma_{V,s,t,y}-\sigma_{D,s,t,y})^2 ds + \hat{E}\left\{1\{\tau_\ell<\tau_n\}(\tau_n-\tau_n\wedge\tau_\ell)(\sigma_{D,\tau_\ell,t,y}-\sigma_{V,\tau_\ell,t,y})^2\right\}\\
&\le \hat{E}\int_t^{\tau_n}(\sigma_{V,s,t,y}-\sigma_{D,s,t,y})^2 ds + \hat{E}\left\{1\{\tau_\ell<T\}\sup_{t\le s\le T}(\sigma_{D,s,t,y}-\sigma_{V,s,t,y})^2\right\}. \qquad (47)
\end{aligned}
$$

What we need is to evaluate the expectations in (47). Like notation $\hat{\sigma}^2_{D,s,t,y}$ used in the proof of Lemma 4.1, we denote by $\hat{\sigma}^2_{V,s,t,y}$ the expression of $\sigma^2_{V,s,t,y}$ with $B_{V,s}-B_{V,t}$ replaced by $\hat{B}_{V,s}$. Then the probabilistic behavior of $\hat{\sigma}^2_{\Lambda,s,t,y}$ under $\hat{P}$ is the same as $\sigma^2_{\Lambda,s,t,y}$ under P, and from Proposition A3.1 we have

$$
\log\sigma^2_{V,s,t,y}-\log\sigma^2_{D,s,t,y} = \log\hat{\sigma}^2_{V,s,t,y}-\log\hat{\sigma}^2_{D,s,t,y}+H_{s,t}, \qquad (48)
$$

where

$$
H_{s,t} = \rho\beta_2\int_t^{[ns-1]/n}\left(\alpha_1^{[ns]-[nu]-1}-e^{\beta_1(s-u)}\right)\sigma_{D,u\wedge\tau_\ell,t,y}\,du-\rho\beta_2\int_{[ns-1]/n}^{s}e^{\beta_1(s-u)}\,\sigma_{D,u\wedge\tau_\ell,t,y}du.
$$

Lemma A3.1 together with (45) and continuity of σ_D imply

$$
\sup_{t\le s\le\tau_n}\left|H_{s,t}\right| \;\le\; \Upsilon n^{-1}\sup_{t\le s\le\tau_n}\sigma_{D,s,t,y}\le n^{-1}\exp\left(\Upsilon_1+\Upsilon_2\left|\log y\right|+\Upsilon_3\log^{1/2} n\right). \qquad (49)
$$

Now we evaluate the expectations in (47). First,

$$
\hat{E}\int_t^{\tau_n}(\sigma_{V,s,t,y}-\sigma_{D,s,t,y})^2 ds
$$

$$
\begin{aligned}
&= \hat{E}\int_t^{\tau_n} \sigma^2_{D,s\wedge\tau_\ell,t,y}\left(1+e^{H_{s,t}}\hat{\sigma}^2_{V,s,t,y}/\hat{\sigma}^2_{D,s,t,y}-2e^{H_{s,t}/2}\,\hat{\sigma}_{V,s,t,y}/\hat{\sigma}_{D,s,t,y}\right)ds\\
&\le \hat{E}\int_t^{\tau_n} \hat{\sigma}^2_{D,s\wedge\tau_\ell,t,y}\left(1+e^{a/n}\hat{\sigma}^2_{V,s,t,y}/\hat{\sigma}^2_{D,s,t,y}-2e^{-0.5\,a/n}\hat{\sigma}_{V,s,t,y}/\hat{\sigma}_{D,s,t,y}\right)ds\\
&\le \int_t^{T}\hat{E}\left(\hat{\sigma}^2_{D,s,t,y}+e^{a/n}\hat{\sigma}^2_{V,s,t,y}-2e^{-0.5a/n}\hat{\sigma}_{V,s,t,y}\hat{\sigma}_{D,s,t,y}\right)ds\\
&\le n^{-1}\exp\left(\Upsilon_1+\Upsilon_2\,|\log y|+\Upsilon_3\log^{1/2}n\right), \qquad (50)
\end{aligned}
$$

where, of above five lines of equation array, the equality in the second line is from (48), the inequality in the third line is because of (33) and (49) with

$$
a=\exp\left(\Upsilon_1+\Upsilon_2\,|\log y|+\Upsilon_3\log^{1/2}n\right),
$$

and the inequality in the last line is due to the identical distribution of $(\hat{\sigma}_D,\hat{\sigma}_V)$ under $\hat{P}$ and (σ_D,σ_V) under P, and direct calculation by Proposition A3.2 that $(\log\sigma_{V,s,t,y},\log\sigma_{D,s,t,y})$ follows a bivariate normal distribution with marginal means and variances differing by an order of n^{-1} and correlation $1+O(n^{-1})$.

Second,

$$
\begin{aligned}
&\hat{E}\left\{1\{\tau_\ell<T\}\sup_{t\le s\le T}(\sigma_{D,s,t,y}-\sigma_{V,s,t,y})^2\right\}\le 2\left[\hat{P}(\tau_\ell<T)\hat{E}\left\{\sup_{t\le s\le T}(\sigma^4_{D,s,t,y}+\sigma^4_{V,s,t,y})\right\}\right]^{1/2}\\
&\le 2\left[\hat{P}(\tau_\ell<T)\hat{E}\left\{\sup_{t\le s\le T}(\hat{\sigma}^4_{D,s,t,y}+\hat{\sigma}^4_{V,s,t,y})\right\}\right]^{1/2}\\
&\le \exp\left(\Upsilon_1+\Upsilon_2\,|\log\,y|\right)\left[\hat{P}(\tau_\ell<T)\right]^{1/2}, \qquad (51)
\end{aligned}
$$

where the first inequality is from Cauchy-Schwartz inequality, the second inequality is from (33) and $\sigma^2_{V,s,t,y}\le\hat{\sigma}^2_{V,s,t,y}$ [which can be proved as easily as (33)], and the third inequality is from Proposition A3.3 and the identical distribution of $(\hat{\sigma}_D,\hat{\sigma}_V)$ under $\hat{P}$ and (σ_D,σ_V) under P.

Third, as $\rho\le 0$,

$$
\int_t^s e^{\beta_1(s-u)}dB_{V,u}=\int_t^s e^{\beta_1(s-u)}d\hat{B}_{V,u}+\rho\int_t^s e^{\beta_1(s-u)}\,\sigma_{D,u\wedge\tau_\ell,t,y}du\le\int_t^s e^{\beta_1(s-u)}d\hat{B}_{V,u},
$$

and hence from (31) we obtain

$$\begin{aligned}
\hat{P}(\tau_\ell < T) &= \hat{P}\left(\max_{t\le s\le T}\int_t^s e^{\beta_1(s-u)}\,dB_{V,u} > \ell\right)\\
&\le \hat{P}\left(\max_{t\le s\le T}\int_t^s e^{\beta_1(s-u)}d\hat{B}_{V,u} > \ell\right)\\
&\le \Upsilon_1 \ell^{-1}\exp(-\Upsilon_2\ell^2), \qquad (52)
\end{aligned}$$

where the probability in the second line of above equation array is calculated from the distribution of the maximum of Brownian motion $\hat{B}_V$ under $\hat{P}$.

Finally, plugging (50)-(52) into (47) and then combining it with (44) and (46), we arrive at

$$\begin{aligned}
&E\left|M_D(T,t,\sigma_{D,s\wedge\tau_\ell,t,y}) - M_V(T,t,\sigma_{V,s\wedge\tau_\ell,t,y})\right|\\
&\le n^{-1/2}\,exp\left(\Upsilon_1 + \Upsilon_2\,|log\,y| + \Upsilon_3\log^{1/2}n\right) + \exp\left(\Upsilon_4 + \Upsilon_5|\log y| - \Upsilon_6\ell^2\right)\ell^{-1/2}.
\end{aligned}$$

To complete the proof, we let $\ell \to \infty$ and apply Fatou's lemma.

4.6 Option pricing for the GARCH model

Theorem 4.2 in section 4.5 provides convergence speed for the SV option price to the diffusion option price, and Theorem 4.3 below derives the convergence speed for the GARCH option price to the SV option price. Thus, the main result, Theorem 2.1, is a consequence of Theorems 4.2 and 4.3.

Theorem 4.3

$$\mathcal{C}_G(t,x;T,K;y) \;\; = \;\; \mathcal{C}_V(t,x;T,K;y) + O(\zeta_n),$$

where $\zeta_n = n^{-1/2}\,\exp(\Upsilon\,\log^{1/2}\,n)$, *as specified in Theorem 4.2.*

Proof. The put-call parity implies that it is equivalent to establish order ζ_n for the difference between the prices of a European put option under the SV and GARCH models. Denote by

$\mathcal{P}_\Lambda(t, x; T, K; y)$ the price of a European call option with price x at current time t and strike price K at terminal time T.

From Duan (1995, corollary 2.3), we have that the European put option has the price

$$\begin{aligned}
\mathcal{P}_G(t, x; T, K; y) &= E\left[e^{-r\,(T-t)}\,(K - S_{G,T})_+ | (S_{G,t}, \sigma_{G,t+1/n}) = (x, y)\right] \\
&= e^{-r\,(T-t)}\,\Big\{K\,P\left(S_{G,T} \le K | (S_{G,t}, \sigma_{G,t+1/n}) = (x, y)\right) \\
&\qquad - E\left[S_{G,T} 1(S_{G,T} \le K) | (S_{G,t}, \sigma_{G,t+1/n}) = (x, y)\right]\Big\}. \qquad (53)
\end{aligned}$$

Note $t_* = [n\,t]/n$ and $T_* = [n\,T]/n$. With $\Lambda = G, V$, $(S_{\Lambda,s}, \sigma_{\Lambda,s})$ are step functions, $(S_{\Lambda,t}, \sigma_{\Lambda,t}) = (S_{\Lambda,t_*}, \sigma_{\Lambda,t_*})$, and $(S_{\Lambda,T}, \sigma_{\Lambda,T}) = (S_{\Lambda,T_*}, \sigma_{\Lambda,T_*})$, so we just need to prove the case of $t = t_*$ and $T = T_*$. From now on, we assume $t = t_*$ and $T = T_*$. From (15) and (16) we get

$$S_{\Lambda,T} = S_{\Lambda,t}\,\exp\left(r\,(T-t) + \int_t^T \sigma_{\Lambda,s} dW_{1,s}^{(n)} - \frac{1}{2}\int_t^T \sigma_{\Lambda,s}^2 ds\right). \qquad (54)$$

$$\int_t^T \sigma_{\Lambda,t}\,d\,W_{1,t} = \int_t^T \sigma_{\Lambda,t}\,d\,W_{1,t}^{(n)} = n^{-1/2} \sum_{i=n\,t+1}^{n\,T} \sigma_{\Lambda,t_i}\,\varepsilon_i, \quad \int_t^T \sigma_{\Lambda,s}^2 ds = n^{-1} \sum_{i=n\,t+1}^{n\,T} \sigma_{\Lambda,t_i}^2. \qquad (55)$$

Since $(S_{G,t}, \sigma_{G,t+1/n})$ are independent of $\varepsilon_{n\,t+1}, \cdots, \varepsilon_{n\,T}$, then the conditional probability and the conditional expectation in (53) can be evaluated by plugging the value (x, y) of $(S_{G,t}, \sigma_{G,t+1/n})$ into $\sigma_{G,s}$ and $S_{G,T}$, and then calculating the resulting probability and expectation with respect to $\varepsilon_{n\,t+1}, \cdots, \varepsilon_{n\,T}$. Given $S_{G,t} = x$ and $\sigma_{G,t+1/n} = y$, from (54)-(55) we can replace $\sigma_{G,s}$ by $\sigma_{G,s,t+1/n,y}$ and $S_{G,T}$ by $x\,e^{r\,(T-t)+A_G}$, and obtain from (53)

$$\begin{aligned}
\mathcal{P}_G(t, x; T, K; y) &= e^{-r\,(T-t)}\,K\,P\left\{A_G \le \log(K/x) - r\,(T-t)\right\} \\
&\quad - x\,E\left[e^{A_G}\,1\left\{A_G \le \log(K/x) - r\,(T-t)\right\}\right], \qquad (56)
\end{aligned}$$

where

$$A_G = A_G(T, t) = y\,n^{-1/2}\,\varepsilon_{[n\,t]+1} - \frac{y^2}{2\,n} + \int_{t+1/n}^T \sigma_{G,s,t+1/n,y}\,dW_{1,s}^{(n)} - \frac{1}{2}\int_{t+1/n}^T \sigma_{G,s,t+1/n,y}^2\,ds,$$

and $\sigma_{G,u,t,y}$ are defined in Proposition A3.1. For the SV model, from (34) and put-call parity, we conclude that the European put option has the price under the SV model

$$\begin{aligned}\mathcal{P}_V(t,x;T,K;y) = e^{-r\,(T-t)}\,K\,P\left\{A_V \le \log\left(K/x\right) - r\,(T-t)\right\}\\ - x\,E\left[e^{A_V}\,1\{A_V \le \log\left(K/x\right) - r\,(T-t)\}\right],\end{aligned}\tag{57}$$

where

$$A_V = A_V(T,t) = \int_t^T \sigma_{V,s,t,y}\,dW_{1,s}^{(n)} - \frac{1}{2}\int_t^T \sigma_{V,s,t,y}^2\,ds = \int_t^T \sigma_{V,s,t,y}\,dW_{1,s} - \frac{1}{2}\int_t^T \sigma_{V,s,t,y}^2\,ds.$$

From (17), $W_{1,s}$ and $B_{V,s}$ have correlation ρ. Conditional on $B_{V,s}, s \in [t,T]$, A_V follows a normal distribution with mean

$$\rho\int_t^T \sigma_{V,u,t,y}\,dB_{V,u} - \frac{1}{2}\int_t^T \sigma_{V,u,t,y}^2\,du,$$

and variance $(1-\rho^2)\int_t^T \sigma_{V,u,t,y}^2\,du$. Thus,

$$\begin{aligned}&P\left\{A_V \le \log\left(K/x\right) - r\,(T-t)\right\}\\ &= E\left[\Phi\left(\frac{\log\left(K/x\right) - r\,(T-t) - \rho\int_t^T \sigma_{V,s,t,y}\,dB_{V,s} + \int_t^T \sigma_{V,s,t,y}^2\,ds/2}{\sqrt{(1-\rho^2)\int_t^T \sigma_{V,s,t,y}^2\,ds}}\right)\right],\end{aligned}\tag{58}$$

and

$$\begin{aligned}&E\left[e^{A_V}\,1\{A_V \le \log\left(K/x\right) - r\,(T-t)\}\right]\\ &= E\int_{-\infty}^{\log(K/x) - r\,(T-t)} e^u\,\phi\left(\frac{u - \rho\int_t^T \sigma_{V,s,t,y}\,dB_{V,s} + \int_t^T \sigma_{V,s,t,y}^2\,ds/2}{\sqrt{(1-\rho^2)\int_t^T \sigma_{V,s,t,y}^2\,ds}}\right)du.\end{aligned}\tag{59}$$

Comparing (56) with (57), we see that the theorem will be proved if

$$P\left\{A_G \le \log\left(K/x\right) - r\,(T-t)\right\} = P\left\{A_V \le \log\left(K/x\right) - r\,(T-t)\right\} + O(\zeta_n),\tag{60}$$

$$E\left[e^{A_G}1\{A_G \le \log\left(K/x\right)\}\right] = E\left[e^{A_V}1\{A_V \le \log\left(K/x\right)\}\right] + O(\zeta_n).\tag{61}$$

Now we will show (60) and (61). By Lemma 4.4 below, we have

$$P\left\{A_G \le \log\left(K/x\right) - r\,(T-t)\right\} = P\left(A_G \le \log\left(K/x\right) - r\,(T-t), |A_G - A_V| \le \zeta_n\right) + O(n^{-1}).\tag{62}$$

Note that if $|A_G - A_V| \le \zeta_n$, the set $\{A_G \le \log(K/x) - r(T-t)\}$ is sandwiched between the two sets $\{A_V \le \log(K/x) - r(T-t) + \zeta_n\}$ and $\{A_V \le \log(K/x) - r(T-t) - \zeta_n\}$. Then the probability on the right hand side of (62) is between the two probabilities

$$P\left(A_V \le \log(K/x) - r(T-t) \pm \zeta_n, |A_G - A_V| \le \zeta_n\right).$$

To prove (60) we need to show that above two probabilities have ζ_n order difference with the probability on the right hand side of (60). Indeed,

$$\begin{aligned}
&P\left(A_V \le \log(K/x) - r(T-t) \pm \zeta_n, |A_G - A_V| \le \zeta_n\right) \\
&= P\left(A_V \le \log(K/x) - r(T-t) \pm \zeta_n\right) + O(n^{-1}) \\
&= E\Phi\left(\frac{\log(K/x) - r(T-t) - \rho\int_t^T \sigma_{V,s,t,y}\, dB_{V,s} + \int_t^T \sigma^2_{V,s,t,y}\, ds/2 \pm \zeta_n}{\sqrt{(1-\rho^2)\int_t^T \sigma^2_{V,s,t,y}\, ds}}\right) + O(n^{-1}) \\
&= E\Phi\left(\frac{\log(K/x) - r(T-t) - \rho\int_t^T \sigma_{V,s,t,y}\, dB_{V,s} + \int_t^T \sigma^2_{V,s,t,y}\, ds/2}{\sqrt{(1-\rho^2)\int_t^T \sigma^2_{V,s,t,y}\, ds}}\right) + O(\zeta_n)\, E\left(\sup_{0\le t\le s\le 1} \sigma^{-1}_{V,s,t,y}\right) \\
&= P\left\{A_V \le \log(K/x) - r(T-t)\right\} + O(\zeta_n),
\end{aligned}$$

where the first equality is from Lemma 4.4, the second equality is due to the conditional normality of A_V given B_V, the third equality is because of the mean value theorem and the boundness of the standard normal density function, and the last equality is from (58) and Proposition A3.3.

To prove (61), using Lemma 4.4 we obtain

$$\begin{aligned}
&x\, E\left[e^{A_G} 1\{A_G \le \log(K/x) - r(T-t), |A_G - A_V| > \zeta_n\}\right] \\
&\qquad \le KP(|A_G - A_V| > \zeta_n) = O(n^{-1}). \qquad (63)
\end{aligned}$$

Thus,

$$\begin{aligned}
&x\, E\left[e^{A_G} 1\{A_G \le \log(K/x) - r(T-t)\}\right] \\
&= x\, E\left[e^{A_G} 1\{A_G \le \log(K/x) - r(T-t), |A_G - A_V| \le \zeta_n\}\right] + O(n^{-1}).
\end{aligned}$$

The same argument proving (60) again shows that the expectation on the right hand side of above equation is between the following two expectations,

$$x\, E\left[e^{A_V \pm \zeta_n} 1\{A_V \le \log(K/x) - r\,(T-t) \pm \zeta_n, |A_G - A_V| \le \zeta_n\}\right],$$

and to establish (61) we need to prove that above two expectations differ by an order of ζ_n from the expectation on the right hand side of (61). In fact,

$$\begin{aligned}
& x\, E\left[e^{A_V \pm \zeta_n} 1\{A_V \le \log(K/x) - r\,(T-t) \pm \zeta_n, |A_G - A_V| \le \zeta_n\}\right] \\
& = x\, e^{\pm \zeta_n}\, E\left[e^{A_V} 1\{A_V \le \log(K/x) - r\,(T-t) \pm \zeta_n\}\right] + O(n^{-1}) \\
& = x\, e^{\pm \zeta_n}\, E \int_{-\infty}^{\log(K/x) - r\,(T-t) \pm \zeta_n} e^u\, \phi\left(\frac{u - \rho \int_t^T \sigma_{V,s,t,y}\, dB_{V,s} + \int_t^T \sigma^2_{V,s,t,y}\, ds/2}{\sqrt{(1-\rho^2) \int_t^T \sigma^2_{V,s,t,y}\, ds}}\right) du + O(n^{-1}) \\
& = x\, E \int_{-\infty}^{\log(K/x) - r\,(T-t)} e^u\, \phi\left(\frac{u - \rho \int_t^T \sigma_{V,s,t,y}\, dB_{V,s} + \int_t^T \sigma^2_{V,s,t,y}\, ds/2}{\sqrt{(1-\rho^2) \int_t^T \sigma^2_{V,s,t,y}\, ds}}\right) du + O(\zeta_n) \\
& = x\, E\left[e^{A_V}\, 1\{A_V \le \log(K/x) - r\,(T-t)\}\right] + O(\zeta_n),
\end{aligned}$$

where the first equality is using the fact that similar to (63), we have

$$\begin{aligned}
& x\, E\left[e^{A_V \pm \zeta_n} 1\{A_V < \log(K/x) - r\,(T-t) \pm \zeta_n, |A_G - A_V| > \zeta_n\}\right] \\
& \le e^{2\,\zeta_n}\, K\, P(|A_G - A_V| > \zeta_n) = O(n^{-1}),
\end{aligned}$$

the second equality is due to the conditional normality of A_V given B_V, the third equality is because of the mean value theorem, Proposition A3.3 and the fact that the standard normal density function is bounded, and the last equality is from (59).

Lemma 4.4 *Suppose $t = [n\,t]/n$ and $T = [n\,T]/n$. Then*

$$P\left(\max_{0 \le t \le T \le 1} |A_G(T,t) - A_V(T,t)| > \zeta_n\right) = O(n^{-1}).$$

Proof. Note that

$$\begin{aligned}
A_G(T,t) - A_V(T,t) \;\; &= \;\; (y - \sigma_{V,t+1/n,t,y})\, n^{-1/2}\, \varepsilon_{[n\,t]+1} + \left(\sigma^2_{V,t+1/n,t,y} - y^2\right)/(2\,n) \\
& \quad + n^{-1/2}\, U_T + n^{-1}\, Z_T/2, \qquad (64)
\end{aligned}$$

where

$$U_s = \int_{t+1/n}^{s} n^{1/2} \left(\sigma_{G,u,t+1/n,y} - \sigma_{V,u,t,y}\right) dW_{1,u}^{(n)}, \qquad Z_s = \int_{t+1/n}^{s} n \left(\sigma_{V,u,t,y}^2 - \sigma_{G,u,t+1/n,y}^2\right) du.$$

For the first term in (64), we have

$$\begin{aligned}
&P\left(\sup_{0\le t\le 1}\left|(y-\sigma_{V,t+1/n,t,y})\, n^{-1/2}\, \varepsilon_{[n\,t]+1}\right| > \sqrt{2}\, n^{-1/2} \log n\right) \\
&\le P\left(\sup_{0\le t\le 1}\left|(y-\sigma_{V,t+1/n,t,y})\, \varepsilon_{[n\,t]+1}\right| > \sqrt{2} \log n,\ \sup_{0\le t\le 1}|\varepsilon_{[n\,t]}| \le \sqrt{2} \log^{1/2} n\right) + O(n^{-1} \log^{-1/2} n) \\
&\le P\left(\sup_{0\le t\le 1}\left|y-\sigma_{V,t+1/n,t,y}\right| > \log^{1/2} n\right) + o(n^{-1}) \\
&\le P\left(\sup_{0\le t\le 1}\sigma_{V,t+1/n,t,y} > \log^{1/2} n\right) + o(n^{-1}) \\
&\le P\left(\sup_{0\le t\le 1}(B_{V,t}-B_{V,t-1/n}) > \beta_2^{-1}\{\log\log n - (\beta_0+\beta_1 \log y^2)/n\}\right) + o(n^{-1}) \\
&\le P\left(\sup_{0\le s\le 1} B_{V,s} > n^{1/2}\, \beta_2^{-1}\{\log\log n - (\beta_0+\beta_1 \log y^2)/n\}\right) + o(n^{-1}) = o(n^{-1}),
\end{aligned} \tag{65}$$

where the first inequality is because of

$$P\left(\sup_{0\le t\le 1}|\varepsilon_{[n\,t]}| > \sqrt{2} \log^{1/2} n\right) \sim n^{-1} \log^{-1/2} n,$$

the fourth inequality is due to

$$\sigma_{V,t+1/n,t,y} = y \exp\left\{(\beta_1 \log y + \beta_0/2)/n + \beta_2\,(B_{V,t}-B_{V,t-1/n})/2\right\},$$

which is implied by Proposition A3.1, the fifth inequality is from the rescaling property of Brownian motion B_V, and the final order is derived from the distribution of the maximum of Brownian motion B_V. Similarly, for the second term in (64) we get

$$\begin{aligned}
&P\left(\sup_{0\le t\le 1}\left|\sigma_{V,t+1/n,t,y}^2 - y^2\right|/(2\,n) > n^{-1/2} \log n\right) \le P\left(\sup_{0\le t\le 1}\sigma_{V,t+1/n,t,y}^2 > 2\,n^{1/2} \log n\right) + o(n^{-1}) \\
&\le P\left(\sup_{0\le s\le 1} B_{V,s} > n^{1/2}\, \beta_2^{-1}\{0.5 \log n - (\beta_0+\beta_1 \log y^2)/n\}\right) + o(n^{-1}) = o(n^{-1}).
\end{aligned} \tag{66}$$

From (64)-(66) we conclude

$$
\begin{aligned}
&P\left(\sup_{t\le T\le 1}|A_G(T,t)-A_V(T,t)|>3\,n^{-1/2}\,\exp(d\,\log^{1/2}\,n)\right)\\
&\le P\left(\sup_{t+1/n\le T\le 1}|U_T+n^{-1/2}\,Z_T/2|>2\,\exp(d\,\log^{1/2}\,n)\right)+o(n^{-1})\\
&\le P\left(\sup_{t+1/n\le T\le 1}(U_T+n^{-1/2}\,Z_T/2)>2\,\exp(d\,\log^{1/2}\,n)\right)\\
&\quad+P\left(\sup_{t+1/n\le T\le 1}(-U_T-n^{-1/2}\,Z_T/2)>2\,\exp(d\,\log^{1/2}\,n)\right)+o(n^{-1})\\
&\le P\left(\sup_{t+1/n\le T\le 1}(U_T-[U,U]_T/2)>\exp(d\,\log^{1/2}\,n)\right)\\
&\quad+P\left(\sup_{t+1/n\le T\le 1}(-U_T-[U,U]_T/2)>\exp(d\,\log^{1/2}\,n)\right)+2\,P\left([U,U]_1>\exp(d\,\log^{1/2}\,n)\right)\\
&\quad+2\,P\left(\sup_{t+1/n\le s\le 1}|n^{-1/2}\,Z_T|>\exp(d\,\log^{1/2}\,n)\right)+o(n^{-1}),
\end{aligned}
\tag{67}
$$

where d is a positive constant that will be chosen later on, and $[U,U]$ is the variation process of U_s given by

$$
[U,U]_s=\int_{t+1/n}^{s} n\,(\sigma_{G,u,t+1/n,y}-\sigma_{V,u,t,y})^2\,du.
$$

By Proposition A3.6, we obtain that for $\Upsilon_3\,\log^{1/2}\,n>\Upsilon_1+\Upsilon_2\,|\log\,y|$,

$$
\begin{aligned}
&P\left(\sup_{t+1/n\le T\le 1}|n^{-1/2}\,Z_T|>\exp(2\,\Upsilon_3\,\log^{1/2}\,n)\right)\\
&\le P\left(\sup_{t+1/n\le s\le 1}|\sigma^2_{G,s,t+1/n,y}-\sigma^2_{V,s,t,y}|>n^{-1/2}\,\exp(2\,\Upsilon_3\,\log^{1/2}\,n)\right)<\Upsilon\,n^{-1},
\end{aligned}
\tag{68}
$$

$$
\begin{aligned}
&P\left([U,U]_1>\exp(4\,\Upsilon_3\,\log^{1/2}\,n)\right)\\
&\le P\left(\sup_{t+1/n\le s\le 1}|\sigma_{G,s,t+1/n,y}-\sigma_{V,s,t,y}|>n^{-1/2}\,\exp(2\,\Upsilon_3\,\log^{1/2}\,n)\right)<\Upsilon\,n^{-1}.
\end{aligned}
\tag{69}
$$

For $t_k\ge t+1/n$, $\sigma^2_{G,t_k,t+1/n,y}$ depends only on $\varepsilon_{n\,t+2},\cdots,\varepsilon_{k-1}$, and $\sigma^2_{V,t_k,t,y}$ relies on $\varepsilon_{n\,t+1},\cdots,\varepsilon_{k-1}$ and $\delta_{n\,t+1},\cdots,\delta_{k-1}$, thus ε_k is independent of $\sigma^2_{G,t_k,t+1/n,y}$ and $\sigma^2_{V,t_k,t,y}$. Using conditional argument and the normality of ε_k, we calculate

$$
E\left[\exp\left\{\pm(U_{t_k}-U_{t_{k-1}})-([U,U]_{t_k}-[U,U]_{t_{k-1}})/2\right\}\Big|\,\varepsilon_{n\,t+1},\delta_{n\,t+1},\cdots,\varepsilon_{k-1},\delta_{k-1}\right]
$$

$$= E\left[\exp\left\{\pm(\sigma_{G,t_k,t+1/n,y} - \sigma_{V,t_k,t,y})\,\varepsilon_k - (\sigma_{G,t_k,t+1/n,y} - \sigma_{V,t_k,t,y})^2/2\right\}\middle|\,\varepsilon_{n\,t+1}, \delta_{n\,t+1}, \cdots, \varepsilon_{k-1}, \delta_{k-1}\right]$$
$$= 1.$$

This leads to that

$$\exp\left(\pm U_{t_k} - [U,U]_{t_k}/2\right), t + 1/n \le t_k \le 1$$

is a martingale with mean one. Hence,

$$\begin{aligned}
&P\left(\sup_{t+1/n \le T \le 1} (\pm U_T - [U,U]_T) > exp(\Upsilon_3\, log^{1/2}\, n)\right) \\
&\le P\left(\sup_{t+1/n \le t_k \le 1} \exp\left\{\pm U_{t_k} - [U,U]_{t_k}/2\right\} > \exp\left\{\exp(\Upsilon_3\, \log^{1/2}\, n)\right\}\right) \\
&\le \exp\left\{-\exp(\Upsilon_3\, \log^{1/2}\, n)\right\} = o(n^{-1}), \qquad (70)
\end{aligned}$$

where the first inequality is because of piecewise constant of U_s on $[t_k, t_{k+1})$, and the second inequality is due to an application of the martingale maxima inequality to $\exp(\pm U_{t_k} - [U,U]_{t_k}/2)$. Finally, we prove the lemma by plugging (68)-(70) into (67) and taking $d = 4\,\Upsilon_3$.

Appendix: Probability inequalities for price and volatility processes

To make the proofs self-complete, in this section we provide detailed analysis of price and volatility processes for the three models under the risk neutral probability.

A1 Embed normal noises in Brownian motions

Since ε_j and δ_j in the GARCH and SV models are i.i.d. standard normal random variables, and they correspond to Brownian motions W_1 and W_2, respectively, in the diffusion model. Thus, we realize $(\varepsilon_j, \delta_j)$ and (W_1, W_2) on common probability spaces and take $n^{1/2}\,\varepsilon_j =$

$W_{1,t_j} - W_{1,t_{j-1}}$, and $n^{1/2}\,\delta_j = W_{2,t_j} - W_{2,t_{j-1}}$, and then

$$W^{(n)}_{1,t_k} = W_{1,t_k}, \qquad W^{(n)}_{2,t_k} = W_{2,t_k}, \qquad B^{(n)}_{V,t_k} = B_{V,t_k}, \qquad k = 0, \cdots, n.$$

Lemma A1.1 *On some common probability spaces we can realize that for $i = 1, 2$, $W^{(n)}_{i,t_j} = W_{i,t_j}$, and $B^{(n)}_{V,t_j} = B_{V,t_j}$, $j = 0, \cdots, n$,*

$$\sup_{0 \le s \le 1} \left| B^{(n)}_{V,s} - B_{V,s} \right| = \max_{1 \le j \le n} \sup_{t_j \le s \le t_{j+1}} \left| B_{V,s} - B_{V,t_j} \right|,$$

$$P\left(\max_{1 \le j \le n} \sup_{t_j \le s \le t_{j+1}} \left| B_{V,s} - B_{V,t_j} \right| > 2\, n^{-1/2}\, \log^{1/2}\, n \right) \le n^{-2}\, \log^{-1/2}\, n, \qquad \text{(A1.1)}$$

where B_V, $W^{(n)}_i$ and $B^{(n)}_V$ are defined in (17), (19) and (20).

Proof. The embedding of $(W^{(n)}_1, W^{(n)}_2, B^{(n)}_V)$ in (W_1, W_2, B_V) is given in Section 2, so we just need to prove (A1.1). Let

$$\eta_j = n^{1/2} \sup_{t_{j-1} \le s \le t_j} [B_{V,s} - B_{V,t_j}].$$

Because of the rescaling property and independent and stationary increments of Brownian motion, η_j are i.i.d. with distribution equal to the maximum of a Brownian motion over $[0, 1]$. The maximum of a standard Brownian motion over $[0, a]$ has the distribution function

$$2\,\Phi(u/\sqrt{a}) - 1 = 2\,a^{-1/2} \int_0^u \phi(x/\sqrt{a})\,dx, \qquad u > 0. \qquad \text{(A1.2)}$$

Thus,

$$P\left(\max_{1 \le j \le n} |\eta_j| > 2\, \log^{1/2}\, n \right) = 2 - 2\left[1 - 2 \int_{2\,\log^{1/2}\, n}^{\infty} \phi(x)\,dx \right]^n \le n^{-2}\, log^{-1/2}\, n.$$

This proves (A1.1).

A2 Strong approximation of partial sum processes for the GARCH and SV noises

Lemma A2.1 *On some common probability spaces we can construct ε_j, δ_i and ξ_j to preserve all the relationship among noises in price and volatility processes of the GARCH and SV models, and to satisfy*

$$P\left(\sup_{0\le s\le 1}\left|B_{V,s}^{(n)}-B_{G,s}^{(n)}\right|>\Upsilon_1\, n^{-1/2}\,\log\, n\right)<\Upsilon\, n^{-1},$$

where $B_V^{(n)}$ and $B_G^{(n)}$ are defined in (20) and (21), ε_j are standard normal random noise for the price processes in the GARCH and SV models, and δ_j and ξ_j [defined in (22)] are parts of random noise sources for the volatility processes in the GARCH and SV models, respectively.

Proof. Denote the partial sum process

$$U_s^{(n)}=(1-2/\pi)^{-1/2}\,n^{-1/2}\sum_{j=1}^{[n\,s]}\left\{|\varepsilon_j|-(2/\pi)^{1/2}\right\},$$

$$Z_s^{(n)}=(1-2/\pi)^{-1/2}\,n^{-1/2}\sum_{j=1}^{[n\,s]}\left(|\varepsilon_j-\lambda\, n^{-1/2}|-|\varepsilon_j|\right).$$

Then from (19)-(21) we get

$$\left|B_{V,s}^{(n)}-B_{G,s}^{(n)}\right|\le\left|W_{2,s}^{(n)}-U_s^{(n)}\right|+\left|Z_s^{(n)}\right|. \tag{A2.3}$$

Applying KMT strong approximation (Komlós, Major and Tusnády 1975, 1976) to $U_s^{(n)}$ and $W_{2,s}^{(n)}$, we have that on some probability spaces

$$P\left(\sup_{0\le s\le 1}\left|W_{2,s}^{(n)}-U_s^{(n)}\right|>\Upsilon_1\, n^{-1/2}\,\log\, n+n^{-1/2}\,x\right)<\Upsilon_2\, e^{-x}. \tag{A2.4}$$

As $|\varepsilon_j-\lambda\, n^{-1/2}|-|\varepsilon_j|$ are i.i.d. random variables, simple calculations show

$$a_n=(1-2/\pi)^{-1/2}\,E\left(|\varepsilon_j-\lambda\, n^{-1/2}|-|\varepsilon_j|\right)=(1-2/\pi)^{-1/2}\,(2\,\pi)^{-1/2}\,\lambda^2\, n^{-1}+O(n^{-2}),$$

$$b_n^2 = (1-2/\pi)^{-1}\, Var\left(|\varepsilon_j - \lambda\, n^{-1/2}| - |\varepsilon_j|\right) = (1-2/\pi)^{-1}\, \lambda^2\, n^{-1} + O(n^{-3/2}).$$

The median of a random variable is always within one standard deviation from its mean, then

$$\begin{aligned}
&\left| median\left(Z_1^{(n)} - Z_s^{(n)}\right)\right| \le \left| E\left(Z_1^{(n)} - Z_s^{(n)}\right)\right| + \left\{ Var\left(Z_1^{(n)} - Z_s^{(n)}\right)\right\}^{1/2} \\
&\le n^{1/2}\,|a_n| + b_n = (1-2/\pi)^{-1/2}\,(\lambda + \lambda^2\,(2/\pi)^{-1/2})\, n^{-1/2} + O(n^{-1}),
\end{aligned}$$

and

$$\begin{aligned}
&P\left(\sup_{0\le s\le 1} \left|Z_s^{(n)}\right| \ge \Upsilon_1\, n^{-1/2}\, \log\, n\right) \\
&\le P\left(\sup_{0\le s\le 1} \left\{\left|Z_s^{(n)} - median\left(Z_1^{(n)} - Z_s^{(n)}\right)\right|\right\} \ge \Upsilon_1\, n^{-1/2}\, \log\, n - n^{1/2}\, a_n - b_n\right) \\
&\le 2\, P\left(\left|Z_1^{(n)}\right| \ge \Upsilon_1\, n^{-1/2}\, \log\, n\right) \\
&\le 2\, P\left(b_n^{-1}\, \left|Z_1^{(n)} - n^{1/2}\, a_n\right| \ge \Upsilon_1\, \log\, n\right) \\
&\le 4\, [1 - \Phi(\Upsilon_1\, \log\, n)] + \Upsilon_1\, \log^2\, n\, \phi(\Upsilon_1\, \log\, n)\, n^{-1/2} \\
&\le \Upsilon\, n^{-1},
\end{aligned}$$

where the second inequality is from the Lévy inequality, and the fourth inequality is due to an application of Edgeworth expansion to $b_n^{-1}\,(Z_1^{(n)} - n^{1/2}\, a_n)$. Now the inequality in the lemma is proved by taking $x = \log\, n$ in (A2.4) and then combining it with (A2.3) and above inequality.

Finally, we show that ε_j, δ_j and ξ_j can be constructed to retain all the relationship specified by the GARCH and SV models: that is, ε_j is independent of δ_i and uncorrelated with $|\varepsilon_j|$. Indeed, first, the standard KMT construction creates $|\varepsilon_j|$ from δ_j to maintain (A2.4), and second from $|\varepsilon_j|$ we construct normal random variables ε_j by multiplying it independently and equal likely with -1 and 1. Since the multiplied sign random variables are independent of δ_j as well as the constructed $|\varepsilon_j|$, ε_j is uncorrelated with $|\varepsilon_j|$ and δ_j. Because ε_j and δ_j are normally distributed, zero correlation means independence.

A3 Analysis of volatility processes

For $s \geq t$, denote by $\sigma^2_{\Lambda,s,t,y}$ the expression of $\sigma^2_{\Lambda,s}$ when $\sigma^2_{\Lambda,t} = y^2$.

Proposition A3.1

$$log\, \sigma^2_{D,s,t,y} \quad = \quad e^{\beta_1\,(s-t)}\, log\, y^2 + \beta_0 \int_t^s e^{\beta_1\,(s-u)}\, du + \beta_2 \int_t^s e^{\beta_1\,(s-u)}\, dB_{2,u},$$

$$\begin{aligned} log\, \sigma^2_{V,s,t,y} &= \alpha_1^{[n\,s]-[n\,t]}\, log\, y^2 + \beta_0\, n^{-1} \sum_{j=[n\,t]+1}^{[n\,s]} \alpha_1^{[n\,s]-j} + \beta_2 \sum_{j=[n\,t]+1}^{[n\,s]} \alpha_1^{[n\,s]-j} \left(B_{V,t_{j-1}} - B_{V,t_{j-2}}\right) \\ &= \alpha_1^{[n\,s]-[n\,t]}\, log\, y^2 + \beta_0 \int_{[n\,t]/n}^{[n\,s]/n} \alpha_1^{[n\,s]-[n\,u]}\, du + \beta_2 \int_{[n\,t-1]/n}^{[n\,s-1]/n} \alpha_1^{[n\,s]-[n\,u]-1}\, dB_{V,u}, \end{aligned}$$

$$log\, \sigma^2_{G,s,t,y} = \alpha_1^{[n\,s]-[n\,t]}\, log\, y^2 + \beta_0\, n^{-1} \sum_{j=[n\,t]+1}^{[n\,s]} \alpha_1^{[n\,s]-j} + \beta_2 \sum_{j=[n\,t]+1}^{[n\,s]} \alpha_1^{[n\,s]-j} \left(B^{(n)}_{G,t_{j-1}} - B^{(n)}_{G,t_{j-2}}\right),$$

where $\alpha_1 = 1 + n^{-1}\,\beta_1$.

Proof. The proposition is easily derived from equations (14) and (16), and embedding of $B_V^{(n)}$ in B_V given by Lemma A1.1.

Lemma A3.1

$$\max_{0\leq t\leq s\leq 1} \left|(1+\beta_1/n)^{[n\,s]-[n\,t]} - e^{\beta_1\,(s-t)}\right| \leq \Upsilon/n.$$

Proof. It can be verified by direction calculation.

Proposition A3.2 *$log\, \sigma^2_{D,s,t,y}$ and $log\, \sigma^2_{V,s,t,y}$ are normally distributed with*

$$E(log\, \sigma^2_{D,s,t,y}) = e^{\beta_1\,(s-t)}\, log\, y^2 + \beta_0\,(e^{\beta_1\,(s-t)}-1)/\beta_1, \quad Var(log\, \sigma^2_{D,s,t,y}) = \beta_2^2\,(e^{2\,\beta_1\,(s-t)}-1)/(2\,\beta_1),$$

$$E(log\, \sigma^2_{V,s,t,y}) = E(log\, \sigma^2_{D,s,t,y}) + O(n^{-1}\, log\, y), \qquad Var(log\, \sigma^2_{V,s,t,y}) = Var(log\, \sigma^2_{D,s,t,y}) + O(n^{-1}),$$

$$Correlation(log\, \sigma^2_{D,s,t,y}, log\, \sigma^2_{V,s,t,y}) = 1 + O(n^{-1}).$$

In particular, $log\, \sigma^2_{V,s,t,y} - log\, \sigma^2_{D,s,t,y}$ follows a normal distribution with mean of order $n^{-1}\, log\, y$ and variance of order n^{-1}.

Proof. The proposition is a simple consequence of Lemma A3.1 and the expressions of $log\, \sigma^2_{\Lambda,s,t,y}$ in Proposition A3.1.

Proposition A3.3 *For the diffusion and SV volatility processes (i.e. $\Lambda = D, V$),*

$$P\left(\sup_{0\le t\le s\le 1} |log\, \sigma^2_{\Lambda,s,t,y}| > \Upsilon_1 + \Upsilon_2\, |log\, y| + \Upsilon_3\, log^{1/2}\, n\right) < \Upsilon\, n^{-1}\, log^{-1/2}\, n,$$

and for real number a,

$$E\left[\sup_{0\le t\le T\le 1}\left(\frac{1}{T-t}\int_t^T \sigma^2_{\Lambda,s,t,y}\, ds\right)^a\right] \le E\left[\sup_{0\le t\le s\le 1}(\sigma^2_{\Lambda,s,t,y})^a\right] \le exp\left\{|a|\, \Upsilon_1 + |a|\, \Upsilon_2\, |log\, y| + a^2\, \Upsilon_3\right\}.$$

Proof. First we show for $\Lambda = D, V$,

$$\sup_{0\le t\le s\le 1} |log\, \sigma^2_{\Lambda,s,t,y}| \le |\beta_0/\beta_1|\, (e^{|\beta_1|} - 1) + 2\, e^{|\beta_1|}\, |log\, y| + \beta_2\, e^{|\beta_1|} \sup_{0\le t\le s\le 1} |B_{V,s} - B_{V,t}|. \quad \text{(A3.5)}$$

As the diffusion is the limit of the SV model, we need to prove (A3.5) for the SV case only. Applying Abel transformation we have for $s \ge t$,

$$\begin{aligned}
&\left|\sum_{j=[n\,t]+1}^{[n\,s]} \alpha_1^{[n\,s]-j}\, \left(B_{V,t_{j-1}} - B_{V,t_{j-2}}\right)\right| \\
&= \left|\sum_{j=[n\,t]+1}^{[n\,s]-1} \left(\alpha_1^{[n\,s]-j} - \alpha_1^{[n\,s]-j-1}\right)\, \left(B_{V,t_j-1/n} - B_{V,t-1/n}\right) + \left(B_{V,s-1/n} - B_{V,t-1/n}\right)\right| \\
&\le e^{|\beta_1|\,(s-t)} \sup_{0\le t\le u\le 1} |B_{V,u} - B_{V,t}|.
\end{aligned}$$

Using above inequality and the expression of $log\, \sigma^2_{V,s,t,y}$ in Proposition A3.1 we obtain that for $s \ge t$,

$$|log\, \sigma^2_{V,s,t,y}| \le 2\, e^{\beta_1\,(s-t)}\, |log\, y| + \beta_2\, e^{|\beta_1|\,(s-t)} \sup_{0\le t\le u\le 1} |B_{V,u} - B_{V,t}| + |\beta_0|\, \left[e^{\beta_1\,(s-t)} - 1\right]/\beta_1.$$

This proves (A3.5) for the SV model. The two inequalities in the proposition are now easily proved by combining the established inequality (A3.5) with the following two inequalities for the maximum of $|B_{V,s} - B_{V,t}|$,

$$P\left\{\max_{0\le t\le s\le 1} |B_{V,s} - B_{V,t}| > (2\, log\, n)^{1/2}\right\} \le 2\int_{[2\, log\, n]^{1/2}}^{\infty} \phi(u)\, du \sim n^{-1}\, log^{-1/2}\, n,$$

$$E exp\left\{|\eta| \max_{0\le t\le s\le 1} |B_{V,s} - B_{V,t}|\right\} \le 2\, e^{\eta^2/2},$$

which are a consequence of independent and stationary increment property of $B_{V,s}$ and its maximum distribution given by (A1.2).

Proposition A3.4

$$P\left(\sup_{0\le t\le s\le 1} \left|log\, \sigma^2_{D,s,t,y} - log\, \sigma^2_{V,s,t,y}\right| > \Upsilon_1\, n^{-1/2}\, log^{1/2}\, n\right) < \Upsilon\, n^{-1}\, log^{-1/2}\, n.$$

Proof. From Proposition A3.1 we have

$$log\, \sigma^2_{V,s,t,y} - log\, \sigma^2_{D,s,t,y} = A_{1,s,t} + A_{2,s,t},$$

where

$$A_{1,s,t} = \left(\alpha_1^{[n\,s]-[n\,t]} - e^{\beta_1\,(s-t)}\right)\, log\, y^2 + (\beta_0/\beta_1)\, \left(\alpha_1^{[n\,s]-[n\,t]} - e^{\beta_1\,(s-t)}\right),$$

$$\begin{aligned} A_{2,s,t} &= \beta_2 \int_t^{[n\,s-1]/n} \left(\alpha_1^{[n\,s]-[n\,u]-1} - e^{\beta_1\,(s-u)}\right)\, dB_{V,u} + \beta_2 \int_{[n\,t-1]/n}^{t} \alpha_1^{[n\,s]-[n\,u]-1}\, dB_{V,u} \\ &\quad - \beta_2 \int_{[n\,s-1]/n}^{s} e^{\beta_1\,(s-u)}\, dB_{V,u}. \end{aligned}$$

Lemma A3.1 indicates that $A_{1,s,t}$ is of order $n^{-1}\,|log\, y|$. $A_{2,s,t}$ is bounded by

$$\Upsilon_1\, n^{-1} \max_{t\le u\le v\le s-1/n} |B_{V,v}-B_{V,u}| + \beta_2\, e^{|\beta_1|/n} \max_{t-2/n\le u\le v\le t} |B_{V,v}-B_{V,u}| + \beta_2\, e^{|\beta_1|/n} \max_{s-2/n\le u\le v\le s} |B_{V,v}-B_{V,u}|,$$

and the maximum distribution of $B_{V,s}$ shows that each of above three terms is bounded by $\Upsilon_1\, n^{-1/2}\, log^{1/2}\, n$ with probability exceeding $1 - \Upsilon\, n^{-1}\, log^{-1/2}\, n$. This completes the proof.

Proposition A3.5

$$P\left(\sup_{0\le t\le s\le 1} \left|log\, \sigma^2_{G,s,t,y} - log\, \sigma^2_{V,s,t,y}\right| > \Upsilon_1\, n^{-1/2} log\, n\right) < \Upsilon\, n^{-1}.$$

Proof. Using the expressions of $log\, \sigma^2_{G,s,t,y}$ and $log\, \sigma^2_{V,s,t,y}$ given by Proposition A3.1, and applying Abel transformation, we obtain

$$|log\, \sigma^2_{V,s,t,y} - log\, \sigma^2_{G,s,t,y}| = \beta_2 \left| \sum_{j=[n\,t]+1}^{[n\,s]} \alpha_1^{[n\,s]-j} \left(B^{(n)}_{V,t_{j-1}} - B^{(n)}_{G,t_{j-1}} - B^{(n)}_{V,t_{j-2}} + B^{(n)}_{G,t_{j-2}}\right)\right|$$

$$\leq \beta_2 \sum_{j=[n\,t]+1}^{[n\,s]-1} \Big| \Big(\alpha_1^{[n\,s]-j} - \alpha_1^{[n\,s]-j-1}\Big) \Big(B^{(n)}_{V,t_j-1/n} - B^{(n)}_{G,t_j-1/n}\Big)$$

$$+ \Big(B^{(n)}_{V,s-1/n} - B^{(n)}_{G,s-1/n}\Big) - \alpha_1^{[n\,s]-[n\,t]-1} \Big(B^{(n)}_{V,t-1/n} - B^{(n)}_{G,t-1/n}\Big)\Big|$$

$$\leq \beta_2\, e^{|\beta_1|} \sup_{0\leq s\leq 1} |B^{(n)}_{V,s} - B^{(n)}_{G,s}|.$$

Now the proposition is an immediate consequence of Lemma A2.1.

Proposition A3.6 *With probability exceeding* $1 - \Upsilon\, n^{-1}$,

$$\sup_{0\leq t+1/n\leq s\leq 1} \Big|\sigma^2_{G,s,t+1/n,y} - \sigma^2_{V,s,t,y}\Big| \leq n^{-1/2}\, exp\Big(\Upsilon_1 + \Upsilon_2\, |log\, y| + \Upsilon_3\, log^{1/2}\, n\Big),$$

$$\sup_{0\leq t+1/n\leq s\leq 1} \Big|\sigma_{G,s,t+1/n,y} - \sigma_{V,s,t,y}\Big| \leq n^{-1/2}\, exp\Big(\Upsilon_1 + \Upsilon_2\, |log\, y| + \Upsilon_3\, log^{1/2}\, n\Big).$$

Proof. Note that

$$\Big|log\, \sigma^2_{G,s,t+1/n,y} - log\, \sigma^2_{V,s,t,y}\Big| \leq \Big|log\, \sigma^2_{G,s,t+1/n,y} - log\, \sigma^2_{V,s,t+1/n,y}\Big|$$

$$+ \Big|log\, \sigma^2_{V,s,t+1/n,y} - log\, \sigma^2_{D,s,t+1/n,y}\Big| + \Big|log\, \sigma^2_{D,s,t,y} - log\, \sigma^2_{V,s,t,y}\Big|$$

$$+ \Big|log\, \sigma^2_{D,s,t+1/n,y} - log\, \sigma^2_{D,s,t,y}\Big|.$$

Propositions A3.4 and A3.5 imply that on the right hand side of above inequality, each of the first three terms is bounded by $\Upsilon_1\, n^{-1/2}$ log n with probability exceeding $1 - \Upsilon\, n^{-1}$. For the last term, by Proposition A3.1 we have

$$|log\, \sigma^2_{D,s,t+1/n,y} - log\, \sigma^2_{D,s,t,y}| \leq 2\, \Big|e^{-\beta_1/n} - 1\Big|\, e^{\beta_1\,(s-t)}\, |log\, y| + |\beta_0| \int_t^{t+1/n} e^{\beta_1\,(s-u)}\, du$$

$$+ \beta_2 \left| \int_t^{t+1/n} e^{\beta_1\,(s-u)}\, dB_{V,u} \right|,$$

where the first two deterministic terms are of order n^{-1}, and the third stochastic term is bounded by

$$\Upsilon_1 \max_{0\leq t\leq u\leq v\leq t+1/n} |B_{V,v} - B_{V,u}|,$$

which in turn is bounded by $\Upsilon_1 \, n^{-1/2} \log^{1/2} n$ with probability exceeding $1 - \Upsilon \, n^{-1}$. This proves that with probability exceeding $1 - \Upsilon \, n^{-1}$,

$$\sup_{0 \le t+1/n \le s \le 1} \left| log\, \sigma^2_{G,s,t+1/n,y} - log\, \sigma^2_{V,s,t,y} \right| \le \Upsilon_1 \, n^{-1/2} log\, n,$$

that is,

$$exp\{-\Upsilon_1 \, n^{-1/2} \, log\, n\} \le \sigma^2_{G,s,t+1/n,y} / \sigma^2_{V,s,t,y} \le exp\{\Upsilon_1 \, n^{-1/2} \, log\, n\}.$$

Since $n^{-1/2} \, log\, n < 1$, and

$$exp\{-\Upsilon_1 \, n^{-1/2} \, log\, n\} \ge 1 - \Upsilon_1 \, n^{-1/2} \, log\, n, \qquad exp\{\Upsilon_1 \, n^{-1/2} \, log\, n\} \le 1 + \Upsilon_1 \, e^{\Upsilon_1} \, n^{-1/2} \, log\, n,$$

we immediately obtain that with probability exceeding $1 - \Upsilon \, n^{-1}$,

$$\sup_{0 \le t+1/n \le s \le 1} |\sigma_{G,s,t+1/n,y} / \sigma_{V,s,t,y} - 1| \le \Upsilon_1 \, e^{\Upsilon_1} \, n^{-1/2} \, log\, n,$$

$$\sup_{0 \le t+1/n \le s \le 1} |\sigma^2_{G,s,t+1/n,y} / \sigma^2_{V,s,t,y} - 1| \le \Upsilon_1 \, e^{\Upsilon_1} \, n^{-1/2} \, log\, n.$$

Now the proposition is easily proved by combining above two inequalities with the first inequality in Proposition A3.3.

A4 Analysis of price processes

Proposition A4.1

$$\sup_{0 \le s \le 1} |S_{G,s} - S_{D,s}| = O_p(n^{-1/2} \, log\, n), \qquad \sup_{0 \le s \le 1} |S_{V,s} - S_{D,s}| = O_p(n^{-1/2} \, log^{1/2} \, n).$$

Proof. Note that

$$P\left(\sup_{0 \le s \le 1} |log\,(S_{G,s}/S_{V,s})| > 2\, d\, n^{-1/2} \, log\, n \right)$$
$$\le P\left(\sup_{0 \le s \le 1} \left| \sigma^2_{G,s,0,\sigma_0} - \sigma^2_{V,s,0,\sigma_0} \right| > 2\, d\, n^{-1/2} \, log\, n \right) + P\left(\sup_{0 \le s \le 1} |U_s| > d\, n^{-1/2} \, log\, n \right),$$

where

$$U_s = \int_0^s (\sigma_{G,s,0,\sigma_0} - \sigma_{V,s,0,\sigma_0})\, dW_{1,s}.$$

U_s has variation process

$$[U, U]_s = \int_0^s (\sigma_{G,s,0,\sigma_0} - \sigma_{V,s,0,\sigma_0})^2\, du.$$

By Proposition A3.5, we obtain

$$\sup_{0\le s\le 1} \left|\sigma^2_{G,s,0,\sigma_0} - \sigma^2_{V,s,0,\sigma_0}\right| = O_p(n^{-1/2}\, log\, n), \qquad [U, U]_1 = O_p(n^{-1}\, log^2\, n).$$

By the Lenglart inequality we have

$$P(\sup_{0\le s\le 1} |U_s| > d\, n^{-1/2}\, log\, n) \le \frac{1}{d} + P([U, U]_1 > d\, n^{-1}\, log^2\, n) \to 0,$$

as $n \to \infty$ and then $d \to \infty$.

The second result can be proved by the same arguments with Proposition A3.4.

REFERENCES

Aït-Sahalia, Y. (2002). Telling from discrete data whether the underlying continuous-time model is a diffusion. J. Finance **57**, 2075-2122.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics **31**, 307-327.

Duan, J. C. (1995). The GARCH option pricing model. Mathematical Finance **5**, 13-32.

Duan, J. C. (1997). Augmented GARCH (p,q) process and its diffusion limit. Journal of Econometrics **79**, 97-127.

Duffie, D. (1992). Dynamic Asset Pricing Theory. Princeton University Press, Princeton, N.J. Second edition.

Dumas, B., Fleming, F. and Whaley, R. (1998). Implied volatility functions: empirical tests. Journal of Finance **53**, 2059-2106.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. Econometrica **50**, 987-1007.

Engle, R. F. and Ng, V. (1993). Measuring and Testing the Impact of News on Volatility. Journal of Finance **48**, 1749-1778.

Fouque, J. P., Papanicolaou, G. and Sircar, K. R. (2000). Derivatives in Financial Markets with Stochastic Volatility. Cambridge University Press.

Heston, S. (1993). A closed form solution for options with stochastic volatility with applications to bond and currency options. Review of Financial Studies **6**, 327-343.

Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. Journal of Finance **42**, 281-300.

Jacod, J. and Shiryaev, A. N. (1987). Limit Theorems for Stochastic Processes. Springer.

Johnson, H. and Shanno, D. (1987). Option pricing when the variance is changing. J. Financial and Quantitative Analysis **22**, 143-151.

Komlós, J. Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent RV's and the sample DF. I. Z. Wahrsch. Verw. Gebiete **32**, 111-131.

Komlós, J. Major, P. and Tusnády, G. (1976). An approximation of partial sums of independent RV's and the sample DF. II. Z. Wahrsch. Verw. Gebiete **34**, 33-58.

Lamoureux, C. G. and Lastrapes, W. D. (1993). Forecasting stock return variance: toward an understanding of stochastic implied volatilities. Review of Financial Studies **6**, 293-326.

Lyuu, Y. and Wu, C. (2005). On Accurate and Provably Efficient GARCH Option Pricing Algorithms. Quantitative Finance **5**, 181-198.

Melino, A. and Turnbull, S. (1990). Pricing foreign currency options with stochastic volatility. J. Econometrics **45**, 239-265.

Nelson, D. B. (1990). ARCH models as diffusion approximations. Journal of Econometrics **45**, 7-38.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: a new approach. Econometrica **59**, 347-370.

Ritchken, P. and Trevor, R. (1999). Pricing Options Under Generalized GARCH and Stochastic Volatility Processes. Journal of Finance **54**, 377-402.

Rossi, P. E. (1996). Modeling Stock Market Volatility. Academic Press.

Scott, L. (1987). Option pricing when the variance changes randomly: Theory, estimation and an application. J. Financial and Quantitative Analysis **22**, 419-438.

Stein, and Stein, (1991). Stock price distributions with stochastic volatility: An analytic approach. Review of Financial Studies **4**, 727-752.

Wang, Y. (2002). Asymptotic nonequivalence of GARCH models and diffusions. The Annals of Statistics 30, 754-783.

Wiggins, J. (1987). Option values under stochastic volatility: Theory and empirical estimates. J. Financial Economics **19**, 351-372.