Approximation of A Jump-Diffusion Process

Sanghoon Lee*
SungKyunKwan University

03 May 2003

Abstract

We present a weak convergence of a discrete time process to a jump-diffusion process as the length of sampling interval, h, goes to zero. There is an example given for the weak convergency with using $GARCH$ (1, 1)-M model by Engle and Bollerslev(1986). It is shown that $ARCH$ type models can be used as discrete time approximations of jump-diffusion processes. We use Exponential $ARCH$ with Poisson Jump component as an example for the approximation. Therefore, we may use a discrete time $ARCH$ process as an approximation of a jump-diffusion process in estimation and forecasting. And we may use the jump-diffusion process as an approximation of $ARCH$ process when there is distributional results available for the jump-diffusion limit of the sequence of $ARCH$ processes.

JEL classification: C22

Key word(s): Weak Convergence, ARCH Model, Jump-Diffusion Process

1 Introduction.

During the last couple of decades or so, many researchers have found that the value of option prices is not continuous with probability one. Cox and Ross(1975) assumed that the new information arriving at a market is a lump sum causing a discrete jump in the value of options and derived the option pricing formula by using a Poisson jump process. Merton, in his works $(1976a, b)$, decomposed the total change in the stock prices into two components: 1)systematic risk which is typically modelled by a Brownian motion, 2) nonsystematic risk which represents the arrivals of new information, in other word, shock, to the market, which can be modelled by a Poisson jump process, and derived the option pricing formula

*Sanghoon Lee, Department of Economics, SungKyunKwan University, 53 Myungnyun-Dong 3-Ga, Chonghno-Gu, Seoul, Korea. Tel: +82-2-760-1294, Fax: +82-2-744-5717, e-mail: sanghoon_l@hotmail.com

with jump-diffusion process. After these researches, as the solution to a stochastic asset optimization problem, the jump-diffusion process is popularly used in the dynamic asset pricing literature [e.g., Oldfield, Rogalski, and Jarrow(1977), Ball and Torous(1983, 1985), Jarrow and Rosenfeld(1984), Amin(1993), Kim, Oh, and Brooks(1994), Chang(1995)] as well as other financial economic literature such as in term structure of interest rates [e.g., Ahn and Thompson(1988), Das(1997)], foreign exchange rates [Jorion(1988), Ball and Roma(1993), Park, Ahn and Fujihara(1993), Vlaar and Palm(1993)] and etc.

For example, Ball and Torous(1983) considered the Bernoulli process to model the arrivals of information in a market and estimated the model with 47 NYSE(New York Stock Exchange) listed stocks each with 500 daily return observations. They compared estimates produced by Beckers′ cumulant method and Bernoulli cumulant method, then reported maximum likelihood estimates. Beckers′ cumulant method produced negative variances, $\widehat{\sigma}^2$ and $\widehat{\delta}^2$ in 60 percent of sampled stocks, which, respectively, denote the variances of the diffusion part and the jump size. However, they are reduced 20 percent by Bernoulli cumulant method. As it is expected, the maximum likelihood method did not produce any negative variances. When the cumulant method produce positive variances, the parameter estimates were similar to those of maximum likelihood estimates. And they considered the likelihood ratio test,

$$\Lambda = -2\left(\ln L\left(\underline{x} : \gamma^*\right) - \ln L\left(\underline{x} : \gamma^o\right)\right),$$

where γ^* and γ^o denote the maximum likelihood estimates of γ in the presence and the absence of Bernoulli jump structure, respectively. Under the null hypothesis that security returns were consistent with a lognormal diffusion process without Bernoulli jump components, Λ was asymptotically distributed χ^2 with two degree of freedom. Only five stocks did not demonstrate the presence of jumps at the 5 percent significant level. Moreover, over 78 percent of the stocks indicated the presence of jumps at 1 percent significant level. The result, therefore, confirms the presence of jumps in the majority of the sampled common stock returns. And, in another work of theirs (1985), they estimated the Poisson jump-diffusion process with 30 daily common stock returns. They also found the evidence that jump components were present in a majority of the stocks examined.

Another point we need to consider is that the financial time series is found to be highly heteroscedastic over time. There are massive amount of literature documenting the heteroscedastic nature of the financial time series data. With $ARCH$ models introduced by Engle(1983), the heteroscedastic nature of the data is well explained by the $ARCH$ type models. Since, then, the $ARCH$ type models are developed with many different parameterization for σ^2, including the $ARCH\left(p\right)$ model of Engle(1982), the $GARCH$ model of Bollerslev(1986), the $GARCH-M$ model of Bollerslev and Engle(1986), the log $-GARCH$ model of Pantula(1986), and Geweke(1986), and the *Exponential* $ARCH$ model of

Nelson(1989). As we cannot overview the whole of literature of $ARCH$ models in this paper, I would like to introduce the couple of survey papers here. The readers will find that the survey papers are well overviewed the $ARCH$ models in the various types of parameterization. The one is written by Bollerslev, Chou, and Kroner(1992) and the other by Bollerslev, Engle, and Nelson(1994). The first paper is overviewed $ARCH$ type models specifically used to model in the financial economics area and the second one including $ARCH$ type models used in the various area in Economics.

In this paper, we are trying to develop the relationship between the continuous time stochastic differential equation used in the theoretical literature of financial economics and the discrete time difference equation used in the lots of empirical works. There has been done relatively little work on the relation between the continuous time nonlinear stochastic differential equation systems, used in so much of the theoretical literature, and the $ARCH$ stochastic difference equation systems, favored by empirical workers. Indeed, the two literatures have developed quite independently, with little attempt to reconcile the discrete and continuous models. The reason why the empiricist being turned away from the continuous time stochastic differential equation system is difficult to derive the likelihood function of such models, especially when there are unobservable state variables in the system.[1] By the contrast, the reason why the empiricist being favored the discrete time $ARCH$ type models is that all economic data are collected at discrete time intervals and a discrete time $ARCH$ likelihood function is usually easy to derive and maximize. Therefore, we will develop conditions under which $ARCH$ stochastic difference equation systems converge in distribution to Ito process (Jump-Diffusion process) as the length of the discrete time interval goes to zero. What we hope to gain from this work are following: First, it may easier to estimate and forecast the parameters with $ARCH$ models with jump components observed at discrete time intervals. So, we may want to use $ARCH$ models with jump components as jump-diffusion approximations. Second, in some cases we may find that distributional results are available for the jump-diffusion limit of a sequence of $ARCH$ process with jump components, while the discrete time $ARCH$ models with jump components themselves are not available. In such cases, we may be able to use jump-diffusion processes as $ARCH$ approximations with jump components.

The structure of the paper is following. In section 2, we will sketch the main results in weak convergence of a sequence of stochastic difference equation to a jump-diffusion process. As an example, we use the $GARCH-M$ model of Bollerslev and Engle(1986). In section 3, an $ARCH$ jump-diffusion approximation will be presented. We show that it can approximate a wide variety of Generalized Ito process which is a jump-diffusion process. We will examine an example based on AR(1) Exponential ARCH model of Nelson(1991). Finally,

[1] See Lo(1988) for deriving the likelihood function such models.

we will conclude this paper in section 4.

Some conditions for the non-explosion in finite time stated in Appendix A. These conditions will be useful to prove the results in weak convergence. All the proofs of theorems are delegated to Appendix B.

2 Weak Convergence of the processes.

2.1 The Main Result in Weak Convergence

Here we want to show the weak convergence of a discrete time process to a jump-diffusion process. The basic theoretical setup is following.

Let $D([0,\infty), R^n)$ be the space of mappings from $[0,\infty)$ into R^n that are right continuous having finite left limits and let $B(R^n)$ denote the Borel sets on R^n. With introduction of an appropriate Skorohod metric, $D([0,\infty), R^n)$ becomes a complete metric space.[2] For each $h > 0$, let M_{kh} be the σ-algebra generated by kh, ${}_hX_0$, ${}_hX_h$, ${}_hX_{2h}$, $\ldots$, ${}_hX_{kh}$, and let ν_h be a probability measure on $(R^n, B(R^n))$. For each $h > 0$, and each $k = 0, 1, 2, \ldots$, define $\Pi_{h,kh}(x,\cdot)$, called a transition function on R^n, as follows:

i) $\Pi_{h,kh}(x,\cdot)$ is a probability measure on $(R^n, B(R^n))$ for all $x \in R^n$,

ii) $\Pi_{h,kh}(\cdot,\Gamma)$ is $B(R^n)$ measurable for all $\Gamma \in B(R^n)$.

For each $h > 0$, let P_h be the probability measure on $D([0,\infty), R^n)$ such that

$$P_h[{}_hX_0 \in \Gamma] = \nu_h(\Gamma) \text{ for any } \Gamma \in B(R^n), \tag{2.1}$$

$$P_h[{}_hX_t = {}_hX_{kh},\ kh \le t < (k+1)h] = 1, \tag{2.2}$$

$$P_h\left[{}_hX_{(k+1)h} \in \Gamma \mid \mathsf{M}_{kh}\right] = \Pi_{h,kh}({}_hX_{kh}, \Gamma) \text{ almost surely under } P_h \text{ for all } k \ge 0 \text{ and } \Gamma \in B(R^n). \tag{2.3}$$

Here, for each $h > 0$, we specify the distribution of the random starting point by (2.1) and form a continuous time process ${}_hX_t$ from the discrete time process ${}_hX_{kh}$ by (2.2) making ${}_hX_t$ a step function with jumps at time h, $2h$, $3h$, and so on. (2.3) specifies the transition probabilities of n-dimensional discrete time Markov process ${}_hX_{kh}$.[3]

[2] See [2] Kushner (1984) Section 4.3 in Chapter 2.

[3] In this somewhat complicated setup, out notation must keep track of three distinct kinds of processes:

(a) the sequence of discrete time processes $\{{}_hX_{kh}\}$ that depend both on h and on a discrete time index kh, k = 0,1,2, ...,

(b) the sequence of continuous time processes $\{{}_hX_t\}$ formed as step functions from the discrete time processes in (a) using (2.2). This process also depends on h and on a continuous time index t, $t \geq 0$, Now, define, for each $h > 0$,

$$a_h^{ij}(x,t) \equiv h^{-1}\int_{\|y-x\|\le 1}(y_i - x_i)(y_j - x_j)\,\Pi_{h,h[t/h]}(x,dy), \tag{2.4}$$

$$b_h^{i}(x,t) \equiv h^{-1}\int_{\|y-x\|\le 1}(y_i - x_i)\,\Pi_{h,h[t/h]}(x,dy), \tag{2.5}$$

$$\Delta_h^{\varepsilon}(x,t) \equiv h^{-1}\int_{\|y-x\|>\varepsilon}\Pi_{h,h[t/h]}(x,dy). \tag{2.6}$$

$$g_h(x,u) = x_t - x_{t^-}, \tag{2.7}$$

where $[t/h]$ is the integer part of t/h and $x_{t^-} = \lim_{s\to t} x_s$ for $s < t$.

$a_h(x,t)$ is a measure of the truncated second moment per unit of time, $b_h(x,t)$ is a measure of truncated drift per unit of time, $\Delta_h^{\varepsilon}(x,t)$ is a probability that the process has a jump of magnitude greater than ε per unit of time. We define the truncated first and second moment for the process x, since the usual conditional moments for the process may not be finite. For example, if $X_t = \exp[\exp W_t]$, where W_t is a Wiener process, X_t is a diffusion process, but there exist no moments of any order. And $g_h(x,u)$ measures the size of jump per unit of time. We suppose that a jump occurs with probability $\lambda h + o(h)$[4] in the time interval $[t, t+h)$. As we assume the process is right continuous with finite left limit, there will be only discontinuity of the first kind (i.e., discrete jumps) and the jump size will be finite.

Now, we state the assumptions which are required to obtain the weak convergence result. Let S^n denote the space of $n \times n$ matrices and let S_+^n denote the space of $n \times n$ symmetric nonnegative definite matrices.

Assumption 1. Let $a(x,t) : R^n \times [0,\infty) \to S_+^n$, $b(x,t) : R^n \times [0,\infty) \to R^n$ and $g(x,t) : R^n \times [0,\infty) \to R^n$ be continuous measurable mappings which are continuous in x for each $t \ge 0$. We assume that, for all $R > 0$, $T > 0$ and $\varepsilon > 0$,

$$\lim_{h\downarrow 0}\sup_{\|x\|\le R,\ 0\le t\le T}\|a_h(x,t) - a(x,t)\| = 0 \tag{2.8}$$

$$\lim_{h\downarrow 0}\sup_{\|x\|\le R,\ 0\le t\le T}\|b_h(x,t) - b(x,t)\| = 0 \tag{2.9}$$

$$\lim_{h\downarrow 0}\sup_{\|x\|\le R,\ 0\le t\le T}\|g_h(x,t) - g(x,t)\| = 0 \tag{2.10}$$

$$\lim_{h\downarrow 0}\sup_{\|x\|\le R,\ 0\le t\le T}\Delta_h^{\varepsilon}(x,t) = \lambda \tag{2.11}$$

(c) a limiting jump-diffusion process X_t to which, under conditions given below, the sequence of processes $\{{}_hX_t\}$ weakly converges as $h \to 0$.

To accommodate these three different processes, we indicate dependence on the length of sampling interval, h, to lower left of X and dependence on the time index to lower right.

[4] Let $f(h)$ and $g(h)$ be functions of h. $f(h) = o(g(h))$, if $\lim_{h\to 0}[f(h)/g(h)] = 0$ and $f(h) = O(g(h))$, if $\lim_{h\to 0}[f(h)/g(h)]$ is bounded.

This assumption requires that the second moment, drift, and jumps per unit of time converge uniformly on compact sets to well-behaved functions of time and the state variables x. And the probability of jump of size greater then ε converges to a constant λ. So, the sample paths of the limit process will have only discontinuity of the first kind with probability one.

Assumption 2. Let $\sigma(x,t) : R^n \times [0,\infty) \to S^n$ be a continuous measurable mapping such that, for all $x \in R^n$ and all $t \geq 0$,

$$a(x,t) = \sigma(x,t)\,\sigma(x,t)'. \tag{2.12}$$

This assumption requires that the function $a(x,t)$, the second moment per unit of time of the limit process, has a well-behaved matrix square root $\sigma(x,t)$.

Assumption 3. As $h \to 0$, ${}_hX_0$ converges in distribution to a random variable X_0 with probability measure ν_0 on $(R^n, B(R^n))$.

This assumption requires that the probability measure ν_h of the random starting points ${}_hX_0$ to converge to a limit measure ν_0 as $h \to 0$.

With all the assumptions we made above, we specified a initial probability measure ν_0 of the limit process, an instantaneous drift function $b(x,t)$, an instantaneous covariance matrix $a(x,t)$, and a jump amplitude $g(x,t)$. And we have supposed that the sample path of the process is discontinuous with probability one. However, there is no guarantee that a limit process is finite or is uniquely defined. There are a number of works considering the conditions under which ν_0, $a(x,t)$, and $b(x,t)$ uniquely define a diffusion limit process. Especially Strook and Varadhan(1979) studied extensively about the diffusion limit process. Ethier and Kurtz(1986) considers the martingale problems with Levy measure. We provide the conditions for a jump-diffusion limit process being finite in finite time in appendix A.

Assumption 4. ν_0, $a(x,t)$, $b(x,t)$, and $g(x,t)$ uniquely specify the distribution of a jump diffusion process X_t with initial distribution ν_0, diffusion matrix $a(x,t)$, drift vector $b(x,t)$, and jump amplitude $g(x,t)$.

Theorem 1 Under the assumptions 1 through 4, the sequence of ${}_hX_t$ processes defined by (2.1) to (2.3) converges weakly as $h \to 0$ to the X_t process defined by the stochastic integral equation

$$\begin{aligned} X_t &= X_0 + \int_0^t b(X_s,s)\,ds + \int_0^t \sigma(X_s,s)\,dW(s) \\ &\quad + \int_0^t \int g(X_{s^-},s)\,\widetilde{N}_\lambda(ds,dg) \end{aligned} \tag{2.13}$$

where W_t is an n-dimensional standard Brownian motion, independent of X_0, $\widetilde{N}_\lambda(ds, du)$ is the compensated Poisson process defined as $\widetilde{N}_\lambda(ds, du) = N(ds, du) - \lambda ds$[5] and where for any $\Gamma \in B(R^n)$, $P(X_0 \in \Gamma) = v_0(\Gamma)$. Such an X_t process exists and is distributionally unique. This distribution does not depend on the choice of $\sigma(\cdot, \cdot)$ made in Assumption 2. Finally, X_t remains finite in finite time intervals almost surely, i.e. for all $T > 0$,

$$P\left[\sup_{0 \le t \le T} \|X_t\| < \infty\right] = 1. \tag{2.14}$$

Now, we want to make the above result a bit more general. For each i, $i = 1, 2, \ldots, n$, each $\delta > 0$, and each $h > 0$, define

$$\gamma_{h,i,\delta}(x,t) \equiv h^{-1} \int_{R^n} |(y-x)_i|^{2+\delta} \Pi_{h,h[t/h]}(x, dy), \tag{2.15}$$

where $(y-x)_i$ is the ith element of the vector $(y-x)$. If, for some $\delta > 0$ and all i, $i = 1, 2, \ldots, n$, $\gamma_{h,i,\delta}(x,t)$ is finite, then the following integral will be well-defined and finite:

$$\begin{aligned} a_h^*(x,t) &\equiv h^{-1} \int_{R^n} (y-x)(y-x)' \Pi_{h,h[t/h]}(x, dy), \\ b_h^*(x,t) &\equiv h^{-1} \int_{R^n} (y-x) \Pi_{h,h[t/h]}(x, dy). \end{aligned}$$

They are the same measures as $a_h(\cdot,\cdot)$, $b_h(\cdot,\cdot)$ and $g_h(\cdot,\cdot)$, but integration is taken over R^n rather than $|y-x| \le 1$.

Assumption 1′. There exist $\delta > 0$ such that for each $R > 0$, each $T > 0$, and each i, $i = 1, 2, \ldots, n$,

$$\lim_{h \downarrow 0} \sup_{\|x\| \le R,\ 0 \le t \le T} \gamma_{h,i,\delta}(x,t) = 0. \tag{2.16}$$

Further, let $a(x,t) : R^n \times [0,\infty) \to S_+^n$, $b(x,t) : R^n \times [0,\infty) \to R^n$ and $g(x,t) : R^n \times [0,\infty) \to R^n$ be continuous measurable mappings which are continuous in x for each $t \ge 0$. We assume that for all $R > 0$, $T > 0$

$$\lim_{h \downarrow 0} \sup_{\|x\| \le R,\ 0 \le t \le T} \|a_h^*(x,t) - a(x,t)\| = 0, \tag{2.17}$$

$$\lim_{h \downarrow 0} \sup_{\|x\| \le R,\ 0 \le t \le T} \|b_h^*(x,t) - b(x,t)\| = 0, \tag{2.18}$$

$$\lim_{h \downarrow 0} \sup_{\|x\| \le R,\ 0 \le t \le T} \|g_h(x,t) - g(x,t)\| = 0, \tag{2.19}$$

$$\lim_{h \downarrow 0} \sup_{\|x\| \le R,\ 0 \le t \le T} \Delta_h^\varepsilon(x,t) = \lambda \tag{2.20}$$

[5] $N(ds, du)$ is a Poisson process, and λ is a constant probability of jump in the Poisson process, and $\lambda > 0$.

Theorem 2 Under Assumptions 1′, and 2 through 4, the conclusion of Theorem 1 hold.

As stated in Merton(1990, Ch.3), the assumption 1′ implies that the moments higher than two are vanishes to zero at an appropriate rate as $h \to 0$.

2.2 Example: $GARCH\,(1,1)$-M Model.

In Engle and Bollerslev(1986) , they presented the $GARCH\,(1,1)$-M process for the cumulative excess returns y_t on a portfolio. The excess return process is

$$\begin{aligned} y_t &= y_{t-1} + \mu\sigma_t^2 + \sigma_t Z_t, \\ \sigma_{t+1}^2 &= \omega + \sigma_t^2 \left[\beta + \alpha Z_t^2\right]. \end{aligned}$$

where Z_t ~i.i.d. $N\,(0,1)$.

Let us suppose that a stochastic process in discrete time is including the jump process as follows;

$$\begin{aligned} y_t &= y_{t-1} + \mu\sigma_t^2 + \sigma_t Z_t + c\eta_t, && (2.21) \\ \sigma_{t+1}^2 &= \omega + \sigma_t^2 \left(\beta + \alpha Z_t^2\right), && (2.22) \end{aligned}$$

where $Z_t \sim iid\ N\,(0,1)$ and η_t ~Bernoulli distributed with $\Pr(\eta_t = 0) = 1 - \lambda dt + o\,(dt)$ and $\Pr(\eta_t = 1) = \lambda dt + o\,(dt)$. And c denote the jump size of the process when a jump occurs.

Now, we partition the time interval more and more finely and examine the properties of the stochastic difference equation system. We allow the parameter α, β, and ω to depend on h and make the drift term in (2.21) proportional to h. Then we may rewrite the system of stochastic processes (2.21) and (2.22) as

$$\begin{aligned} {}_hy_{kh} &= {}_hy_{(k-1)h} + h\,\mu_h\,{}_h\sigma_{kh}^2 + h^{1/2}\,{}_h\sigma_{kh}\,{}_hZ_{kh} \\ &\quad + {}_h\eta_{kh}\left(c_h + \phi_h\,{}_hZ_{kh}\right), && (2.23) \\ {}_h\sigma_{(k+1)h}^2 &= \omega_h + {}_h\sigma_{kh}^2\left(\beta_h + \alpha_h\,{}_hZ_{kh}^2\right), && (2.24) \end{aligned}$$

and

$$\Pr\left[\left({}_hy_0,\ {}_h\sigma_0^2\right) \in \Gamma\right] = \upsilon_h\,(\Gamma) \qquad \text{for any } \Gamma \in B\left(R^2\right), \qquad (2.25)$$

where ${}_hZ_{kh}$ ~i.i.d. $N\,(0,1)$ and ${}_h\eta_{kh}$ ~Bernoulli distributed with $\Pr\left[{}_h\eta_{kh} = 0\right] = 1 - \lambda h + o\,(h)$ and $\Pr\left[{}_h\eta_{kh} = 1\right] = \lambda h + o\,(h)$. υ_h satisfy the assumption 3 as $h \to 0$ and for each $h \geq 0$, $\upsilon_h\left(\left(y_0, \sigma_0^2\right) : \sigma_0^2 > 0\right) = 1$. And we create the continuous time process ${}_hy_t$ and ${}_h\sigma_t^2$ by

$${}_hy_t \equiv {}_hy_{kh} \text{ and } {}_h\sigma_t^2 \equiv {}_h\sigma_{kh}^2 \quad \text{for } kh \leq t < (k+1)\,h. \qquad (2.26)$$

We want to find out which sequences $\{\omega_h,\ \alpha_h,\ \beta_h\}$ make the $\{{}_h\sigma_t^2,\ {}_hy_t\}$ process converge in distribution an jump-diffusion mixed process as $h \to 0$.

Let M_{kh} is the σ-algebra generated by kh, ${}_hy_0$, ${}_hy_h$, $\ldots$, ${}_hy_{(k-1)h}$, and ${}_h\sigma_0^2$, ${}_h\sigma_h^2$, $\ldots$, ${}_h\sigma_{kh}^2$. Then, the first moment of the process is

$$E\left[h^{-1}\left({}_hy_{kh} - {}_hy_{(k-1)h}\right)|\mathsf{M}_{kh}\right] = \mu_h\, {}_h\sigma_{kh}^2 + \lambda c_h \tag{2.27}$$

$$E\left[h^{-1}\left({}_h\sigma_{(k+1)h}^2 - {}_h\sigma_{kh}^2\right)|\mathsf{M}_{kh}\right] = h^{-1}\omega_h + h^{-1}\sigma_{kh}^2\left(\beta_h + \alpha_h - 1\right) \tag{2.28}$$

To satisfy the Assumption 1', we require the following limits exist and be finite;

$$\lim_{h\downarrow 0} h^{-1}\omega_h = \omega \tag{2.29}$$

$$\lim_{h\downarrow 0} h^{-1}\left(1 - \beta_h - \alpha_h\right) = \theta \tag{2.30}$$

As it is stated in Bollerslev(1986) , it is necessary to require that ω_h, α_h, and β_h be nonnegative because σ_t^2 should be remain positive with probability one. Therefore, $\omega \geq 0$ while θ could be of either sign.

Then,

$$\lim_{h\downarrow 0} E\left[h^{-1}\left({}_hy_{kh} - {}_hy_{(k-1)h}\right)|\mathsf{M}_{kh}\right] = \mu\sigma^2 + \lambda c \tag{2.31}$$

$$\lim_{h\downarrow 0} E\left[h^{-1}\left({}_h\sigma_{(k+1)h}^2 - {}_h\sigma_{kh}^2\right)|\mathsf{M}_{kh}\right] = \omega - \theta\sigma^2 \tag{2.32}$$

The second moment per unit of time is

$$\begin{aligned}
&E\left[h^{-1}\left({}_hy_{kh} - {}_hy_{(k-1)h}\right)^2|\mathsf{M}_{kh}\right] \\
&= h\mu_h^2\, {}_h\sigma_{kh}^4 + {}_h\sigma_{kh}^2 + \lambda\left(c_h^2 + \phi_h^2\right) + 2\lambda\, h\, c_h\, \mu_h\, {}_h\sigma_{kh}^2, \quad (2.33)\\
&E\left[h^{-1}\left({}_h\sigma_{(k+1)h}^2 - {}_h\sigma_{kh}^2\right)^2|\mathsf{M}_{kh}\right] \\
&= h^{-1}\omega_h^2 + h^{-1}\, {}_h\sigma_{kh}^4\left(\alpha_h + \beta_h - 1\right)^2 + 2h^{-1}\alpha_h^2\, {}_h\sigma_{kh}^4 \\
&+h^{-1}2\omega_h^2\, {}_h\sigma_{kh}^4\left(\alpha_h + \beta_h - 1\right), \quad (2.34)\\
&E\left[h^{-1}\left({}_hy_{kh} - {}_hy_{(k-1)h}\right)\left({}_h\sigma_{(k+1)h}^2 - {}_h\sigma_{kh}^2\right)|\mathsf{M}_{kh}\right] \\
&= \mu_h\, {}_h\sigma_{kh}^2\, \omega_h + \mu_h\, {}_h\sigma_{kh}^4\left(\alpha_h + \beta_h - 1\right) \\
&+\lambda_h\, c_h\, \omega_h + \lambda_h\, c_h\, {}_h\sigma_{kh}^2\left(\alpha_h + \beta_h - 1\right). \quad (2.35)
\end{aligned}$$

With (2.29) and (2.30) and assuming that

$$\lim_{h\downarrow 0} 2h^{-1}\alpha_h^2 = \alpha^2, \tag{2.36}$$

exist and is finite. And α^2 is always greater than 0. Then we have

$$E\left[h^{-1}\left({}_hy_{kh} - {}_hy_{(k-1)h}\right)^2|\mathsf{M}_{kh}\right] = {}_h\sigma_{kh}^2 + \lambda\left(c_h^2 + \phi_h^2\right) + o\left(1\right), \tag{2.37}$$

$$E\left[h^{-1}\left({}_h\sigma_{(k+1)h}^2 - {}_h\sigma_{kh}^2\right)^2|\mathsf{M}_{kh}\right] = \alpha_h^2\, {}_h\sigma_{kh}^4 + o\left(1\right), \tag{2.38}$$

$$E\left[h^{-1}\left({}_hy_{kh} - {}_hy_{(k-1)h}\right)\left({}_h\sigma_{(k+1)h}^2 - {}_h\sigma_{kh}^2\right)|\mathsf{M}_{kh}\right] = o\left(1\right) \tag{2.39}$$

where if $\psi(h) = o(1)$, then $\lim_{h \to 0} \psi(h) = 0$. We can show that the third and fourth moments of the process ${}_h\sigma_t^2$ exist and converge to zero as $h \to 0$, and those of the process ${}_hy_t$ exist and $O(h)$.[6]

Then, we can define the coefficients in the jump-diffusion mixed process as

$$b\left(y, \sigma^2\right) \equiv \begin{bmatrix} \mu\sigma^2 + \lambda c \\ \omega - \theta\sigma^2 \end{bmatrix} \tag{2.40}$$

$$a\left(y, \sigma^2\right) \equiv \begin{bmatrix} \sigma^2 + \lambda\left(c^2 + \phi^2\right) & 0 \\ 0 & \alpha^2\sigma^4 \end{bmatrix} \tag{2.41}$$

$$g\left(y, \sigma^2\right) \equiv \begin{bmatrix} c \\ 0 \end{bmatrix} \tag{2.42}$$

and, if α_h, β_h and ω_h satisfy the conditions in (2.29), (2.30) and (2.36), then Assumption $1'$ holds. If we suppose that $\sigma(\cdot, \cdot)$ in Assumption 2 is the element-by-element square root of $a(y, \sigma^2)$, then Assumption 2 holds as well. From (2.40) - (2.42), we can write the jump-diffusion mixed limit as

$$dy_t = \left(\mu\sigma^2 + \lambda c\right) dt + \left[\sigma^2 + \lambda\left(c^2 + \phi^2\right)\right]^{1/2} dW_{1,t} + c\eta_t \tag{2.43}$$

$$d\sigma^2_{(t+1)} = \left(\omega - \theta\sigma^2\right) dt + \alpha\sigma^2 dW_{2,t} \tag{2.44}$$

$$P\left[\left(y_0, \sigma_0^2\right) \in \Gamma\right] = \nu_0(\Gamma) \quad \text{for any } \Gamma \in B\left(R^2\right) \tag{2.45}$$

where $W_{i,t}$, $i = 1, 2$, are independent standard Wiener processes and are also independent of the Bernoulli process, η_t. And all those three independent processes are independent of the initial values (y_0, σ_0^2).

3 Jump-Diffusion Approximation.

In this section, we will present that $ARCH$ models can be used as approximation of generalized Ito process (jump-diffusion process).

3.1 $ARCH$ Jump-Diffusion Approximation

Define the stochastic differential equation system

$$\begin{aligned} dy_t = {} & f(s_t, y_t, t)\, dt + g(s_t, y_t, t)\, dW_{1,t} \\ & + \left\{k(s_t, y_t, t) + dt^{-1/2}\phi(s_t, y_t, t)\, dW_{1,t}\right\} d\eta_t, \end{aligned} \tag{3.1}$$

$$ds_t = F(s_t, y_t, t)\, dt + G(s_t, y_t, t)\, dW_{2,t}, \tag{3.2}$$

$$\begin{bmatrix} dW_{1,t} \\ dW_{2,t} \end{bmatrix} \begin{bmatrix} dW_{1t} & dW'_{2,t} \end{bmatrix} = \begin{bmatrix} 1 & \Omega_{1,2} \\ \Omega_{2,1} & \Omega_{2,2} \end{bmatrix} dt = \Omega dt, \tag{3.3}$$

$$\begin{bmatrix} d\eta_t \\ 0 \end{bmatrix} \begin{bmatrix} d\eta_t & 0' \end{bmatrix} = \begin{bmatrix} \lambda & 0_{1,2} \\ 0_{2,1} & 0_{2,2} \end{bmatrix} dt = \Lambda dt, \quad \text{and} \tag{3.4}$$

$$\Sigma dt = \Omega dt + \Lambda dt \tag{3.5}$$

[6] The detailed calculations are available on request.

where Ω and Λ are $(n+1)\times(n+1)$ positive semi-definite matrices of rank two or less, $0_{1,2}, 0_{2,1}, 0_{2,2}$ are $(n\times 1)$ column vector, $(1\times n)$ row vector, and $(n\times n)$ matrix of zeros, respectively, s_t is an n-dimensional vector of unobservable state variables, y is an observable scalar process, $W_{1,t}$ is one dimensional standard Wiener process, $W_{2,t}$ is an n-dimensional standard Wiener process, η_t is a Poisson process with intensity λ, $f(s_t, y_t, t)$, $g(s_t, y_t, t)$, and $k(s_t, y_t, t)$ are real-valued continuous scalar functions, and $F(s_t, y_t, t)$ and $G(s_t, y_t, t)$ are real, continuous $n\times 1$ and $n\times n$ valued functions, respectively. The initial values of the process (y_0, s_0) is assumed to be random and independent of $W_{1,t}$, $W_{2,t}$, and η_t, and $W_{1,t}$, $W_{2,t}$, and η_t, are independent of each other.

Define the matrix and vector functions a, b, and c by

$$a(y,s,t) = \begin{bmatrix} g^2+\lambda\left(k^2+\phi^2\right) & g\Omega_{1,2}G' \\ G\Omega_{2,1}g & G\Omega_{2,2}G' \end{bmatrix}, \tag{3.6}$$

$$b(y,s,t) = \begin{bmatrix} f+\lambda k & F' \end{bmatrix}', \tag{3.7}$$

$$c(y,s,t) = \begin{bmatrix} k & 0' \end{bmatrix}' \tag{3.8}$$

where 0 is an $n\times 1$ vector of zeros. Then, $a(y,s,t)$ is $(n+1)\times(n+1)$ matrix and $b(y,s,t)$ and $c(y,s,t)$ are $(n+1)\times 1$ vectors.

Now, we define a sequence of approximating processes that converge to (3.1)-(3.3) in distribution as $h\to 0$.

$$\begin{aligned} {}_h y_{kh} = {} & {}_h y_{(k-1)h} + f({}_h y_{kh}, {}_h s_{kh}, kh)\, h + g({}_h y_{kh}, {}_h s_{kh}, kh)\, {}_h Z_{kh} \\ & + \eta_{kh}\left(k({}_h y_{kh}, {}_h s_{kh}, kh) + h^{-1/2}\phi({}_h y_{kh}, {}_h s_{kh}, kh)\, {}_h Z_{kh}\right), \end{aligned} \tag{3.9}$$

$${}_h s_{(k+1)h} = {}_h s_{kh} + F({}_h y_{kh}, {}_h s_{kh}, kh)\, h + G({}_h y_{kh}, {}_h s_{kh}, kh)\, {}_h Z^*_{kh}, \tag{3.10}$$

where

$${}_h Z_{kh} \sim \text{i.i.d. } N(0,h), \tag{3.11}$$

$${}_h Z^*_{kh} = \begin{bmatrix} \theta_1\, {}_h Z_{kh} + \gamma_1\left[\left|{}_h Z_{kh}\right| - \left(\frac{2h}{\pi}\right)^{1/2}\right] \\ \theta_n\, {}_h Z_{kh} + \gamma_n\left[\left|{}_h Z_{kh}\right| - \left(\frac{2h}{\pi}\right)^{1/2}\right] \end{bmatrix}, \tag{3.12}$$

and the coefficients $\{\theta_1, \gamma_1, \ldots, \theta_n, \gamma_n\}$ are selected so that

$$E\begin{bmatrix} {}_h Z_{kh} + \eta_{kh} \\ {}_h Z^*_{kh} \end{bmatrix}\begin{bmatrix} {}_h Z_{kh} + \eta_{kh} & {}_h Z^*_{kh} \end{bmatrix} = \Sigma dt. \tag{3.13}$$

Now we can convert the discrete time process $[{}_h y_{kh}, {}_h s'_{kh}]$ into a continuous time process by defining

$${}_h y_t \equiv {}_h y_{kh},\ {}_h s_t \equiv {}_h s_{kh} \quad \text{for} \quad kh \le t < (k+1)h. \tag{3.14}$$

Theorem 3 If $a(y, s, t)$, $b(y, s, t)$ and $c(y, s, t)$ satisfy Assumption 4, with $x \equiv [y, s']$, and if the joint probability measures υ_h of the starting values $({}_h y_0, {}_h s'_0)$ converges to the measure υ_0 as $h \to 0$, then $({}_h y_t, {}_h s'_t) \Rightarrow (y_t, s'_t)$ as $h \to 0$.

The proof of this theorem is a direct application of Theorem 2. $\left({}_h y_{kh} - {}_h y_{(k-1)h}\right)$, $\left({}_h s_{(k+1)h} - {}_h s_{kh}\right)$ and h are discrete correspondence of dy, ds, and dt, respectively. The theorem 3.2 in Nelson (1990b) shows that ${}_h Z_{kh}$ and ${}_h Z^*_{kh}$ are the discrete time counterpart of dW_1 and dW_2.

3.2 AR(1) Exponential ARCH.

Here we consider a jump-diffusion process with conditional variance following a continuous time AR(1) process. Let's define a system of stochastic differential equations

$$d(\ln S_t) = \theta\sigma_t^2 dt + \sigma_t dW_{1,t} + k_t d\eta_t \tag{3.15}$$

$$d(\ln \sigma_t^2) = -\beta\left[(\ln \sigma_t^2) - \alpha\right] dt + dW_{2,t} \tag{3.16}$$

$$P\left[(\ln S_0, (\ln \sigma_0^2)) \in \Gamma\right] = \upsilon_0(\Gamma) \quad \text{for any } \Gamma \in B(R^2) \tag{3.17}$$

where S_t is the value of portfolio at time t, $W_{1,t}$ and $W_{2,t}$ are Wiener processes with

$$\begin{bmatrix} dW_{1,t} \\ dW_{2,t} \end{bmatrix} \begin{bmatrix} dW_{1,t} & dW_{2,t} \end{bmatrix} = \begin{bmatrix} 1 & \Omega_{1,2} \\ \Omega_{1,2} & \Omega_{2,2} \end{bmatrix} dt \equiv \Omega dt \tag{3.18}$$

and $\Omega_{2,2} \geq \Omega_{1,2}^2$, and η_t is a Poisson process with parameter λ and

$$\begin{bmatrix} d\eta_t \\ 0 \end{bmatrix} \begin{bmatrix} d\eta_t & 0 \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix} dt \equiv \Lambda dt. \tag{3.19}$$

Then, the variance matrix of the system of equation is

$$\Sigma dt = \Omega dt + \Lambda dt. \tag{3.20}$$

We want to find a sequence of $ARCH$ models converging weakly to (3.15)-(3.20) by using Theorem 3. As we assume that $(\ln \sigma_t^2)$ in (3.16) follows a continuous time $AR(1)$ process, $(\ln {}_h\sigma_{kh}^2)$ in the discrete counterpart of (3.16) will also follow an $AR(1)$ process. For each $h > 0$, we have

$$\begin{aligned} (\ln_h S_{kh}) &= \left(\ln_h S_{(k-1)h}\right) + h\theta\, {}_h\sigma_{kh}^2 + {}_h\sigma_{kh}\, {}_h Z_{kh} \\ &\quad + \eta_{kh}\left(k_h + h^{-1/2}\phi_h\, {}_h Z_{kh}\right), \end{aligned} \tag{3.21}$$

$$\begin{aligned} \left(\ln_h \sigma_{(k+1)h}^2\right) &= (\ln_h \sigma_{kh}^2) - \beta\left[(\ln {}_h\sigma_{kh}^2) - \alpha\right] h + \Omega_{1,2}\, {}_h Z_{kh} \\ &\quad + \gamma\left[|{}_h Z_{kh}| - \left(\tfrac{2h}{\pi}\right)^{1/2}\right], \end{aligned} \tag{3.22}$$

$$P\left[(\ln S_0, (\ln \sigma_0^2)) \in \Gamma\right] = \upsilon_0(\Gamma) \quad \text{for any } \Gamma \in B(R^2), \tag{3.23}$$

where $\gamma = [(\Omega_{2,2} - \Omega_{1,2}) / (1 - 2/\pi)]^{1/2}$ and ${}_hZ_{kh} \sim$i.i.d. $N(0, h)$. Then, we have

$$\begin{aligned} & E \left[\begin{array}{c} {}_hZ_{kh} + \eta_{kh} \\ \Omega_{1,2}\, {}_hZ_{kh} + \gamma \left[|{}_hZ_{kh}| - \left(\frac{2h}{\pi}\right)^{1/2} \right] \end{array} \right] \\ & \times \left[\begin{array}{cc} {}_hZ_{kh} + \eta_{kh} & \Omega_{1,2}\, {}_hZ_{kh} + \gamma \left[|{}_hZ_{kh}| - \left(\frac{2h}{\pi}\right)^{1/2} \right] \end{array} \right] \\ = & \left[\begin{array}{cc} 1 + \lambda & \Omega_{1,2} \\ \Omega_{1,2} & \Omega_{2,2} \end{array} \right] h \qquad (3.24) \\ \equiv & \Sigma h \end{aligned}$$

which is discrete counterpart of (3.20). As before, we define a continuous time step function

$${}_hS_t \equiv {}_hS_{kh},\ {}_h\sigma_t^2 \equiv {}_h\sigma_{kh}^2 \quad \text{for } kh \le t < (k+1)\, h.$$

And, the discrete time counterparts of $d[\ln S]$, $d[\ln \sigma^2]$, $dW_{1,t}$, $dW_{2,t}$, $d\eta_t$ and dt are $(\ln_h S_{kh}) - \left(\ln_h S_{(k-1)h}\right)$, $\left(\ln_h \sigma^2_{(k+1)h}\right) - (\ln_h \sigma^2_{kh})$, ${}_hZ_{kh}$, η_{kh}, and h, respectively.

Theorem 4 If the distribution of random starting point, υ_h, converges to υ_0 as $h \to 0$, then $\{{}_hS_t,\, {}_h\sigma_t^2\} \Rightarrow \{S_t,\, \sigma_t^2\}$ as $h \to 0$.

4 Conclusions

In this paper, we have shown that a stochastic difference equation converges weakly to a stochastic differential equation with jump component as length of sampling interval, h, goes to zero. We presented that, as an example, $GARCH(1,1)$-M process converges weakly to a jump-diffusion limit as h goes to zero. That is, a $ARCH$ type process can be approximated by stochastic jump-diffusion process. It may much easier to estimate and forecast with a discrete time stochastic difference equation than with a continuous time stochastic differential equation. So, when we observe a jump-diffusion model at discrete time intervals, we may use $ARCH$ model to estimate and forecast as jump-diffusion approximation.

It is shown that $ARCH$ process is a discrete approximation of jump-diffusion process with using Exponential $ARCH$ process with Poisson jump component. If there are distributional results available for the jump-diffusion limit of a sequence of discrete time stochastic difference equations and there is not available for the discrete time stochastic difference process, such as $ARCH$ process, then we may use jump-diffusion process as an approximation of $ARCH$ process.

Therefore, we may use a discrete time $ARCH$ process as an approximation of a jump-diffusion process in estimation and forecasting. And we may use the jump-diffusion process as an approximation of $ARCH$ process when there is distributional results available for the jump-diffusion limit of the sequence of $ARCH$ processes.

While we show the weak convergence of those processes, we fixed the Poisson jump intensity λ as a constant. This may be relaxed for the jump intensity to vary over time with a certain probability structure.

Appendix A : Conditions for Non-Explosion.

Theorem 10.2.1 in Strook and Varadhan(1975) provides a non-explosion condition for the limit process. This condition ensures that the limit process does not explode in finite time. In the theorem, the condition is given for the case of diffusion process. But, we adopt this condition for the jump-diffusion process with replacing the infinitesimal operator for a jump-diffusion process with that of a diffusion process.

Suppose that there exist a nonnegative function $\varphi(x,t)$ which is twice differentiable with respect to x and differentiable with respect to t such that for each $T>0$,

$$\lim_{|x|\to\infty}\inf_{0\le t\le T}\varphi(x,t)=\infty \tag{A.1}$$

and there exist positive locally bounded function $M(T)$ such that, for each $T>0$, for all $x\in R^n$, and for all t, $0\le t\le T$,

$$\left(\frac{\partial}{\partial t}+\mathsf{L}_t\right)\varphi(x,t)\le M(T)\,\varphi(x,t)\,, \tag{A.2}$$

where

$$\begin{aligned}\mathsf{L}\varphi(x,t) &= \sum_{i=1}^{n}b_i(x,t)\frac{\partial\varphi(x,t)}{\partial x_i}+\frac{1}{2}\sum_{i,j=1}^{n}a_{ij}(x,t)\frac{\partial^2\varphi(x,t)}{\partial x_i\partial x_j}\\ &\quad+\int\Bigl[\varphi(x+g(x,u),t)-\varphi(x,t)\\ &\qquad-\sum_{i=1}^{n}g_i(x,u)\frac{\partial\varphi(x,t)}{\partial x_i}\Bigr]N(h,du)\,. \end{aligned}\tag{A.3}$$

If we assume that $X_t=x$, $(B.2)$ ensures the instantaneous drift of $\varphi(X_t,t)$ grows linearly with $\varphi(X_t,t)$. Therefore, it guarantees that $\varphi(X_t,t)$ does not explode in finite time. Also, $(B.1)$ will guarantee that if $\varphi(X_t,t)$ does not explode, neither will X_t.

Appendix B : Proofs of The Theorems.

Proof of Theorem 1

Let the infinitesimal operator for the jump-diffusion process be

$$\begin{aligned}\mathsf{L}f(x) &= \sum_{i=1}^{n}b_i(x,t)\frac{\partial f}{\partial x_i}+\frac{1}{2}\sum_{i,j=1}^{n}a_{ij}(x,t)\frac{\partial^2 f}{\partial x_i\partial x_j}\\ &\quad+\int\left[f(x+c(x,u))-f(x)-\sum_{i=1}^{n}g_i(x,t)\frac{\partial f}{\partial x_i}\right]N(h,du)\,, \end{aligned}\tag{B.1}$$

then, by the *Lemma* 11.2.1 in Strook and Varadhan(1979) , the assumption 1 is equivalent to the condition that for each $f \in C_0^\infty(R^n)$

$$\frac{1}{h} A_h f \to \mathsf{L} f,$$

where the infinitesimal operator for a discrete Markov process A_h defined as

$$A_h f(x) = \int [f(y) - f(x)] \Pi_{h,h[t/h]}(x, dy).$$

Let's define a random process M_f as

$$M_f(t) = f(x(t)) - f(x(0)) - \int_0^t \mathsf{L} f(x(s)) ds.$$

If $M_f(t)$ is a martingale, then there exist a Wiener process $W(t)$, and a Poisson process $N_\lambda(t, \Gamma)$ with independent increments and identically distributed jumps which solve a jump diffusion model

$$x(t) = x(0) + \int_0^t b(x,s) ds + \int_0^t \sigma(x,s) dW(s) + \int_0^t \int g(x,u) \widetilde{N}_\lambda(ds, du), \tag{B.2}$$

where $g(x,u)$ is a bounded continuous function and $\widetilde{N}_\lambda(dt, du)$ is a compensated Poisson process defined as $\widetilde{N}_\lambda(ds, du) = N(ds, du) - \lambda ds$ with jump probability of λ.

In assumption 4, the distribution of X_t is specified by v_0, $a(x,t)$, $b(x,t)$, and $g(x,t)$. As $\sigma(x,t)$ only enters the equation through $a(x,t)$ function, the distribution of X_t does not depend on the choice of $\sigma(x,t)$ as long as $\sigma(x,t)\sigma'(x,t) = a(x,t)$. ■

Proof of **Theorem 2.**

To prove the theorem, we need to show that the Assumption 1′ implies Assumption 1. Then Theorem 2 follows immediately by Theorem 1. To do so, we only need to prove that

$$\lim_{h\to 0} \sup_{|x|\le R,\, 0\le t\le T} \frac{1}{h} \int_{\|y-x\|>1} (y-x)_i^2 \Pi_{h,h[t/h]}(x, dy) = 0, \tag{B.3}$$

$$\lim_{h\to 0} \sup_{|x|\le R,\, 0\le t\le T} \frac{1}{h} \int_{\|y-x\|>1} |y-x|_i \Pi_{h,h[t/h]}(x, dy) = 0, \tag{B.4}$$

since the conditions for $c(x,u)$ and $\Delta_h^\varepsilon(x,t)$ remain same as in Assumption 1.

By Hőlder's Integral Inequality,

$$\begin{aligned} & \frac{1}{h} \int_{\|y-x\|>1} |y-x|_i \Pi_{h,h[t/h]}(x, dy) \\ \le\ & \left[\gamma_{h,i,\delta}(x,t)\right]^{1/(2+\delta)} \left[\Delta_h^\varepsilon(x,t)\right]^{(1+\delta)/(2+\delta)}. \end{aligned} \tag{B.5}$$

And by (2.16), there is some $\delta > 0$ such that for all R, $T > 0$, the right hand side of the inequality vanishes to zero for every ε as $h \to 0$ uniformly on $\|x\| \leq R$, $0 \leq t \leq T$, proving (B.4). Again, by Hőlder's Integral Inequality

$$\begin{aligned} & \frac{1}{h}\int_{\|y-x\|>1} (y-x)_i^2 \, \Pi_{h,h[t/h]}(x,dy) \\ \leq \; & \left[\gamma_{h,i,\delta}(x,t)\right]^{2/(2+\delta)} \left[\Delta_h^{\varepsilon}(x,t)\right]^{\delta/(2+\delta)} \end{aligned} \tag{B.6}$$

which vanishes in the same manner as (B.3). ■

Proof of Theorem 3.

To prove this theorem, we need to show the Assumption 2 and 1 are satisfied. First we can factor $a(y,s,t)$ into $\sigma(y,s,t)\,\sigma'(y,s,t)$ which satisfies Assumption 2. To show this

$$\begin{aligned} a(y,s,t) &= \begin{bmatrix} g^2+\lambda\left(k^2+v^2\right) & g\Omega_{1,2}G' \\ G\Omega_{2,1}g & G\Omega_{2,2}G' \end{bmatrix} \\ &= \begin{bmatrix} g^2 & g\Omega_{1,2}G' \\ G\Omega_{2,1}g & G\Omega_{2,2}G' \end{bmatrix} + \begin{bmatrix} \lambda\left(k^2+v^2\right) & 0_{1,2} \\ 0_{2,1} & 0_{2,2} \end{bmatrix} \\ &= \left\{ \begin{bmatrix} gdW_{1,t} \\ GdW_{2,t} \end{bmatrix} + \begin{bmatrix} \left(k+dt^{-1/2}vdW_{1,t}\right)d\eta_t \\ 0 \end{bmatrix} \right\} \\ &\quad \times \left\{ \begin{bmatrix} gdW_{1t} & dW_{2,t}'G' \end{bmatrix} + \begin{bmatrix} \left(k+dt^{-1/2}vdW_{1,t}\right)d\eta_t & 0' \end{bmatrix} \right\} \\ &= \begin{bmatrix} gdW_{1,t} + \left(k+dt^{-1/2}vdW_{1,t}\right)d\eta_t \\ GdW_{2,t} \end{bmatrix} \\ &\quad \times \begin{bmatrix} gdW_{1t} + \left(k+dt^{-1/2}vdW_{1,t}\right)d\eta_t & dW_{2,t}'G' \end{bmatrix} \\ &= \sigma\sigma' \end{aligned} \tag{B.7}$$

Now we must show that Assumption 1′ is satisfied. That is, we need to show that $a_h^*(y,s,t)$, $b_h^*(y,s,t)$, and $c_h(y,s,t)$ converge to $a(y,s,t)$, $b(y,s,t)$ and $c(y,s,t)$ respectively, and $\gamma_{h,i,\delta}(y,s,t)$ converges to zero uniformly on compacts as $h \to 0$. Since

$$b_h^*(y,s,t) = \begin{bmatrix} f(s,y,t)+\lambda k(y,s,t) \\ F(s,y,t) \end{bmatrix}, \tag{B.8}$$

$b_h^*(s,y,t) = b(s,y,t)$.

$$a_h^*(y,s,t) = \begin{bmatrix} f^2h+g^2+\lambda\left(k^2+v^2\right) & fhF'+G\Omega_{1,2}G' \\ Ffh+G\Omega_{2,1}g & G\Omega_{2,2}G' \end{bmatrix}, \tag{B.9}$$

which will converge to $a(y,s,t)$ as $h \to 0$, since f, F, g and G are locally bounded.

Finally, if we choose $\delta = 1$, then

$$\begin{aligned}\gamma_{h,1}(y,s,t) &= h^{-1}E\left[\frac{\left|hf + g_h Z_{kh} + \eta_{kh}\left(k + h^{-1/2}v_h Z_{kh}\right)\right|^3}{\left|hF + G_h Z^*_{kh}\right|^3}\right] \\ &= \left[\frac{\sqrt{h}\left(g^2 + \lambda\left(k^2 + v^2\right)\right)\sqrt{fh + g^2 + \lambda\left(k^2 + v^2\right)}}{\sqrt{h}G\Omega_{2,2}G'\sqrt{FF'h + G\Omega_{2,2}G'}}\right] \\ &= O\left(h^{1/2}\right)\end{aligned}$$

uniformly on compacts.

Therefore, we showed that Assumption 2 and 1' are satisfied. ■

Proof of Theorem 4.

To proof the theorem, we need to show the Assumption 4 is satisfies, then the result follows immediately by theorem 2 and theorem 4. To do so, we need to show that the system of stochastic differential equation has unique solution.

i) Show that the martingale problem for $a(\cdot,\cdot)$, $b(\cdot,\cdot)$, and $c(\cdot,\cdot)$ is well posed.

ii) Show that the limit process does not explode.

By Chapter 8, Theorem 3.3 in Ethier and Kurtz(1986) and Theorem 11.2.3 in Strook and Varadhan(1979), we can prove the statement i). To prove the statement ii), define for $K > 0$,

$$\varphi \equiv K + f(S)|S| + f(V)\exp(|V|),$$

where

$$f(x)\begin{cases} \equiv \exp\left(-\frac{1}{|x|}\right), & \text{if } x = 0, \\ \equiv 0, & \text{otherwise.}\end{cases}$$

$\varphi(V,S)$ is nonnegative, arbitrarily differentiable and satisfies (A.1). Its derivatives are locally bounded, so that positive K and M can be chose to satisfy (A.2) on any compact set. For large values of S and V

$$\begin{aligned}\varphi_V(V,S) &\approx sign(V)\exp(|V|), \\ \varphi_{VV}(V,S) &\approx \exp(|V|), \\ \varphi_S(V,S) &\approx sign(S), \\ \varphi_{SS}(V,S) &\approx 0,\end{aligned}$$

so that with $M > 1 + \alpha\beta + \Omega_{2,2}/2 + |\theta| + \lambda k$, there exist a finite k satisfying (A.2). Then, the result follows by Theorem 2. ■

References

[1] Ahn, C. M. and H. E. Thompson (1988), "Jump-Diffusion Processes and the Term Structure of Interest Rates," Journal of Finance, 43, 155 -174.

[2] Amin, K. I.(1993), "Jump Diffusion Option Valuation in Discrete Time," Journal of Finance, 48, 1833-1863.

[3] Arnold, L. (1974), Stochastic Differential Equations: Theory and Applications, John Wiley & Son, New York, NY.

[4] Ball, C. A. and A. Roma (1993), "A Jump Diffusion Model for the European Monetary System," Journal of International Money and Finance, 12, 475-492.

[5] Ball, C. A. and W. N. Torous (1983), "A Simplified Jump Process for Common Stock Returns," Journal of Financial and Quantitative Analysis, 18, 53-65.

[6] ______________________ (1985), "On Jumps in Common Stock Prices and Their Impact On Call Option Pricing," Journal of Finance, 40, 155-173.

[7] Bhattacharya, R. N. and E. C. Waymire (1990), Stochastic Processes with Applications, John Wiley & Son, New York, NY.

[8] Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroscedasticity," Journal of Econometrics, 31, 307-327.

[9] Bollerslev, T, R. Y. Chou, and K. F. Kroner (1992), "*ARCH* modelling in Finance: A Review of Theory and Empirical Evidence," Journal of Econometrics, 52, 5-59.

[10] Bollerslev, T., R. F. Engle, and D. B. Nelson(1994), "*ARCH* Models," in Handbook of Econometrics, Vol IV, eds. R.F. Engle and D. McFadden, Amsterdam, North-Holland, 2959-3038.

[11] Chang, C. W.(1995), "A No-Arbitrage Martingale Analysis for Jump-Diffusion Valuation," Journal of Financial Research, 18, 351-381.

[12] Cox, J. C. and S. A. Ross (1975), "The Pricing of Option for Jump Processes," Rodney L. White Centre Working Paper Series, No. 2-75, University of Pennsylvania, Philadelphia, Penn.

[13] Das, S. R. (1997), "An Efficient Generalized Discrete Time Approach to Poisson-Gaussian Bond Option Pricing in the Heath-Jarrow-Morton Model," NBER Technical Working Paper, No. 212, Cambridge, MA.

[14] Engle, F. and T. Bollerslev, (1986), "Modelling the Persistence of Conditional Variances," Econometrics Reviews, 5, 1-50.

[15] Ethier, S. N. and T. G. Kurtz (1986), Markov Process : Characterization and Convergence, John Wiley & Son, New York, NY,

[16] Gihman, I. I. and A. V. Skorohod (1972), Stochastic Differential Equation, Springer-Verlag, Berlin, Heidelberg.

[17] Jarrow, R. A. and E. R. Rosenfeld (1984), "Jump Risks and the Intertemporal Capital Asset Pricing Model," Journal of Business, 57, 337-351.

[18] Jorion, P. (1989), "On Jump Processes in the Foreign Exchange and Stock Markets," Review of Financial Studies, 1, 427-445.

[19] Kim, M-J, Y-H Oh, and R. Brooks (1994), "Are Jumps in Stock Returns Diversfiable? Evidence and Implications for Option Pricing," Journal of Financial and Quantitative Analysis, 29, 609-631

[20] Kushner, H. J. (1984), Approximation and Weak Convergence Methods for Random Processes, with Applications to Stochastic Systems Theory, MIT Press, Cambridge, MA.

[21] Merton, R. C. (1976*a*), "Option Pricing When Underlying Stock Returns are Discontinuous," Journal of Financial Economics, 3, 125-144.

[22] ___________ (1976*b*), "The Impact on Option Pricing of Specification Error in the Underlying Stock Returns," Journal of Finance, 31, 333-350.

[23] ___________ (1990), Continuous Time Finance, Cambridge, MA: Basil Blackwell.

[24] Nelson, D. B.(1990), "*ARCH* Models as Diffusion Approximations," Journal of Econometrics, 45, 7-38.

[25] ___________(1991), "Conditional Heteroscedasticity in Asset Returns: New Approach," Econometrica, 59, 347-390.

[26] Oldfield, G. S., R. J. Rogalski, and R. A. Jarrow (1977), "An Autoregressive Jump Process for Common Stock Returns," Journal of Financial Economics, 5, 389-418.

[27] Park, K., C. M. Ahn, and R. Fujihara (1993), "Optimal hedged Portfolio: the Case of Jump-Diffusion Risks," Journal of International Money and Finance, 12, 493-510.

[28] Protter, P. (1990), Stochastic Integration and Differential Equations: A New Approach, Springer-Verlag, Berlin, Heidelberg.

[29] Strook, D. W. and S. R. S. Varadhan (1979), Multidimensional Diffusion Processes, Springer-Verlag, New York.

[30] Vlaar, P. J. G. and F. C. Palm (1993), "The Message in Weekly Exchange Rates in the European Monetary System: Mean Reversion, Conditional Heteroscedasticity and Jumps," Journal of Business and Economic Statistics, 11, 351-360.