



A Markov Chain Approximation for American Option Pricing in Tempered Stable-GARCH Models

Xiang Shi^{1*}, Lihua Zhang² and Young S. A. Kim³

¹ Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY, USA, ² Business School of Administration, Zhejiang University of Technology, Hangzhou, China, ³ College of Business, Stony Brook University, Stony Brook, NY, USA

This paper considers the American option pricing problem under the stochastic volatility models. In particular, we introduce the GARCH model with two heavy-tailed distributions: classical tempered stable (CTS) and normal tempered stable (NTS) distribution. Then we apply the Markov chain approach to compute the prices of American style options under these two models. Minimal entropy provides a convenient way to construct equivalent martingale measure (EMM) and allows us to overcome the difficulties in incorporating the Markov chain approximation. The convergence of the approximation is also proved. Both numerical and empirical results are analyzed to show the advantages and drawbacks of our approach.

OPEN ACCESS

Edited by:

Sergei Levendorskii,
University of Leicester, UK

Reviewed by:

Jiayao Xie,
Ernst & Young, Hong Kong
Oleg Kudryavtsev,
Russian Customs Academy Rostov
Branch, Russia

*Correspondence:

Xiang Shi
xiang.shi@stonybrook.edu

Specialty section:

This article was submitted to
Mathematical Finance,
a section of the journal
Frontiers in Applied Mathematics and
Statistics

Received: 20 July 2015

Accepted: 10 December 2015

Published: 08 January 2016

Citation:

Shi X, Zhang L and Kim YSA (2016) A
Markov Chain Approximation for
American Option Pricing in Tempered
Stable-GARCH Models.
Front. Appl. Math. Stat. 1:13.
doi: 10.3389/fams.2015.00013

Keywords: tempered stable distribution, GARCH, American options, Markov chain, minimal entropy

1. INTRODUCTION

It is commonly known that the Black-Scholes model does not describe the real behavior of derivatives very well. The model is not consistent with the observed behavior found in asset returns, such as volatility clustering, skewness and fat tails. Many studies have been conducted to modify Black-Scholes model to incorporate these three empirical features of the asset returns. The α -stable distribution proposed by Mandelbrot [1] is an alternative to the normal distribution for modeling asset returns because it allows for skewness and fat tails. While the α -stable distribution has certain desirable properties, it is not suitable in modeling option prices, because it does not have the second moment or even the first moment. In order to obtain a well-defined model for pricing options, the mean, variance, and exponential moments of the return distribution have to be finite. Therefore, there are many modified distributions proposed for financial modeling based on the α -stable distribution. For example, the classical tempered stable (CTS) distribution [2, 3], normal inverse Gaussian (NIG) distribution [4] and its generalization called normal tempered stable (NTS) distribution [5], the modified tempered stable (MTS) distribution [6] and an extension of the CTS distribution named the KR distribution [7]. These distributions, called the tempered stable distributions, not only can incorporate the fat tails and skewness of asset returns, but also have finite moments for all orders. Furthermore, they are all infinitely divisible, i.e., the random variable can be decomposed to a sequence of i.i.d random numbers from the same family. So each of them uniquely corresponds to a Lévy process. Therefore, by replacing the Brownian motion by these Lévy processes, one can extend Black Scholes model to the so called exponential Lévy models. However, these Lévy exponential models are based on the assumptions that the returns are identically and independently distributed (i.i.d). Therefore, they are not able to explain the volatility clustering phenomena.

GARCH model have been developed to model the asset returns under the assumption of a non- Markovian property, more precisely, the assumption of volatility clustering. In Duan [8, 9], Kallsen[10], Heston [11], and Christoffersen [12] theories have been developed for pricing derivatives contracts using GARCH models. Most of researches assume that the innovations of GARCH are normally distributed. Although one can show that the normal GARCH model has heavier tails than Gaussian white noise, i.e., the kurtosis is greater than 3, it is still not able to fully capture the extreme values and skewness of the return process. To overcome the drawback of normal innovation process, Menn and Rachev [13, 14] introduced an enhanced GARCH model with innovations which follow the smoothly truncated stable distribution. Stentoft [15] proposed a GARCH model with NIG distribution for American option pricing. Kim et al. [7] applied the three tempered stable distributions (CTS, MTS, and KR distributions) to model the residual distribution in the GARCH process, and presented the CTS-GARCH model, MTS-GARCH model, and KR-GARCH model. Kim et al. [16] presented the RDTSGARCH option pricing model. These models can incorporate the three empirical features of the asset returns pretty well. In this paper, we will consider the two most widely used tempered stable distributions to model the residual distribution: classical tempered stable (CTS) and normal tempered stable (NTS).

However, it is difficult to price American options under the GARCH models with tempered stable distributions. It is widely recognized that the existing methods for pricing American options under GARCH are bivariate tree method and least square simulation method presented by Longstaff and Schwartz [17]. Ritchken and Trevor [18], Duan and Simonato [19] have introduced a modified lattice approach and a Markov chain method, respectively. These methods are applied to price American options under the normal-GARCH model. Rachev et al. [20] applied the least square approach to price American options under different TS-GARCH models. However, the least square method might fail to give accurate results when the option is out of the price [21]. Furthermore, increasing the order of the approximating polynomials may not give a good fit to the price when the returns have heavier tails, i.e., more extreme values.

In this paper, we consider the Markov chain approach presented by Duan and Simonato [19] and apply it to the CTS-GARCH and NTS-GARCH models. The idea of the approach is to discretized the stock price and the volatility, and use a Markov chain to approximate the original process. The fixed partition of the state space simplifies the task of option pricing to a sequence of standard matrix operations.

There are two major difficulties in applying this approach to TS-GARCH models. First, the distribution of the residuals changes over time under the original risk neutral argument of Rachev et al. [20]. Thus, we have to recompute the transition matrix each time which makes the Markov chain approach time-consuming. So we apply the minimal entropy method to determine the risk neutral measure. This method provides a large degree of freedom to construct the model

in the risk neutral world. By applying this approach we are allowed to use the same transition matrix when computing the option prices. The second difficulty is how to determine the partition conditions of TS GARCH effectively. The skewness and heavy tails of the tempered stable distributions implies that we need a different way to discretize stock price and volatility in order to make the approximating Markov chain convergent.

The numerical examples shows that the option values computed by the Markov chain method converge to the theoretical value. The computation time of the approach is not as expensive as one might expect since the transition probability matrix is highly sparse. It works well for the TS-GARCH option pricing framework. However, the convergence speed of Markov chain method depends on GARCH parameters. For stocks with relatively higher volatilities, the convergence rate becomes slow even for the normal GARCH model, since we have to increase the range of the stock price and volatilities. It is not strange that the convergence speed under the TS-GARCH model would be slower than the normal GARCH. Nevertheless, the TS-GARCH model captures the stylized facts of the market and therefore provides a better fit to the stock price, see Stentoft [15], Kim et al. [7] and Rachev et al. [20]. In the empirical analysis we illustrate how to apply our approach to compute the American option prices in the real world. Calibrated risk neutral parameters and minimal entropy parameters are computed and compared.

The paper is organized as follows: Section 2 contains the introduction to the GARCH model with the tempered stable distributions. Section 3 presents the Markov chain method under the CTS-GARCH, and NTS-GARCH models. The convergence of the Markov chain method is presented in Section 4. In Section 5, we conduct some numerical experiments and empirical analysis to illustrate our approach. The last section contains the concluding remarks.

2. GARCH MODEL WITH TEMPERED STABLE DISTRIBUTED INNOVATIONS

In this section we briefly review the GARCH model with tempered stable innovations introduced by Rachev et al. [20]. Furthermore, we apply the minimal entropy approach to determine the equivalent martingale measure.

2.1. Tempered Stable Distributions

First we introduce two most widely-used tempered stable distributions: the CTS distribution and the NTS distribution.

The CTS distribution with parameter $\alpha, C, \lambda_+, \lambda_-, m$ is defined by its characteristic function:

$$\begin{aligned} \phi_{CTS}(u; \alpha, C, \lambda_+, \lambda_-, m) \\ = \exp(ium - iuC\Gamma(1 - \alpha)(\lambda_+^{\alpha-1} - \lambda_-^{\alpha-1}) + \\ C\Gamma(-\alpha)((\lambda_+ - iu)^\alpha - \lambda_+^\alpha + (\lambda_- + iu)^\alpha - \lambda_-^\alpha)), \end{aligned}$$

where $C, \lambda_+, \lambda_- > 0, \alpha \in (0, 2) \setminus \{1\}$ and $m \in \mathbb{R}$. We denote $X \sim \text{CTS}(\alpha, C, \lambda_+, \lambda_-, m)$ if the random variable X follows

CTS distribution. The parameters m and C are the location and scale parameters respectively. α is the key parameter related to the heavy tail. And λ_+ and λ_- control the rate of decay on the positive and negative tails respectively. The distribution is symmetric if $\lambda_+ = \lambda_-$. A Lévy process induced from the CTS distribution is called the classical tempered stable process with parameters $(\alpha, C, \lambda_+, \lambda_-, m)$.

The mean and the variance of the CTS distribution are given by

$$E[X] = m;$$

$$Var(X) = C\Gamma(2 - \alpha)(\lambda_+^{\alpha-2} - \lambda_-^{\alpha-2}).$$

Therefore, if we set $m = 0$ and

$$C = (\Gamma(2 - \alpha)(\lambda_+^{\alpha-2} - \lambda_-^{\alpha-2}))^{-1},$$

then X has zero mean and unit variance. In this case, we say that X follows the *standard CTS distribution* with parameters $(\alpha, \lambda_+, \lambda_-)$ and denoted by $X \sim \text{stdCTS}(\alpha, \lambda_+, \lambda_-)$.

Let $\alpha \in (0, 2)$, $\theta, \gamma > 0$ and $\beta, m \in \mathbb{R}$. X is said to follow the normal tempered stable (NTS) distribution, if the characteristic function of X is given by:

$$\phi_{NTS}(u; \alpha, \theta, \gamma, \beta, m)$$

$$= \exp\left(iu(m - \beta) - \frac{2\theta^{1-\frac{\alpha}{2}}}{\alpha} \left(\left(\theta - i\beta u + \frac{\gamma^2 u^2}{2} \right)^{\frac{\alpha}{2}} - \theta^{\frac{\alpha}{2}} \right)\right).$$

We denote an NTS distributed random variable X by $X \sim \text{NTS}(\alpha, \theta, \gamma, \beta, m)$. α and θ control the rate of tail decay; γ and β control the variance; and β determines the skewness. If $\beta > 0$, the distribution is skewed to the left. The mean and variance of NTS are given by

$$E[X] = m;$$

$$Var(X) = \gamma^2 + \beta^2 \left(\frac{2 - \alpha}{2\theta} \right).$$

If $|\beta| < \sqrt{\frac{2\theta}{2-\alpha}}$, then we are able to set $m = 0$ and $\gamma = \sqrt{1 - \beta^2 \left(\frac{2-\alpha}{2\theta} \right)}$ so that X has zero mean and unit variance. In this case, we say that X follows the *standard NTS distribution* and denoted by $X \sim \text{stdNTS}(\alpha, \theta, \beta)$.

The tempered stable distributions, including CTS and NTS, have heavier tails than Gaussian distribution. On the other hand, their tails are thinner than the α -stable distribution whose tail has a power decay. One can show that the moments of CTS and NTS are all finite. For details we refer to Cont and Tankov [22] and Rachev et al. [20].

2.2. GARCH Option Pricing Model

2.2.1. Under the Market Measure \mathbb{P}

In this section we introduce the GARCH model with heavy tailed distributions under the market measure. We do not consider dividends here for simplicity. Let $(\Omega, \{\mathcal{F}_t\}_{t=0,1,\dots,T}, \mathbb{P})$ denote the standard probability space, where \mathcal{F}_t is the sigma algebra

represents the information up to time t . We assume that the dynamics of the stock price is given by

$$X_t := \log\left(\frac{S_t}{S_{t-1}}\right) = \mu_t + \sigma_t \varepsilon_t, \quad t = 1, 2, \dots, T,$$

where S_t is the stock price and X_t is the log return at time t . Both μ_t and σ_t are predictable which represent conditional mean and conditional variance on t respectively. The well-known GARCH(1,1) process is given by

$$\sigma_t^2 = \alpha_0 + \alpha_1(X_{t-1} - \mu_{t-1})^2 + \beta_1 \sigma_{t-1}^2, \quad t = 1, 2, \dots, T,$$

$$\varepsilon_0 = 0, \tag{1}$$

where α_0, α_1 , and β_1 are positive parameters satisfies $\alpha_1 + \beta_1 < 1$ since we want the series to be stationary.

There are different ways to model μ_t under market measure \mathbb{P} . For example, if $\{\varepsilon_t\}_{t=1,2,\dots,T}$ are i.i.d normally distributed, we can set the conditional mean to be $\mu_t = r + \lambda \sigma_t - \sigma_t^2/2$ for $t = 1, 2, \dots, T$, where r is the risk free interest rate and λ is the market price of risk. Then we are able to construct a risk neutral dynamics in a way similar to Black Scholes model, see Duan [8]. If $\{\varepsilon_t\}_{t=1,2,\dots,T}$ follows infinitely divisible distribution like CTS or NTS, we can extend Duan's model by setting $\mu_t = r + \lambda_t \sigma_t - g(\sigma_t)$ where $g(x) := \log E[\exp(x\varepsilon_t)]$. For tempered stable distributions the moment generating function $E[\exp(x\varepsilon_t)]$ may not be well-defined on the real line. In that case, we require σ_t^2 to be smaller than a positive number ρ otherwise $g(\sigma_t)$ will turn to infinity. So we would rewrite the Equation (1) as:

$$\sigma_t^2 = \min(\alpha_0 + \alpha_1 \sigma_{t-1}^2 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \rho) \quad t = 1, 2, \dots, T,$$

In practice it is very unlikely that σ_t^2 will reach the upper bound ρ . So the properties of GARCH process such as volatility clustering still apply to our modified model.

2.2.2. Under the Risk Neutral Measure \mathbb{Q}

Now we want to find an equivalent risk neutral measure \mathbb{Q} so that the dynamics of stock prices $S_t, t = 1, 2, \dots, T$ is a martingale, i.e.,

$$S_{t-1} = E^{\mathbb{Q}}[e^{-r} S_t | \mathcal{F}_{t-1}] = S_{t-1} E^{\mathbb{Q}}[e^{X_t - r} | \mathcal{F}_{t-1}] \quad t = 2, \dots, T.$$

Rachev et al. [20] define the risk neutral measure \mathbb{Q} to be a product of a sequence of independent measures $\mathbb{Q}_1, \dots, \mathbb{Q}_T$, such that $\xi_t := \varepsilon_t + k_t$ is a standard tempered stable distribution under \mathbb{Q}_t , where

$$k_t = \lambda_t + \frac{1}{\sigma_t} (\tilde{g}(\sigma_t) - g(\sigma_t)).$$

where \tilde{g} is the log moment generating function of ξ_t . Therefore, we can obtain the risk neutral dynamic of S_t :

$$\log\left(\frac{S_t}{S_{t-1}}\right) = r - \tilde{g}(\sigma_t) + \sigma_t \xi_t, \quad t = 1, 2, \dots, T,$$

so that

$$E^{\mathbb{Q}}[e^{X_t - r} | \mathcal{F}_{t-1}] = e^{-\tilde{g}(\sigma_t) + \sigma_t \xi_t} = 1.$$

Under this setting, however, the distribution of the innovations ξ_t varies over time. This makes it very difficult to implement the Markov chain method which we will see later, since we have to recompute the probability transition matrix at each time. Furthermore, it is hard to identify the dynamics of λ_t under \mathbb{P} . We can simply assume it to be a constant, but this assumption would put strong restrictions on the dynamics of the conditional mean under \mathbb{P} , that is, we would have only one free parameter in the estimation.

2.2.3. Minimal Entropy Martingale Measure

Luckily, the TS-GARCH model which is a discrete time process with stochastic volatility implies that the market is incomplete, that is, the equivalent martingale measure (EMM) is not unique. In fact, there are several methods to find the “best” EMM of tempered stable distributions. One classical method in selecting an EMM is the Esscher transform presented by Gerber and Shiu [23, 24]. Fujiwara and Miyahara [25] presented the method to find “minimal entropy martingale measure.” Cont and Tankov [22] proffered the “the least-squares calibration with a prior” to find an EMM of the market measure that minimizes the least square error of the model option prices relative to the market option price.

In this paper, we apply the minimal entropy approach which is more flexible than the previous methods. We simply assume that μ_t is modeled by the well-known ARMA(1,1) process:

$$\mu_t = c + \phi X_{t-1} + \psi \sigma_{t-1} \epsilon_{t-1}, \quad t = 1, 2, \dots, T.$$

Under the risk neutral measure \mathbb{Q} , we assume that our model is given by

$$\tilde{X}_t = \tilde{\mu}_t + \tilde{\sigma}_t \xi_t, \quad t = 1, 2, \dots, T,$$

where $\tilde{\mu}_t$ is such that the process is a discrete martingale:

$$\tilde{\mu}_t = r_t - \tilde{g}(\tilde{\sigma}_t), \quad t = 1, 2, \dots, T.$$

$\tilde{\sigma}_t$ is still a GARCH(1,1) process:

$$\tilde{\sigma}_t^2 = \tilde{\alpha}_0 + \tilde{\alpha}_1(\tilde{X}_{t-1} - \tilde{\mu}_{t-1} - \lambda)^2 + \tilde{\beta}_1 \tilde{\sigma}_{t-1}^2, \quad t = 1, 2, \dots, T,$$

where, λ is the market price of risk or the skewness parameter of GARCH, and $\xi_t, t = 1, \dots, T$ are i.i.d random variables whose distribution is in the same parametric family of ϵ_t . We denote the density function by $f(\cdot | \eta)$ where η is the vector of parameters. For example, $\eta = (\alpha, \lambda_+, \lambda_-)$ for standard CTS and $\eta = (\alpha, \theta, \beta)$ for standard NTS.

Let p and q be the joint density of the two processes, then the Kullback-Leibler divergence or relative entropy between q and p is given by

$$\begin{aligned} D_{KL}(\mathbb{Q} | \mathbb{P}) &= E^{\mathbb{Q}} \left[\log \frac{q(\tilde{X}_0, \tilde{X}_1, \dots, \tilde{X}_T)}{p(\tilde{X}_0, \tilde{X}_1, \dots, \tilde{X}_T)} \right] \\ &= E^{\mathbb{Q}} \left[\sum_{t=1}^T \log \frac{q(\tilde{X}_t | \tilde{X}_{t-1})}{p(\tilde{X}_t | \tilde{X}_{t-1})} \right] \\ &= \sum_{t=1}^T E^{\mathbb{Q}} \left[\log \frac{f((\tilde{X}_t - \tilde{\mu}_t) / \tilde{\sigma}_t | \tilde{\eta})}{f((\tilde{X}_t - \mu_t) / \sigma_t | \eta)} \right]. \end{aligned}$$

The density function f of the tempered stable distribution can be calculated via fast Fourier transform (FFT). The expectation in the equation can be computed numerically.

Finally, we calibrate the risk neutral parameters by minimizing the Kullback-Leibler divergence:

$$(\lambda, \tilde{\alpha}_0, \tilde{\alpha}_1, \tilde{\beta}_1, \tilde{\eta}) = \arg \min D_{KL}(\mathbb{Q} | \mathbb{P}).$$

3. MARKOV CHAIN METHOD UNDER THE TS-GARCH MODELS

In this section we introduce the Markov chain method proposed by Duan and Simonato [19], and implement this approach to our TS-GARCH models. And finally we will prove the convergence of Markov chain method under TS-GARCH framework.

3.1. Markov Chain Method

We start with the risk neutral GARCH model introduced in the last section:

$$X_t = r_t - g(\sigma_t) + \sigma_t \epsilon_t,$$

$$\sigma_t^2 = \alpha_0 + \alpha_1(X_{t-1} - r_t + g(\sigma_t) - \lambda)^2 + \beta_1 \sigma_{t-1}^2, \quad t = 1, 2, \dots, T.$$

Note that the expression is actually a bivariate Markovian system $\{(S_t, \sigma_{t+1}), t = 1, 2, \dots, T\}$. We denote the value of the American option at time t by $V(S_t, \sigma_{t+1}, t)$. Let $f: \mathbb{R}^+ \rightarrow \mathbb{R}$ be the payoff function, for call option we have $f(S_T) = (S_T - K)^+$ and for put option $f(S_T) = (K - S_T)^+$. Since the bivariate process $\{(S_t, \sigma_{t+1}), t = 1, 2, \dots, T\}$ is Markovian, then the value of the American option is given by invoking the following dynamic programming:

$$\begin{aligned} V(S_T, \sigma_{T+1}, T) &= f(S_T); \\ V(S_t, \sigma_{t+1}, t) &= \max(f(S_t), e^{-r} E^{\mathbb{Q}}[V(S_{t+1}, \sigma_{t+2}, t+1) | \\ &\quad S_t, \sigma_{t+1}]). \quad t = 1, 2, \dots, T. \end{aligned} \quad (2)$$

The Markov property of the bivariate process allows us to approximate the asset price and conditional volatility (S_t, σ_{t+1}) by a Markov chain.

First let us assume that under the risk neutral measure \mathbb{Q} the stock price and volatility follows a discrete Markov chain. The stock price S_t takes m different values: $s(1) < s(2) < \dots < s(m)$; and the volatility σ_{t+1} takes n different values: $\sigma(1) < \sigma(2) < \dots < \sigma(n)$. So we have overall mn status. Let Π be the $mn \times mn$ transition probability matrix for the Markov chain:

$$\Pi = \begin{pmatrix} \pi(1, 1; 1, 1) & \pi(1, 1; 2, 1) & \cdots & \pi(1, 1; m, 1) & \pi(1, 1; 1, 2) & \cdots & \pi(1, 1; m, 2) & \cdots & \pi(1, 1; m, n) \\ \pi(2, 1; 1, 1) & \pi(2, 1; 2, 1) & \cdots & \pi(2, 1; m, 1) & \pi(2, 1; 1, 2) & \cdots & \pi(2, 1; m, 2) & \cdots & \pi(2, 1; m, n) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \pi(m, 1; 1, 1) & \pi(m, 1; 2, 1) & \cdots & \pi(m, 1; m, 1) & \pi(m, 1; 1, 2) & \cdots & \pi(m, 1; m, 2) & \cdots & \pi(m, 1; m, n) \\ \pi(1, 2; 1, 1) & \pi(1, 2; 2, 1) & \cdots & \pi(1, 2; m, 1) & \pi(1, 2; 1, 2) & \cdots & \pi(1, 2; m, 2) & \cdots & \pi(1, 2; m, n) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \pi(m, n; 1, 1) & \pi(m, n; 2, 1) & \cdots & \pi(m, n; m, 1) & \pi(m, n; 1, 2) & \cdots & \pi(m, n; m, 2) & \cdots & \pi(m, n; m, n) \end{pmatrix} \quad (3)$$

where

$$\pi(i, j; k, l) = \mathbb{Q}(S_t = s(i), \sigma_{t+1} = \sigma(j) | S_{t-1} = s(k), \sigma_t = \sigma(l)).$$

Although the transition matrix is usually very large, we are able to do matrix operation efficiently since it is a highly sparse. There are two facts to ensure the sparsity of the matrix. First, $\pi(i, j; k, l)$ would be close to zero when the absolute difference between $s(i)$ and $s(k)$ is too large. Secondly, as we will discuss later, there are at most m non-zero probability for every row of Π even though there are mn entries in each row since σ_{t+1} would be completely determined by S_t if (S_{t-1}, σ_t) is given. The sparsity of the transition matrix is a very important property when implementing this method, because it can save storage and computing time. Note that the transition matrix for TS-GARCH is relatively less sparse than normal GARCH due to the heavy tails. But it is still good enough for efficient computing.

However, the distribution of (S_t, σ_{t+1}) under TS-GARCH model is continuous. To construct a Markov chain for approximating the TS-GARCH process, let $\{(c(i-1), c(i))\}_{i=1, \dots, m}, \{(d(j-1), d(j))\}_{j=1, \dots, n}$ be two sequences of small disjoint intervals such that $s(i) \in (c(i-1), c(i)]$ and $\sigma(j) \in (d(j-1), d(j)]$, the transition matrix of the approximating Markov chain is given by:

$$\pi(i, j; k, l) = \mathbb{Q}(S_t \in (c(i-1), c(i)], \sigma_{t+1} \in (d(j-1), d(j)] | S_{t-1} = s(k), \sigma_t = \sigma(l)), k = 1, \dots, m, l = 1, \dots, n.$$

So the conditional expectation in Equation (2) can be approximated as follows:

$$E^{\mathbb{Q}}[V(S_t, \sigma_{t+1}, t) | S_{t-1} = s(k), \sigma_t = \sigma(l)] \approx \sum_{i,j} V(s(i), \sigma(j), t) \pi(i, j; k, l).$$

Because σ_{t+1} is uniquely determined by S_{t-1}, S_t, σ_t , then $j \mapsto \pi(i, j; k, l)$ is either zero or equal to $P(S_t \in (c(i-1), c(i)] | S_{t-1} = s(k), \sigma_t = \sigma(l))$. In the latter case, simple calculations show that

$$\pi(i, j; k, l) = \mathbb{Q}(\tilde{c}(i-1) < \xi_t \leq \tilde{c}(i)) = F(\tilde{c}(i)) - F(\tilde{c}(i-1)),$$

where

$$\tilde{c}(i) = \frac{1}{\sigma(l)} \left(\log \frac{c(i)}{s(k)} - r + g(\sigma(l)) \right),$$

F denotes the cumulative distribution function, and g denotes the log moment generating function of the innovation ε_t under \mathbb{Q} .

Different with the normal distribution, F does not have a closed form of solution for tempered stable distributions. Similar as the density function which can be computed numerically by FFT, Kim et al. [26] show that we can use one step FFT to obtain F :

$$F(x) = \frac{e^{x\theta}}{\pi} \operatorname{Re} \left(\int_0^\infty e^{-ixu} \frac{\Phi(u + i\theta)}{\theta - ui} du \right),$$

where $\Phi(\cdot)$ is the characteristic function of ε_t and θ is a positive number such that $|\Phi(z)| < \infty$ for all z with $\operatorname{Im}(z) = \theta$. To compute F , let K and N be large positive integers with $N > K$ and $x_k = (k - N)/K, k = 0, 1, 2, \dots, 2N$, then we have

$$F(x_k) \approx \frac{K}{N} \sum_{j=0}^{N-1} \omega^{jk} f_j, \quad k = 0, 1, \dots, N - 1.$$

where $\omega = e^{-2\pi j/N}$ and $f_j = (-1)^j \Phi(2\pi jK/N + i\theta) / (\theta - i(2\pi jK/N))$. Then apply FFT to compute this equation and we can obtain the cumulative function of the tempered stable distributions.

Note that we assume that the approximation Markov chain process is time-homogenous since we want the states of stock price and conditional variance take values to be in the same set for every time t . Finally the option value can be approximated as follows:

$$\hat{V}(s(k), \sigma(l), T) = f(s(i));$$

$$\hat{V}(s(k), \sigma(l), t) = \max \left(f(s(k)), e^{-r} \sum_{i,j} V(s(i), \sigma(j), t + 1) \pi(i, j; k, l) \right).$$

3.2. Implementing the Markov Chain Approach to TS-GARCH

The basic idea of implementing the Markov chain method to the TS-GARCH model follows Duan and Simonato [19]. First note that the stock price has a trend in the GARCH process under the locally risk neutral measure \mathbb{Q} , that is, $E^{\mathbb{Q}}[\log(S_t/S_{t-1}) | \mathcal{F}_{t-1}] \neq 0$. This contradicts to our assumption that the Markov chain is homogeneous. Therefore, we need to adjust the stock price such that the Markov process based on the adjusted asset price is closer to homogeneous. For example we can set:

$$p_t = \log S_t - \nu_t,$$

where

$$v_t = E^{\mathbb{Q}}[\log S_t].$$

Thus, the mean of p_t will not change with time. Unlike the normal-GARCH in which $E^{\mathbb{Q}}[\log S_t]$ is a linear function of t if the GARCH process is stationary, the mean of log-price is not linear in TS-GARCH since the log moment generating function $g(\sigma_t)$ can not be assumed to be stationary like σ_t^2 . Furthermore, computing $E^{\mathbb{Q}}[g(\sigma_t)]$ numerically is time consuming. So instead we set

$$v_t = rt - g(\sigma_0)t.$$

Note that all we want is that $E[p_t]$ dose not vary much through time $t = 1, 2, \dots, T$; and that σ_t^2 will be in a neighborhood of σ_0 in a long term. Since the adjustment factor is deterministic, a compounding factor can be easily incorporated later to recover the pre-adjusted stock price. Then we can set the discrete values for the stock price and its conditional variance. To do this, we define an interval that covers a chosen set of discrete stock prices and variances. For the logarithm of the adjusted stock price, consider an interval around p_0 : $[p_0 - I_p, p_0 + I_p]$, Duan and Simonato [19] use the variance of $\frac{1}{t} \sum_{k=1}^t \sigma_k \varepsilon_k$ to determine I_p . Since the variance cannot fully describe the diversification of non-symmetric heavy tail distributions, here we use the quantile instead to determine the range of p . Let F be the cumulative distribution function of $\sum_{t=1}^T \varepsilon_t$, whose distribution belongs to the same family of ε_t due to the infinite divisibility. Then we set $p(1) = p_0 + I_p^-$ and $p(m) = p_0 + I_p^+$, where

$$\begin{aligned} I_p^- &= \sigma_0 F^{-1}(\delta_p(m)), \\ I_p^+ &= \sigma_0 F^{-1}(1 - \delta_p(m)), \end{aligned}$$

F^{-1} is the inverse function of F which is well defined since F is continuous for CTS and NTS. $\delta_p(m)$ is a sequence of decreasing numbers which satisfies the following conditions:

Partition condition. (i) $\delta_p(m) \rightarrow 0$ as $m \rightarrow \infty$; and (ii) $|I_p^+ - I_p^-|/m \rightarrow 0$ as $m \rightarrow \infty$. (iii) $0 < \delta_p(m) \leq 1$.

For example, if the distribution of ξ_t is α -stable for some $\alpha \in (0, 2)$, then it is natural to set $\delta_p(m) = m^{\epsilon - \alpha}$, where $0 < \epsilon < \alpha$. This is because

$$m^\epsilon = \frac{\delta_p(m)}{m^{-\alpha}} = \frac{F(I_p^-/\sigma_0)}{m^{-\alpha}} \propto C \left(\frac{|I_p^-|/\sigma_0}{m} \right)^{-\alpha} \rightarrow \infty, m \rightarrow \infty$$

where C is some positive constant. Here we use the property that the tail of α -stable distribution has a power decay, see Cont and Tankov [22], for details. This implies that $|I_p^-|/m \rightarrow 0$ as $m \rightarrow \infty$. We have similar results for I_p^+ . Thus, $\delta_p(m) = m^{\epsilon - \alpha}$ must satisfy the second condition. Furthermore, any function of m which decays to zero slower than $m^{\epsilon - \alpha}$ also satisfies the partition condition (i) and (ii). For other tempered stable distributions like CTS or NTS, we can still set $m^{\epsilon - \alpha}$ as an ‘‘upper bound’’ of decay speed of $\delta_p(m)$ since their tails are thinner than the α -stable distribution, that is, the inverse cumulative distribution function $F^{-1}(x)$ diverges slower than power law as $x \rightarrow 0$ or 1.

Similar to p_t , we can also set q_t to be the logarithm of σ_t^2 . Luckily we have the upper bound ρ and lower bound α_0 for σ_t^2 , so similar as Duan and Simonato [19] we define the upper and lower bound of q :

$$\begin{aligned} q(1) &= \log(\alpha) + \log(\tilde{\sigma}^2/\alpha)\delta_q(n), \\ q(n) &= \log(\rho) + \log(\tilde{\sigma}^2/\rho)\delta_q(n), \end{aligned}$$

where $\delta_q(n)$ is some function which satisfies (i) $\delta_q(n) \rightarrow 0$ and (ii) $\delta_q(1) = 1$, and $\tilde{\sigma}^2$ is the weighted average of the initial value σ_0^2 and the stationary value σ_{stat}^2 , given by $\sigma_{stat}^2 = \alpha_0/(1 - \alpha_1(1 + k^2) - \beta_1)$:

$$\tilde{\sigma}^2 = \frac{\tau - \min(T, \tau)}{\tau} \sigma_0^2 + \frac{\min(T, \tau)}{\tau} \sigma_{stat}^2,$$

where τ is a preset value. Note that when $n \rightarrow \infty$, $q(1)$ and $q(n)$ converge to $\log(\alpha)$ and $\log(\rho)$ respectively. When $n = 1$, which implies the constant volatility, then $q = \log(\tilde{\sigma}^2)$ would be an ideal choice.

Then we simply take the equal-distant partition of the intervals of adjusted price and volatility and obtain a sequence of discrete values $p(1) < \dots < p(m)$ and $q(1) < \dots < q(n)$. Now consider the m cells for m status of the adjusted price: $C(k) = (c(k), c(k + 1)]$ where $c(1) = -\infty$, $c(m) = +\infty$, and $c(k) = (p(k) + p(k + 1))/2$. Similarly, we have the partition $D(k) = (d(k), d(k + 1)]$ for q , where $d(1) = \log(\alpha_0)$, $d(n) = \log(\rho)$, and $d(k) = (q(k) + q(k + 1))/2$. Let

$$\begin{aligned} h(x, y, z) &:= e^{-z}(x - y + v_t - v_{t-1} - r + g(e^z)) \\ &= e^{-z}(x - y + g(e^z) - g(\sigma_0)); \end{aligned}$$

and

$$H(x, y, z) := \log(\min(\alpha_0 + \alpha_1 e^z(h(x, y, z) - \lambda)^2 + \beta_1 e^z, \rho)).$$

Then the transition probability is given by:

$$\begin{aligned} \pi_t(i, j, k, l) &= \begin{cases} \mathbb{Q}(p_t \in C(i) | p_{t-1} = p(k), q_t = q(l)) \\ \quad \text{if } H(p(i), p(k), q(l)) \in D(j) . \\ 0 & \text{otherwise} \end{cases} \quad (4) \end{aligned}$$

If we set

$$\tilde{c}_t(i; k, l) = h(c(i), p(k), q(l)),$$

then the probability in Equation (4) can be represented as:

$$\begin{aligned} \mathbb{Q}(p_t \in C(i) | p_{t-1} = p(k), q_t = q(l)) \\ = F(\tilde{c}_t(i + 1; k, l)) - F(\tilde{c}_t(i; k, l)), \end{aligned}$$

where F is the cumulative distribution function of innovation ε_t under \mathbb{Q} .

Note that the above approach might not be the best way to discretize the price and volatility. As long as the ranges of discretized price and volatility increase and the sizes of grids

decrease as m and n turns to infinity, we are able to show that the bivariate Markov chain can converge to the objective GARCH process. In the mean time we want to compute the transition matrix only once in order to make the algorithm efficient. So any partitions which meets these criteria may potentially improve the approximation of the Markov chain.

The algorithm is summarized as follows:

1. Compute the discretized adjusted prices and volatilities: $p(1), \dots, p(m)$ and $q(1), \dots, q(n)$.
2. Compute the $mn \times mn$ transition matrix Π given by Equations (3) and (4).
3. Compute the terminal value:

$$\hat{V}_T = e_n \otimes \begin{pmatrix} f(e^{p(1)+v_T}) \\ \vdots \\ f(e^{p(m)+v_T}) \end{pmatrix},$$

where $e_n \in \mathbb{R}^n$ is a vector with all ones and \otimes is the Kronecker product.

4. Given \hat{V}_t , compute

$$\hat{V}_{t-1} = \max \left(\Pi \hat{V}_t, e_n \otimes \begin{pmatrix} f(e^{p(1)+v_{t-1}}) \\ \vdots \\ f(e^{p(m)+v_{t-1}}) \end{pmatrix} \right),$$

where \max denotes the element-wise maximum for two vectors.

5. If $p(k) \leq \log S_0 \leq p(k+1)$, then the option price would be the linear interpolation:

$$V_0 = \frac{\hat{V}_0(k+1) - \hat{V}_0(k)}{p(k+1) - p(k)} (\log S_0 - p(k)) + \hat{V}_0(k).$$

3.3. Convergence of the Markov Chain Approach

Duan and Simonato [19] proves the convergence of the Markov chain method under the normal GARCH model. For TS-GARCH model, the proof is similar. There are two major differences between the two models. First, the $\sigma_t^2/2$ term in the normal GARCH is replaced by a more general form $g(\sigma_t)$. Second, the distribution of ε_t is skewed and heavy tailed in TS-GARCH. Luckily, these differences does not change the idea of the proof.

In this section we denote the adjusted prices and volatilities of the original GARCH process and the approximating Markov chain by (p_t, q_{t+1}) and $(p_t^{(m,n)}, q_{t+1}^{(m,n)})$, $t = 1, \dots, T$, respectively. The distribution function of the Markov chain at time t conditional on the information up to $t - 1$ is given by $F_{t|t-1}^{(m,n)} : \mathbb{R} \times [\log(\alpha_0), \log(\rho)] \rightarrow [0, 1]$; and the distribution of the object adjusted stock price and volatility is $F_{t|t-1} : \mathbb{R} \times [\log(\alpha_0), \log(\rho)] \rightarrow [0, 1]$.

Proposition 1. *Given the partition condition, $F_{t|t-1}^{(m,n)}$ converges to $F_{t|t-1}$ as m and n go to infinity.*

Proof. First note that the marginal distribution of p_t :

$$\mathbb{Q}(p_t \leq x | p_{t-1} = y, q_t = z) = F(h(x, y, z)),$$

where F is the cumulative distribution function of CTS or NTS, is uniformly continuous with respect to $x, y \in \mathbb{R}$ and $z \in [\log(\alpha_0), \log(\rho)]$. This is because that the distribution function F is uniformly continuous on \mathbb{R} , and that $h(x, y, z) = e^{-z}(x - y + g(e^z) - g(\sigma_0))$ is linear with respect to x and y , and $g(e^z)$ is continuous on the closed set $[\log(\alpha_0), \log(\rho)]$. Note that the moment generate functions of CTS and NTS are not defined on ρ , but they have a finite continuous extension on ρ .

Let i^*, k^*, l^* be the integers such that $p(i^*) \leq x < p(i^* + 1)$, $p(l^*) \leq y < p(l^* + 1)$ and $q(k^*) \leq z < q(k^* + 1)$, then the distribution of $p_t^{(m,n)}$:

$$\begin{aligned} \mathbb{Q}(p_t^{(m,n)} \leq x | p_{t-1}^{(m,n)} = p(k^*), \\ q_t^{(m,n)} = q(l^*)) &= \sum_{i=1}^m \sum_{j=1}^n \pi(i, j, k^*, l^*) I_{\{p(i) \leq x\}} \\ &= \sum_{i=1}^m \sum_{j=1}^n \mathbb{Q}(p_t \in C(i) | p_{t-1} = p(k^*), q_t = q(l^*)) \\ &\quad I_{\{H(p(i), p(k^*), q(l^*)) \in D(j)\}} I_{\{p(i) \leq x\}} \\ &= \sum_{i=1}^m \mathbb{Q}(p_t \in C(i) | p_{t-1} = p(k^*), q_t = q(l^*)) I_{\{p(i) \leq x\}} \\ &= \mathbb{Q}(p_t \leq c(i^* + 1) | p_{t-1} = p(k^*), q_t = q(l^*)) \rightarrow \\ &\quad \mathbb{Q}(p_t \leq x | p_{t-1} = y, q_t = z), \end{aligned}$$

as $m, n \rightarrow \infty$. The convergence is uniform, i.e., the rate of convergence dose not depends on the choices of x, y , and z because the uniform continuity.

Now consider the joint distribution of (p_t, q_{t+1}) conditional on $p_{t-1} = y$ and $q_t = z$:

$$\begin{aligned} F_{t|t-1}(x_1, x_2 | y, z) &= \mathbb{Q}(p_t \leq x_1, q_{t+1} \leq x_2 | p_{t-1} = y, q_t = z) \\ &= \mathbb{Q}(p_t \leq x_1, H(p_t, p_{t-1}, q_t) \leq x_2 | p_{t-1} = y, \\ &\quad q_t = z). \end{aligned}$$

where $\alpha_0 \leq x_2 \leq \rho$. On the other side the joint distribution of the Markov chain conditional on $p_{t-1}^{(m,n)} = p(k^*)$ and $q_t^{(m,n)} = q(l^*)$ is given by

$$\begin{aligned} F_{t|t-1}^{(m,n)}(x_1, x_2 | p(k^*), q(l^*)) &= \mathbb{Q}(p_t^{(m,n)} \leq x_1, q_{t+1}^{(m,n)} \leq x_2 | p_{t-1}^{(m,n)} \\ &= p(k^*), q_t^{(m,n)} = q(l^*)) \\ &= \sum_{i=1}^m \sum_{j=1}^n \pi_t(i, j, k^*, l^*) I_{\{q(j) \leq x_2\}} \end{aligned}$$

Let i^*, j^* be the integers such that $p(i^*) \leq x_1 \leq p(i^* + 1)$ and $q(j^*) \leq x_2 \leq q(j^* + 1)$, then

$$\begin{aligned} F_{t|t-1}^{(m,n)}(x_1, x_2|p(k^*), q(l^*)) &= \sum_{i=1}^{i^*} \sum_{j=1}^{j^*} \pi_t(i, j, k^*, l^*) \\ &= \sum_{i=1}^m \sum_{j=1}^n \pi_t(i, j, k^*, l^*) I_{\{H(p(i), p(k^*), q(l^*)) \leq d(j^* + 1)\}} \\ &\quad I_{\{p(i) \leq c(i^* + 1)\}} \\ &= \mathbb{Q}(p_t^{(m,n)} \leq c(i^* + 1), H(p_t^{(m,n)}, p(k^*), q(l^*))) \\ &\leq d(j^* + 1) |p_{t-1}^{(m,n)} = p(k^*), q_t^{(m,n)} = q(l^*)). \end{aligned}$$

Since the distribution of $p_t^{(m,n)}$ conditioning on $t - 1$ converges uniformly to p_t and function H is also uniformly continuous, it is not difficult to show that $F_{t|t-1}^{(m,n)}(x_1, x_2|p(k^*), q(l^*)) \rightarrow F_{t|t-1}(x_1, x_2|y, z)$ uniformly as $m, n \rightarrow \infty$.

From the proposition and the property of weak convergence we can conclude:

Corollary 1. For any bounded continuous function $f : \mathbb{R} \times [\log(\alpha_0), \log(\rho)] \rightarrow \mathbb{R}$,

$$\begin{aligned} E[f(p_t^{(m,n)}, q_{t+1}^{(m,n)})|p_{t-1}^{(m,n)} = p(k^*), q_t^{(m,n)} = q(l^*)] \\ \rightarrow E[f(p_t, q_{t+1})|p_{t-1} = y, q_t = z] \end{aligned}$$

as $m, n \rightarrow \infty$, where k^*, l^* are the integers such that $p(k^*) \leq y < p(k^* + 1)$ and $q(l^*) \leq z < q(l^* + 1)$. The convergence is uniform.

Corollary 2. Let $f^{(m,n)} : \mathbb{R} \times [\log(\alpha_0), \log(\rho)] \rightarrow \mathbb{R}$ be a sequence of bounded continuous function which converges to a bounded continuous function $f : \mathbb{R} \times [\log(\alpha_0), \log(\rho)] \rightarrow \mathbb{R}$ uniformly, then

$$\begin{aligned} E[f^{(m,n)}(p_t^{(m,n)}, q_{t+1}^{(m,n)})|p_{t-1}^{(m,n)} = p(k^*), q_t^{(m,n)} = q(l^*)] \\ \rightarrow E[f(p_t, q_{t+1})|p_{t-1} = y, q_t = z] \end{aligned}$$

as $m, n \rightarrow \infty$, where k^*, l^* are the integers such that $p(k^*) \leq y < p(k^* + 1)$ and $q(l^*) \leq z < q(l^* + 1)$. The convergence is uniform.

Proof. Since $f^{(m,n)}$ converges to f uniformly, for any $\epsilon > 0$ we have

$$\begin{aligned} |E[f^{(m,n)}(p_t^{(m,n)}, q_{t+1}^{(m,n)})|p_{t-1}^{(m,n)} = p(k^*), q_t^{(m,n)} = q(l^*)] \\ - E[f(p_t^{(m,n)}, q_{t+1}^{(m,n)})|p_{t-1}^{(m,n)} = p(k^*), q_t^{(m,n)} = q(l^*)]| \leq \frac{\epsilon}{2} \end{aligned}$$

for all choice of $p(k^*)$ and $q(l^*)$ if m, n is large enough. On the other side corollary 1 ensures that

$$\begin{aligned} |E[f(p_t^{(m,n)}, q_{t+1}^{(m,n)})|p_{t-1}^{(m,n)} = p(k^*), q_t^{(m,n)} = q(l^*)] \\ - E[f(p_t, q_{t+1})|p_{t-1} = y, q_t = z]| \leq \frac{\epsilon}{2} \end{aligned}$$

for all y and z . Combining the two inequalities we prove the the corollary.

Now consider the American type option with bounded continuous payoff f . The Markov chain method first compute the conditional expectation:

$$\begin{aligned} G_{T-1}^{(m,n)}(p(k), q(l)) &:= E[f(p_T^{(m,n)})|p_{T-1}^{(m,n)} = p(k), q_T^{(m,n)} = q(l)], \\ &k = 1, \dots, m, l = 1, \dots, n \end{aligned}$$

Note that this is a sequence of discrete values. But we are able to construct a continuous extension $\tilde{G}_{T-1}^{(m,n)} : \mathbb{R} \times [\log(\alpha_0), \log(\rho)] \rightarrow \mathbb{R}$ such that $\tilde{G}_{T-1}^{(m,n)}(p(k), q(l)) = G_{T-1}^{(m,n)}(p(k), q(l))$, $k = 1, \dots, m, l = 1, \dots, n$, for each m and n . Then we obtain a sequence of continuous bounded function $\tilde{G}_{T-1}^{(m,n)}$ which converges to the true expectation:

$$G_{T-1}(y, z) := E[f(p_T)|p_{T-1} = y, q_T = z]$$

uniformly. Here we assume that G_{T-1} is continuous. This implies that the value of the American option given by Markov chain method:

$$V_{T-1}^{(m,n)}(p(k), q(l)) = \max\{G_{T-1}^{(m,n)}(p(k), q(l)), f(\exp(v_t + p(k)))\}$$

has a continuous extension $\tilde{V}_{T-1}^{(m,n)}$ converges to the true value V_{T-1} uniformly for continuous and bounded payoff function f . The corollary 2 allows us to continue doing this and have $\tilde{V}_0^{(m,n)}$ converges to V_0 .

However, for call and put options whose payoffs are not bounded we may not have these results. But we are able to use options truncated payoff to approximate call and put, thanks for the dominated convergence theorem. Duan and Simonato [19] loosen the condition by showing if the payoff function is bounded by a integrable function then the European option price computed by Markov chain method converges to the true price. Here we consider the bounded payoff for simplicity. In fact, by discretizing the stock price into a finite set of numbers the Markov chain method assumes that the payoff function is bounded. And by increasing m and n this upper bound will become larger and therefore the value of bounded payoff option would be closer to the value of unbound one.

4. NUMERICAL EXAMPLES AND EMPIRICAL RESULTS

In this section we apply the Markov chain method to compute European and American options under the CTS-GARCH and NTS-GARCH models, respectively. The convergence of the Markov chain method will be illustrated by numerical examples and empirical results.

4.1. Numerical Examples

In this subsection we test the Markov chain method numerically by applying it to pricing American option as well as European option. For simplicity, we set the initial conditional variance to be the stationary variance under the market measure \mathbb{P} ; that is $\sigma_0^2 = \alpha_0(1 - \alpha_1 - \beta_1)^{-1}$. The initial stock price is set as $S_0 = 50$, and the annualized risk-free rate is $r = 0.025$.

TABLE 1 | American put options in the CTS-GARCH framework.

K/S_0	Maturity = 30			Maturity = 60			Maturity = 90		
	0.90	1.00	1.10	0.90	1.00	1.10	0.90	1.00	1.10
$m = 150, n = 50$	0.0029	0.7340	5.0684	0.0296	1.0106	5.0982	0.0868	1.2969	5.3111
$m = 180, n = 60$	0.0031	0.7402	5.0612	0.0292	1.0097	5.0881	0.0827	1.2750	5.2704
$m = 210, n = 70$	0.0026	0.7310	5.0560	0.0324	1.0602	5.1959	0.0824	1.2635	5.2407
$m = 240, n = 80$	0.0027	0.7224	5.0520	0.0300	1.0403	5.1773	0.0772	1.2386	5.2179
$m = 270, n = 90$	0.0026	0.7236	5.0489	0.0290	1.0312	5.1627	0.0798	1.2764	5.3124
$m = 300, n = 100$	0.0029	0.7499	5.1056	0.0295	1.0298	5.1508	0.0808	1.2723	5.2877
$m = 250, n = 50$	0.0029	0.7359	5.0741	0.0323	1.0619	5.2049	0.0823	1.2654	5.2517
$m = 300, n = 60$	0.0032	0.7586	5.1056	0.0311	1.0394	5.1508	0.0827	1.2795	5.2877
$m = 350, n = 70$	0.0030	0.7411	5.0775	0.0315	1.0512	5.1838	0.0854	1.2953	5.3147
$m = 400, n = 80$	0.0028	0.7458	5.1015	0.0305	1.0530	5.2094	0.0783	1.2539	5.2572
$m = 450, n = 90$	0.0028	0.7380	5.0797	0.0289	1.0332	5.1722	0.0773	1.2587	5.2824
$m = 500, n = 100$	0.0029	0.7471	5.0993	0.0304	1.0485	5.1944	0.0810	1.2774	5.3031

TABLE 2 | European put options in the CTS-GARCH framework.

K/S_0	Maturity = 30			Maturity = 60			Maturity = 90		
	0.90	1.00	1.10	0.90	1.00	1.10	0.90	1.00	1.10
$m = 150, n = 50$	0.0029	0.7156	4.8301	0.0290	0.9689	4.6796	0.0842	1.2240	4.7533
$m = 180, n = 60$	0.0030	0.7218	4.8244	0.0287	0.9672	4.6715	0.0802	1.2018	4.7153
$m = 210, n = 70$	0.0029	0.7129	4.8184	0.0317	1.0158	4.7734	0.0799	1.1929	4.6918
$m = 240, n = 80$	0.0026	0.7043	4.8136	0.0294	0.9960	4.7516	0.0749	1.1680	4.6633
$m = 270, n = 90$	0.0026	0.7050	4.8107	0.0285	0.9865	4.7368	0.0773	1.2012	4.7430
$m = 300, n = 100$	0.0029	0.7307	4.8670	0.0289	0.9861	4.7276	0.0784	1.1990	4.7260
$m = 250, n = 50$	0.0029	0.7175	4.8360	0.0317	1.0175	4.7813	0.0799	1.1948	4.7003
$m = 300, n = 60$	0.0031	0.7393	4.8681	0.0304	0.9956	4.7313	0.0802	1.2058	4.7299
$m = 350, n = 70$	0.0030	0.7226	4.8398	0.0309	1.0072	4.7613	0.0828	1.2220	4.7566
$m = 400, n = 80$	0.0028	0.7268	4.8627	0.0299	1.0079	4.7814	0.0759	1.1819	4.6971
$m = 450, n = 90$	0.0027	0.7189	4.8413	0.0283	0.9882	4.7450	0.0749	1.1845	4.7148
$m = 500, n = 100$	0.0029	0.7281	4.8608	0.0298	1.0036	4.7681	0.0786	1.2036	4.7389
Monte Carlo	0.0028	0.7196	4.8506	0.0269	0.9790	4.7510	0.0710	1.1666	4.7099
Monte Carlo errors	0.0001	0.0028	0.0064	0.0006	0.0043	0.0088	0.0015	0.0065	0.0118

The parameters of the GARCH process are set to be $\alpha_0 = 10^{-5}$, $\alpha_1 = 0.10$, $\beta_1 = 0.80$. These are the same parameters that Duan and Simonato [19] used in the numerical tests. The facts that $\alpha_0 \approx 0$ and $\alpha_1 < \beta_1$ correspond to most empirical tests in the literature. We set $\alpha = 1.70$, $\lambda_+ = 1.00$, $\lambda_- = 0.35$ to be the parameters of CTS innovation and $\alpha = 1.00$, $\theta = 1.00$, $\beta = -0.20$ to be the parameters of NTS innovation, so that both distributions are skewed to the left and have heavier tails than normal distribution. These parameters are also chosen based on the empirical observations in Rachev et al. [20].

Tables 1, 3 show the results of the Markov chain approach for American put options. The first column of each table represents the number of m, n . We set $m > n$ in order to get a better approximation of the price process when the size of the Markov chain mn is given. As m, n increases, the Markov chain method would converge to the true result which is proved in Section 3. Columns 2–4, 5–7, 8–10 are the prices

of the options with maturity 30, 60, and 90 days respectively. For each maturity we compute the options with strikes $0.9S_0$, $1.0S_0$, and $1.1S_0$, the corresponding columns are marked by the second row. Tables 2, 4 compares the results of the Markov chain approach with the ones of the Monte Carlo with least square approximation. For Monte Carlo method we generate 100,000 scenarios, and the results are given by the last row in Tables 2, 4.

From these results one can observe that the results of Markov chain method with large m and n and the ones of Monte Carlo correspond roughly in two decimal place. Most of the prices in Tables 1, 3 is greater than the corresponding prices in Tables 2, 4, since the American option is always more expensive than the European option. The convergence of American options is not as fast as the one of European options. This because that small differences in accuracy might change the optimal execution time which would make the price different. We do not compare the

TABLE 3 | American put options in the NTS-GARCH framework.

K/S_0	Maturity = 30			Maturity = 60			Maturity = 90		
	0.90	1.00	1.10	0.90	1.00	1.10	0.90	1.00	1.10
$m = 150, n = 50$	0.0029	0.7340	5.0684	0.0296	1.0106	5.0982	0.0868	1.2969	5.3111
$m = 180, n = 60$	0.0031	0.7402	5.0612	0.0292	1.0097	5.0881	0.0827	1.2750	5.2704
$m = 210, n = 70$	0.0029	0.7310	5.0560	0.0324	1.0602	5.1959	0.0824	1.2635	5.2407
$m = 240, n = 80$	0.0027	0.7224	5.0520	0.0300	1.0403	5.1773	0.0772	1.2386	5.2179
$m = 270, n = 90$	0.0026	0.7236	5.0489	0.0290	1.0312	5.1627	0.0798	1.2764	5.3124
$m = 300, n = 100$	0.0029	0.7499	5.1056	0.0295	1.0298	5.1508	0.0808	1.2723	5.2877
$m = 250, n = 50$	0.0029	0.7359	5.0741	0.0323	1.0619	5.2049	0.0823	1.2654	5.2517
$m = 300, n = 60$	0.0032	0.7586	5.1056	0.0311	1.0394	5.1508	0.0827	1.2795	5.2877
$m = 350, n = 70$	0.0030	0.7411	5.0775	0.0315	1.0512	5.1838	0.0854	1.2953	5.3147
$m = 400, n = 80$	0.0028	0.7458	5.1015	0.0305	1.0530	5.2094	0.0783	1.2539	5.2572
$m = 450, n = 90$	0.0028	0.7380	5.0797	0.0289	1.0332	5.1722	0.0773	1.2587	5.2824
$m = 500, n = 100$	0.0029	0.7471	5.0993	0.0304	1.0485	5.1944	0.0810	1.2774	5.3031

TABLE 4 | European put options in the NTS-GARCH framework.

K/S_0	Maturity = 30			Maturity = 60			Maturity = 90		
	0.90	1.00	1.10	0.90	1.00	1.10	0.90	1.00	1.10
$m = 150, n = 50$	0.0029	0.7156	4.8301	0.0290	0.9689	4.6796	0.0842	1.2240	4.7533
$m = 180, n = 60$	0.0030	0.7218	4.8244	0.0287	0.9672	4.6715	0.0802	1.2018	4.7153
$m = 210, n = 70$	0.0029	0.7129	4.8184	0.0317	1.0158	4.7734	0.0799	1.1929	4.6918
$m = 240, n = 80$	0.0026	0.7043	4.8136	0.0294	0.9960	4.7516	0.0749	1.1680	4.6633
$m = 270, n = 90$	0.0026	0.7050	4.8107	0.0285	0.9865	4.7368	0.0773	1.2012	4.7430
$m = 300, n = 100$	0.0029	0.7307	4.8670	0.0289	0.9861	4.7276	0.0784	1.1990	4.7260
$m = 250, n = 50$	0.0029	0.7175	4.8360	0.0317	1.0175	4.7813	0.0799	1.1948	4.7003
$m = 300, n = 60$	0.0031	0.7393	4.8681	0.0304	0.9956	4.7313	0.0802	1.2058	4.7299
$m = 350, n = 70$	0.0030	0.7226	4.8398	0.0309	1.0072	4.7613	0.0828	1.2220	4.7566
$m = 400, n = 80$	0.0028	0.7268	4.8627	0.0299	1.0079	4.7814	0.0759	1.1819	4.6971
$m = 450, n = 90$	0.0027	0.7189	4.8413	0.0283	0.9882	4.7450	0.0749	1.1845	4.7148
$m = 500, n = 100$	0.0029	0.7281	4.8608	0.0298	1.0036	4.7681	0.0786	1.2036	4.7389
Monte Carlo	0.0028	0.7196	4.8506	0.0269	0.9790	4.7510	0.0710	1.1666	4.7099
Monte Carlo error	0.0001	0.0028	0.0064	0.0006	0.0043	0.0088	0.0015	0.0065	0.0118

TABLE 5 | ARMA(1,1)-GARCH(1,1) parameters.

c	ϕ (AR)	ψ (MA)	a_0	a_1	b_1
1.9314×10^{-4}	0.8020	-0.8664	2.6359×10^{-6}	0.8743	0.1053

TABLE 6 | Standard CTS parameters.

α	λ_+	λ_-
1.8123	0.2019	0.1228

American option price with the well known least squares method because the latter is not accurate when the option is out of money. Even for the options which are in the money, the polynomial approximation is only accurate in a small range of stock prices. We would expect more extreme values in GARCH model with heavy tailed distributions. This might make the least square approach more inaccurate.

The fluctuations in the prices are the results of discretizing the stock price and volatility. By comparing the results with the Duan's Normal GARCH model one can find that we need larger m and n to get an acceptable result. This is because we need more grids in a wider interval to approximate the

cumulative distribution functions of heavy tailed distributions. In addition, the rate of convergence become slow when the maturity increases. As Duan and Simonato [19] suggested in his paper, the Markov chain method might not be good for options with long maturity.

4.2. Empirical Analysis

In this section, we illustrate the calibration of risk neutral parameters using Markov chain method. As an example, we consider the SPX put option on June 13, 2013 whose moneyness is between 0.9 and 1.1, and whose maturity is 30 days. There are

TABLE 7 | Standard NTS parameters.

α	θ	β
0.5819	1.5550	-0.2188

TABLE 8 | Minimal entropy parameters (CTS).

λ	α	λ_+	λ_-
0.0121	1.7914	0.3467	0.1122

TABLE 9 | Minimal entropy parameters (NTS).

λ	α	θ	β
0.0261	0.8741	1.1859	-0.2875

TABLE 10 | Errors of minimal entropy measure.

	AAE	APE	ARPE	RMSE
CTS	2.0680	0.1013	0.1621	7.9570
NTS	2.5041	0.1226	0.2127	10.9452

TABLE 11 | Calibrated parameters (CTS).

λ	α	λ_+	λ_-
-0.0005	1.6001	0.14126	0.0892

TABLE 12 | Calibrated parameters (NTS).

λ	α	θ	β
-0.0002	0.4683	1.8390	0.3729

TABLE 13 | Errors of minimal entropy measure.

AAE	APE	ARPE	RMSE	
CTS	1.2077	0.0591	0.1213	3.6134
NTS	1.8629	0.0912	0.3116	5.0308

46 options meet these conditions, the strike price of which ranges from 1480 to 1740. The initial price of S&P 500 index is 1640.42.

The first step is to estimate the parameters under market measure \mathbb{P} . The maximum likelihood estimation (MLE) of ARMA(1,1)-GARCH(1,1) with Gaussian innovation is widely known. For non-Gaussian innovations, we can use the quasi maximum likelihood estimator (QMLE). That is, we first estimate the ARMA(1,1)-GARCH(1,1) parameters under Gaussian assumption. Then we will get a sequence of residuals which are assumed to be i.i.d. Finally we use other distributions like CTS or NTS to fit these i.i.d. residuals. The convergence of QMLE is studied by Lee and Hansen [27] and Lumsdaine [28]. We select 2 years log daily returns of S&P 500 index as

our sample. The data is downloaded from Yahoo finance. The estimated parameters are given by **Tables 5–7**.

Then we can compute the risk neutral parameters in two ways. The first way is to minimize the relative entropy between \mathbb{P} and \mathbb{Q} , as described in Section 2.3. The relative entropy is given by Equation (10), which is computed numerically. Here we assume that the GARCH parameters do not change under \mathbb{Q} . This assumption is widely used in most GARCH option pricing papers, see Duan [8], Stentoft [29], Kim et al. [6] and Rachev et al. [20]. The results of the minimization are shown in **Tables 8, 9**.

Using these parameters we compute the option price via Markov chain method. We compute the average absolute error (AAE), average percentage error (APE), average relative percentage error (ARPE) and root mean square error (RMSE), which are give by **Table 10**.

The second way is to calibrate the risk neutral measure using realized option price, that is, to find the measure which minimizes the errors between model prices and market prices. In this paper we use RMSE to measure the error. The calibrated parameters and their errors are given by **Tables 11–13**. One can observe that the calibrated risk neutral parameters are different from the minimal entropy parameters.

We find that in both risk neutral measures CTS outperform NTS. However, it is not enough to conclude that CTS is better than NTS based on this specific example. In fact, NTS is usually good in the high dimensional case since it is a subordinated distribution and has a natural multivariate extension. It is also not enough to conclude that minimal-entropy measure is not the real risk neutral measure. The minimal entropy approach is based on the assumption that the risk neutral world is very close to the real one. The calibrated parameters might overfit the model if the option is not volatile. Here we use this method because it provides a convenient way to consider risk neutral model under Markov chain framework. Selecting the best measure in the incomplete market is out of the scope of this paper.

5. CONCLUSION

In this paper we employ the Markov chain approximation method to compute the value of American-style options under TS-GARCH model. We apply the minimal entropy approach to find the “best” equivalent martingale measure in an incomplete market. This risk neutral measure also allows us to compute the Markov chain method in efficiently. For heavy-tailed distributions, we use the quantile instead of the variance of the distribution to determine the range of the discrete Markov chain. We also consider the rate of tail decay to find the step size of the discretized stock prices. The convergence of the Markov chain approximation for TS-GARCH models can be proved due to the absolute continuity of TS-distribution functions. The Markov chain method is a good alternative for the least square approximation, since the latter one is less accurate when the option is out of the money and when extreme event happens. Furthermore, although the Markov chain approach for TS-GARCH is slower than the one for normal GARCH, TS-GARCH model performs much better than normal GARCH in option pricing since it has more degrees for freedom to capture the heavy-tails in financial markets.

REFERENCES

- Mandelbrot BB. New methods in statistical economics. *J Pol Econ.* (1963) 71:421–40. doi: 10.1086/258792
- Koponen I. Analytic approach to the problem of convergence of truncated lévy flights towards the gaussian stochastic process. *Phys Rev E.* (1995) 52:1197–9. doi: 10.1103/PhysRevE.52.1197
- Boyarchenko SI, Levendorskiĭ SZ. Option pricing for truncated lévy processes. *Int J Theor Appl Finance.* (2000) 3:549–52. doi: 10.1142/S021902490000541
- Carr P, Geman H, Madan D, Yor M. The fine structure of asset returns: an empirical investigation. *J Bus.* (2002) 75:305–32. doi: 10.1086/338705
- Barndorff-Nielsen OE, Levendorskii S. Feller processes of normal inverse gaussian type. *Quant Finance.* (2001) 1:318–31. doi: 10.1088/1469-7688/1/3/303
- Kim YS, Rachev ST, Chung DM, Bianchi ML. The modified tempered stable distribution, GARCH-models and option pricing. *Probab Math Stat.* (2009) 29:91–117.
- Kim YS, Rachev ST, Bianchi ML, Fabozzi FJ. Financial market models with lévy processes and time-varying volatility. *J Bank Finance.* (2008) 32:1363–78. doi: 10.1016/j.jbankfin.2007.11.004
- Duan JC. The GARCH option pricing model. *Math Finance.* (1995) 5:13–32. doi: 10.1111/j.1467-9965.1995.tb00099.x
- Duan JC. Augmented GARCH (p,q) process and its diffusion limit. *J Econ.* (1997) 79:97–127. doi: 10.1016/S0304-4076(97)00009-2
- Kallsen J, Taqqu MS. Option Pricing in ARCH-type models. *Math Finance.* (1998) 8:13–26. doi: 10.1111/1467-9965.00042
- Heston L, Nandi S. A closed-form GARCH option valuation model. *Rev Finacial Stud.* (2000) 13:585–625. doi: 10.1093/rfs/13.3.585
- Christoffersen P, Jacobs K, Ornathanalai C. GARCH option valuation: theory and evidence. *J Derivat.* (2013) 21:8–41. doi: 10.2139/ssrn.2054859
- Menn C, Rachev ST. A GARCH option pricing model with α -stable innovations. *Eur J Oper Res.* (2005) 163:201–09. doi: 10.1016/j.ejor.2004.01.009
- Menn C, Rachev ST. Smoothly truncated stable distributions, GARCH-models, and option pricing. *Math Methods Oper Res.* (2009) 69:411–38. doi: 10.1007/s00186-008-0245-6
- Stentoft L. American option pricing using GARCH models and the normal inverse Gaussian distribution. *J Financ. Econom.* (2008) 6:540–82. doi: 10.1093/jffinec/nbn013
- Kim YS, Rachev ST, Bianchi ML, Fabozzi FJ. Tempered stable and tempered infinitely divisible GARCH models. *J Bank Finance* (2010) 34:2096–109. doi: 10.1016/j.jbankfin.2010.01.015
- Longstaff FA, Schwartz ES. Valuing American options by simulation: a simple least-squares approach. *Rev Financ Stud.* (2001) 14:113–48. doi: 10.1093/rfs/14.1.113
- Ritchken P, Trevor R. Pricing options under generalized GARCH and stochastic volatility processes. *J Finance.* (1999) 54:377–402. doi: 10.1111/0022-1082.00109
- Duan JC, Simonato JG. American option pricing under GARCH by a Markov chain approximation. *J Econ Dyn Control.* (2001) 25:1689–718. doi: 10.1016/S0165-1889(00)00003-8
- Rachev ST, Kim YS, Bianchi ML, Fabozzi FJ. *Financial Model with Levy Processes and Volatility Clustering.* Hoboken, NJ: John Wiley & Sons (2011).
- Bouchard B, Warin, X. Monte-Carlo valuation of American options: facts and new algorithms to improve existing methods. In: Carmona RA, Del Moral P, Hu P, Oudjane N, editors. *Numerical Methods in Finance.* Berlin; Heidelberg: Springer (2012). p. 215–255.
- Cont R, Tankov P. *Financial Modelling with Jump Processes.* London: Chapman & Hall / CRC (2004).
- Gerber FA, Shiu ESW. Option pricing by Esscher transforms. *Trans Soc Actuar.* (1994) 46:99–191.
- Gerber FA, Shiu ESW. Actuarial bridges to dynamic hedging and option pricing. *Insur Math Econ.* (1996) 18:183–531. doi: 10.1016/0167-6687(96)85007-4
- Fujiwara T, Miyahara Y. The minimal entropy martingale measures for geometric lévy processes. *Finance Stoch.* (2003) 7:509–31. doi: 10.1007/s007800200097
- Kim YS, Rachev S, Bianchi ML, Fabozzi FJ. Computing VaR and AVaR in infinitely divisible distributions. *Probab Math Stat.* (2009) 30:223–45. doi: 10.2139/ssrn.1400965
- Lee SW, Hansen BE. Asymptotic theory for the GARCH (1, 1) quasi-maximum likelihood estimator. *Econom. Theory* (1994) 10:29–52. doi: 10.1017/S0266466600008215
- Lumsdaine RL. Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH (1, 1) and covariance stationary GARCH (1, 1) models. *Econometrica* (1996) 64:575–596. doi: 10.2307/2171862
- Stentoft L. Pricing American options when the underlying asset follows GARCH processes. *J Empir Finance* (2005) 12:576–611. doi: 10.1016/j.jempfin.2004.08.001

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Shi, Zhang and Kim. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.