# Willow tree algorithms for pricing VIX derivatives under stochastic volatility models

Changfu Ma[1], Wei Xu[2†] and Yue Kuen Kwok[3]

[1,2] School of Mathematical Sciences, Tongji University,
Shanghai 200092, China

[3] Department of Mathematics, Hong Kong University of Science and Technology,
Hong Kong, China

## Abstract

VIX futures and options are the most popular contracts traded in the Chicago Board Options Exchange. The bid-ask spreads of traded VIX derivatives remain to be wide, possibly due to lack of reliable pricing models. In this paper, we consider pricing VIX derivatives under the consistent model approach, which considers joint modeling of the dynamics of the S&P index and its instantaneous variance. Under the affine jump-diffusion formulation with stochastic volatility, analytic integral formulas can be derived to price VIX futures and options. However, these integral formulas invariably involve Fourier inversion integrals with cumbersome hyper-geometric functions, thus posing various challenges in numerical evaluation. We propose a unified numerical approach based on the willow tree algorithms to price VIX derivatives under various common types of joint process of the S&P index and its instantaneous variance. Given the analytic form of the characteristic function of the instantaneous variance of the S&P index process in the Fourier domain, we apply the fast Fourier transform algorithm to obtain the transition density function numerically in the real domain. We then construct the willow tree that approximates the dynamics of the instantaneous variance process up to the fourth order moment. Our comprehensive numerical tests performed on the willow tree algorithms demonstrate high level of numerical accuracy, runtime efficiency and reliability for pricing VIX futures and both European and American options under the affine model and 3/2-model. We also examine the implied volatility smirks and the term structures of the implied skewness of VIX options.

Keywords: VIX derivatives, willow tree algorithm, affine jump-diffusion model, 3/2-model.

1

# 1 Introduction

It is well accepted that volatility of the price process of a financial asset or index is a latent (hidden) stochastic process that is not directly observable. Market volatility of the S&P index is a crucial determinant of investment decisions. The volatility index (with the ticker symbol VIX) of the Chicago Board Options Exchange (CBOE) provides the model free measure of the volatility implied by the traded market prices of out-of-the-money options on the S&P 500 index (SPX) over 30 calendar days. Formally, VIX is the square root of the risk neutral expectation of the integrated variance of SPX over the next 30 calendar days, calculated on an annualized basis and expressed in percentage point. Investors use VIX as fear gauge (indicator of market confidence) in the financial markets since VIX tends to stay high after large downward moves of the SPX value. The volatility index VIX and market index SPX are commonly observed to be negatively correlated with each other. An interesting account on the history of the construction of VIX can be found in Carr and Wu (2006). In particular, they discuss the rationale for the replacement of the old volatility index (now called VXO) by the new VIX (launched since 2003). The old volatility index is model dependent since its calculation is based on the Black-Scholes pricing model, which was first computed in 1993 as a linear combination of 8 at-the-money implied volatilities on the S&P 100 options with maturities closest to 30 calendar days. The new VIX is the model free volatility measure of SPX, which would ease the pricing procedures of VIX futures and option.

Investors may directly invest in volatility as an asset class via VIX derivatives, like VIX futures and options. Market practitioners use VIX derivatives to hedge the risks of investment in the S&P 500 index and/or achieve exposure to the S&P 500 volatility without having to delta hedge their S&P 500 option positions. Trading in VIX futures was started in 2004 while that of VIX options was started in 2006. The contract multiplier for each VIX futures contract is $1,000 while that of each VIX option contract is $100. The popularity of trading on VIX futures and options has been growing over the years. On June 24, 2016, in reaction to the Brexit referendum, over 721,000 VIX futures contracts were traded. On November 9, 2016, the trading volume was 644,892 in reaction to the surprise outcome of the US presidential election. For trading of VIX options at the CBOE, the reported record of 2,562,477 contracts were traded on August 10, 2017. Gonzalez-Perez (2015) presents a comprehensive review on the successes and shortcomings of the use of VIX and derivatives on VIX in financial markets as market risk measure, financial products to hedge against volatility risk, and volatility measure to estimate spot volatility dynamics and risk premium.

There are two popular model approaches for pricing VIX derivatives. In the consistent model approach, one considers the stochastic dynamics of SPX under stochastic volatility and derives the dynamics of VIX. The other approach directly models the dynamics of VIX (Psychoyios *et al.*, 2010; Goard and Mazur, 2013; Park, 2016; Li *et al.*, 2017; Detemple and Kitapbayev, 2018). Mencía and Sentana (2013) present an extensive empirical analysis of various VIX derivative valuation models under direct modeling of the dynamics of VIX. Duan and Yeh (2010) discuss the extraction of the jump and volatility risk premiums implied by VIX. Kaeck and Alexander (2013) study the continuous time VIX dynamics and examine the role of stochastic volatility of volatility. Note that VIX is derived from the prices of options on SPX, so the direct modeling approach may ignore the linkage between SPX and VIX.

In this paper, we adopt the consistent model approach that involves modeling of the joint process of SPX and its instantaneous variance. There have been numerous earlier works on pricing VIX derivatives using the consistent model approach. Zhang and Zhu (2006) first consider pricing VIX futures under the Heston stochastic volatility model. Later, Zhu and Zhang (2007) propose the no-arbitrage approach for pricing VIX futures using the term structure of forward variance. Cont and Kokholm (2013) emphasize the importance of adding jumps in both SPX and volatility. They argue that adding jumps in volatility is important in order to produce the positive skew of implied volatilities of VIX options. Under the assumption of the affine stochastic volatility model with simultaneous jumps, Luo and Zhang (2012) analyze the term structure of VIX. Under the assumption of zero jump, they show that the square of VIX can be expressed as the weighted average of the instantaneous variance process and its long term mean. In general, nice analytic tractability exists under the affine model with simultaneous jumps on the index process and its instantaneous process, and various forms of price formulas of VIX futures and options in integral forms can be established (Lin, 2007; Sepp, 2008; Zhu and Lian, 2012; Lian and Zhu, 2013). Wang *et al.* (2017) also manage to derive the integral price formula of VIX futures under the Heston-Nandi GARCH model. To facilitate numerical valuation of VIX derivatives under the affine models, Barletta and Nicolato (2018) derive the orthogonal expansion series for the price functions of VIX derivatives. Kaeck and Alexander (2012) examine the volatility dynamics of SPX under 18 different affine and non-affine stochastic volatility models with jumps on both SPX and its instantaneous variance. Their studies provide clear evidence that the non-affine models out-perform their affine counterparts. Branger *et al.* (2017) use the informational content of VIX derivatives to infer implications on the non-affine modeling of the SPX returns' variance dynamics. They find that

both the non-affine diffusion and jump dynamics are required to capture both the short-term and long-term implied volatility distribution. Baldeaux and Badran (2014) manage to obtain integral representation of the price formulas for VIX futures and options under the 3/2-model with jumps on the index. Based on the 4/2 stochastic volatility model proposed by Grasselli (2017), Lin *et al.* (2017) derive the integral price formulas for VIX derivatives under the 4/2 stochastic volatility plus jump on the index. Later, Lin *et al.* (2018) derive the price formulas of VIX derivatives under their free stochastic volatility model, characterized by an arbitrary power parameter of the instantaneous variance in the dynamic equation of SPX and its instantaneous variance. Their class of models reduce to the 1/2-model (affine model) and 3/2-model when the power parameter assumes the value of 1/2 and 3/2, respectively.

Though integral price formulas can be derived for VIX derivatives under the affine type consistent models, numerical valuation of these price formulas can be quite cumbersome. Even for the numerical implementation of the affine Heston model, Barletta and Nicolato (2018) argue that numerical calculations are not free of complications. Numerical instabilities may arise due to failure of integrability conditions and inappropriate choice of the branch cuts in the software evaluation of multivalued complex logarithm functions. Kwok and Zheng (2018) propose the use of the saddlepoint approximation methods to evaluate the integral price formulas of VIX derivatives under the Heston model. However, numerical accuracy of the saddlepoint approximation may deteriorate when the jump sizes are significant, options are deep-out-of-the-money and at times close to expiration. Therefore, it is desirable to develop reliable and accurate numerical schemes for pricing VIX derivatives under popular types of consistent models.

In this paper, we propose a unified numerical pricing method based on the willow tree algorithm that can be used to price VIX derivatives under common types of joint process of the index and its instantaneous variance. The willow tree algorithm was first proposed by Curran (2001), which has been successfully applied to price various path dependent options and structured derivatives (Xu *et al.*, 2013; Xu and Yin, 2014; Wang and Xu, 2018; Dong *et al.*, 2019). In these research papers, the willow tree method has been demonstrated to be an effective numerical approach for pricing exotic financial derivatives in terms of numerical accuracy, runtime efficiency and reliability. Unlike the finite difference schemes where uniform spacing of layers of nodes is adopted, the first four order moments of the underlying instantaneous variance process are used to determine the layers of nodes in the willow tree. The construction of the willow tree is in the real domain and requires the knowledge of the probability distribution of the underlying instantaneous variance

process. Since only the characteristic function of the underlying instantaneous variance process in the Fourier domain is known in analytic form, our proposed willow tree algorithm includes the fast Fourier transform algorithm to find the first four order moments of the underlying instantaneous variance process. With availability of analytic form of the corresponding characteristic function of the underlying instantaneous variance process and coupled with the fast Fourier-cosine series algorithm (Fang and Oosterlee, 2008), the willow tree approach can be extended to price VIX products and other exotic derivatives under general Lévy processes and stochastic volatility models.

This paper is organized as follows. In the next section, we discuss the model formulation of VIX under two choices of stochastic volatility models: (i) affine jump-diffusion model with simultaneous jumps on both the index and its instantaneous variance, and (ii) 3/2-model with jumps on the index. We show how $\text{VIX}_t^2$ can be expressed in terms of the instantaneous variance $v_t$ and other relevant model parameters under these two models. In Section 3, we present the construction of the willow tree algorithm that price VIX futures and options under the affine model and 3/2-model. We show how to compute the transition probabilities between the willow tree nodes at successive time steps based on the conditional distribution of the instantaneous variance using the fast Fourier cosine algorithm. In Section 4, we present the numerical tests that were performed to assess accuracy, efficiency and reliability of the willow tree algorithms for pricing VIX derivatives. With availability of the effective willow tree pricing algorithm, we examine various pricing behaviors of VIX futures and options, like the implied volatility smirks and the term structure of the implied volatilities of VIX options.

## 2 Model formulation of VIX derivatives

Let $\tau$ denote the time window of the volatility measure of the index value process $S_t$, where $\tau$ is fixed at 30 calendar days by the CBOE. Following Zhang and Zhu (2006), VIX is defined via the relation:

$$\text{VIX}_t^2(\tau) = -\frac{2}{\tau}\mathbb{E}_t^Q[\ln \frac{S_{t+\tau}}{S_t e^{r\tau}}] \times 100^2, \tag{2.1}$$

where $r$ is the risk free interest rate and $\mathbb{E}_t^Q$ is the expectation under a risk neutral measure $Q$ conditional on the filtration $\mathcal{F}_t$. The consistent model specifies the joint dynamics of the index value process $S_t$ and its instantaneous variance process $v_t$. Baldeaux and Badran (2014) argue that the stochastic volatility model for $S_t$ without jumps may not produce the implied volatilities of VIX options with volatility skew structures that are consistent with market observed implied skew

structures inferred from traded VIX options.

## 2.1 Affine jump-diffusion model

For the affine jump-diffusion model, the joint process of $S_t$ and $v_t$ under a risk neutral measure $Q$ is specified by the dynamic equations:

$$\frac{\mathrm{d}S_t}{S_t} = (r - \lambda\bar{\mu}) \,\mathrm{d}t + \sqrt{v_t} \,\mathrm{d}W_t^1 + (e^{J^S} - 1) \,\mathrm{d}N_t \tag{2.2a}$$

$$\mathrm{d}v_t = \eta(\theta - v_t) \,\mathrm{d}t + \sigma_v\sqrt{v_t} \,\mathrm{d}W_t^2 + J^v \,\mathrm{d}N_t, \tag{2.2b}$$

where the Brownian motions $W_t^1$ and $W_t^2$ observe $\mathrm{d}W_t^1\mathrm{d}W_t^2 = \rho \,\mathrm{d}t$. Here, $\rho$ is the constant correlation coefficient between $\mathrm{d}W_t^1$ and $\mathrm{d}W_t^2$, $\theta$ is the constant mean reversion level of $v_t$, $\sigma_v$ is the constant volatility of $v_t$ and $\eta$ is the constant multiplier on the mean reversion drift rate. We assume simultaneous jumps on $S_t$ and $v_t$ and they are modeled by the common Poisson process $N_t$ with constant intensity $\lambda$. The assumed form of the dynamics of $v_t$ with the choice of $\sqrt{v_t}$ in the diffusion term exhibits nice analytic tractability of the conditional characteristic function due to its affine structure (Heston, 1993). It is commonly called the 1/2-model due to the square root term $\sqrt{v_t}$. In our later discussion, we use the simplified term "1/2-model" for this affine jump-diffusion model.

To enhance analytic tractability, it is common to adopt the following assumption on the jump distribution of $S_t$ and $v_t$. Let $J^S$ and $J^v$ denote the respective random jump component on $S_t$ and $v_t$, where $J^S$ and $J^v$ are independent of $N_t$ and both random jump components are correlated with correlation coefficient $\rho_J$. Furthermore, we assume $J^v$ to be exponentially distributed with mean $\mu_v$, where

$$J^v \sim \exp(\mu_v); \tag{2.3a}$$

and $J^S|J^v$ is normally distributed with mean $\mu_S + \rho_J J^v$ and variance $\sigma_S{}^2$, where

$$J^S|J^v \sim N(\mu_S + \rho_J J^v, \sigma_S{}^2). \tag{2.3b}$$

It is seen that

$$\bar{\mu} = \mathbb{E}[e^{J^S} - 1] = \frac{e^{\frac{\sigma_S{}^2}{2} + \mu_S}}{1 - \rho_J\mu_v} - 1,$$

and $\rho_J$ and $\mu_v$ are chosen to observe the technical condition $\rho_J\mu_v < 1$.

Based on the consistent model of the joint process of $S_t$ and $v_t$ under the 1/2-model, we manage to express $\mathrm{VIX}_t^2(\tau)$ in terms of the instantaneous variance $v_t$ and other relevant model parameters.

6

This property is highly desirable in pricing VIX derivatives, the proof of which is outlined below. Firstly, we rewrite (2.2a) in terms of $\ln S_t$, where

$$\mathrm{d}\ln S_t = (r - \lambda\bar{\mu} - \frac{v_t}{2})\,\mathrm{d}t + \sqrt{v_t}\,\mathrm{d}W_t^1 + J^S\,\mathrm{d}N_t.$$

Integrating the dynamic equation from $t$ to $t+\tau$ and substituting the above equation into (2.1), we obtain

$$\begin{aligned}\mathrm{VIX}_t^2(\tau) &= \frac{2}{\tau}\mathbb{E}_t^Q[\int_t^{t+\tau}(\frac{v_u}{2} + \lambda\bar{\mu})\,\mathrm{d}u - J^S\,\mathrm{d}N_u] \times 100^2 \\ &= \left\{\frac{1}{\tau}\mathbb{E}_t^Q[\int_t^{t+\tau} v_u\,\mathrm{d}u] + 2\lambda[\bar{\mu} - (\mu_S + \rho_J\mu_v)]\right\} \times 100^2.\end{aligned}$$

By solving the dynamic equation for $v_t$ in (2.2b), we obtain

$$v_u = e^{-\eta(u-t)}v_t + \theta[1 - e^{-\eta(u-t)}] + \int_t^u \sigma_v e^{-\eta(u-s)}\sqrt{v_s}\,\mathrm{d}W_s^2 + \int_t^u e^{-\eta(u-s)}J^v\,\mathrm{d}N_s, \quad u > t.$$

Taking expectation $\mathbb{E}_t^Q$ on both sides of the above equation, we have

$$\begin{aligned}\mathbb{E}_t^Q[v_u] &= e^{-\eta(u-t)}v_t + \theta[1 - e^{-\eta(u-t)}] + \int_t^u e^{-\eta(u-s)}\mathbb{E}^Q[J^v]\mathbb{E}^Q[\mathrm{d}N_s] \\ &= e^{-\eta(u-t)}v_t + \theta[1 - e^{-\eta(u-t)}] + \frac{\lambda\mu_v}{\eta}[1 - e^{-\eta(u-t)}], \quad u > t.\end{aligned}$$

We then integrate $\mathbb{E}_t^Q[v_u]$ over the time interval $[t, t+\tau]$ and substitute into the equation for $\mathrm{VIX}_t^2(\tau)$ to give

$$\begin{aligned}\mathrm{VIX}_t^2(\tau) &= \left\{\frac{1}{\tau}\int_t^{t+\tau}\mathbb{E}_t^Q[v_u]\,\mathrm{d}u + 2\lambda[\bar{\mu} - (\mu_S + \rho_J\mu_v)]\right\} \times 100^2 \\ &= \left\{\frac{1-e^{-\eta\tau}}{\eta\tau}v_t + (1 - \frac{1-e^{-\eta\tau}}{\eta\tau})(\theta + \frac{\lambda\mu_v}{\eta}) + 2\lambda[\bar{\mu} - (\mu_S + \rho_J\mu_v)]\right\} \times 100^2.\end{aligned}$$

In summary, under the affine 1/2-model, $\mathrm{VIX}_t^2(\tau)$ can be expressed as a linear function in $v_t$ in the form

$$\mathrm{VIX}_t^2(\tau) = [a_0(\tau) + a_1(\tau)v_t] \times 100^2, \tag{2.4}$$

where the $\tau$-dependent parameters are given by

$$\begin{aligned}a_0(\tau) &= (1 - \frac{1-e^{-\eta\tau}}{\eta\tau})(\theta + \frac{\lambda\mu_v}{\eta}) + 2\lambda[\bar{\mu} - (\mu_S + \rho_J\mu_v)], \\ a_1(\tau) &= \frac{1-e^{-\eta\tau}}{\eta\tau}.\end{aligned}$$

Under the case of zero jump, where $\lambda = 0$, $\mathrm{VIX}_t^2(\tau)$ is seen to be the weighted average of the instantaneous variance $v_t$ and its long term mean $\theta$ (Luo and Zhang, 2012). The weight function $a_0(\tau)$ depends on $\eta$ and $\tau$ only. Indeed, $\mathrm{VIX}_t^2(\tau)$ is independent of the parameters that characterize the dynamics of the index process $S_t$.

## 2.2   3/2-model

Based on the empirical studies on the time series data of S&P 100 implied volatilities as a proxy for the instantaneous variance, Bakshi *et al.* (2006) show that nonlinear drift is favored over linear drift and the variance exponent should be approximately 1.3 in the instantaneous variance dynamics. Responding to these empirical studies, Baldeaux and Badran (2014) and Detemple and Kitapbayev (2018) consider pricing of VIX derivatives under the consistent 3/2-model.

Following the formulation of the consistent 3/2-model in earlier works ( Baldeaux and Badran, 2014; Detemple and Kitapbayev, 2018), the dynamics of $S_t$ assumes the same jump-diffusion dynamics as in (2.2a) while the dynamics of $v_t$ takes the following diffusion form without jumps as follows:

$$\mathrm{d}v_t = \eta v_t(\theta - v_t)\,\mathrm{d}t + \sigma_v v_t^{\frac{3}{2}}\,\mathrm{d}W_t^2. \tag{2.5a}$$

The parameters $\eta$, $\theta$ and $\sigma_v$ have similar interpretation as those of the 1/2-model. We emphasize that $S_t$ remains to exhibit jumps while $v_t$ does not have jump. Thanks to the reciprocal transformation $\hat{v}_t = \frac{1}{v_t}$, the dynamic equation of $\hat{v}_t$ is governed by

$$\mathrm{d}\hat{v}_t = \hat{\eta}(\hat{\theta} - \hat{v}_t)\,\mathrm{d}t + \hat{\sigma}_v \hat{v}_t^{\frac{1}{2}}\,\mathrm{d}W_t^2, \tag{2.5b}$$

where $\hat{\eta} = \eta\theta$, $\hat{\theta} = \frac{\eta + \sigma_v^2}{\eta\theta}$ and $\hat{\sigma}_v = -\sigma_v$. The 3/2-power in $v_t$ becomes 1/2-power in $\hat{v}_t$. Working through a similar derivation, $\mathrm{VIX}_t^2(\tau)$ under the 3/2-model can be expressed as

$$\mathrm{VIX}_t^2(\tau) = \left\{\frac{1}{\tau}\int_t^{t+\tau} \mathbb{E}_t^Q[\frac{1}{\hat{v}_u}]\,\mathrm{d}u + 2\lambda(\bar{\mu} - \mu_S)\right\} \times 100^2. \tag{2.6}$$

Unlike the affine 1/2-model, one cannot obtain analytic formula for $\mathbb{E}_t^Q[\frac{1}{\hat{v}_u}]$, so there is no simple relation between $\mathrm{VIX}_t^2(\tau)$ and $v_t$ under the 3/2-model. The integral price formulas for VIX derivatives under the 3/2-model become highly cumbersome, so numerical valuation of these formulas require very daunting tasks and numerical implementation becomes unreliable. In Section 3, we show that an effective willow tree algorithm can be constructed, which requires minimal modification in the algorithmic design when we move from the 1/2-model to 3/2-model. This is because numerical valuation of $\mathbb{E}_t^Q[\frac{1}{\hat{v}_u}]$ can be performed effectively under the framework of the willow tree algorithm.

## 2.3   VIX futures and options

The exchange traded VIX has fixed $\tau$ at the value of 30 calendar days. For notational convenience, we drop $\tau$ in $\mathrm{VIX}_t(\tau)$ in our subsequent discussion if there is no ambiguity on $\tau$. For VIX futures

maturing at $T$, the time-$t$ futures price is given by

$$F_t = \mathbb{E}_t^Q[\text{VIX}_T]. \tag{2.7}$$

Similarly, the time-$t$ prices of the European and American VIX call options maturing at $T$ with strike price $K$ are given by

$$c_t = e^{-r(T-t)}\mathbb{E}_t^Q[(\text{VIX}_T - K)^+], \tag{2.8a}$$

$$C_t = \sup_{u \in [t,T]} \mathbb{E}_t^Q[e^{-r(u-t)}(\text{VIX}_u - K)^+], \tag{2.8b}$$

respectively.

# 3 Pricing VIX derivatives by the willow tree algorithm

In this section, we discuss how to construct the willow tree structure based on the approximation of the distribution of the instantaneous variance process $v_t$ up to the fourth order moment. These order moments of $v_t$ are computed by the fast Fourier inversion algorithm based on the knowledge of the conditional characteristic function of $v_t$. We then employ the usual backward induction procedure in standard lattice tree calculations to price VIX derivatives with their specified terminal payoffs using the willow tree. To incorporate the American early exercise feature, we employ the usual dynamic programming procedure of taking the maximum among the intrinsic value and continuation value at each node. For the 1/2-model, we take the advantage that $\text{VIX}_t^2$ has the nice analytic representation in terms of $v_t$ [see (2.4)]. However, for the 3/2-model, an additional numerical integration procedure is required to find $\text{VIX}_t$ from the willow tree of the instantaneous variance due to the presence of the term $\int_t^{t+\tau} \mathbb{E}_t^Q[\frac{1}{\tilde{v}_u}]\mathrm{d}u$ in (2.6).

## 3.1 Willow tree of the instantaneous variance

The willow tree construction employs the discrete Markov chain to approximate the continuous stochastic process. Figure 1 illustrates the typical willow tree structure with 3 time steps and 4 discrete values of the instantaneous variance at each time step, where each node can jump to one of the four nodes in the next time step. The main challenge is the determination of the transition probability between two nodes across successive time steps based on the dynamics of the instantaneous variance.

We divide the time interval $[0, T]$ into $N$ subintervals with uniform time step $\Delta t = T/N$ and write $t_n = n\Delta t$, $n = 0, 1, \cdots, N$. With regard to the willow tree structure, we assume that there
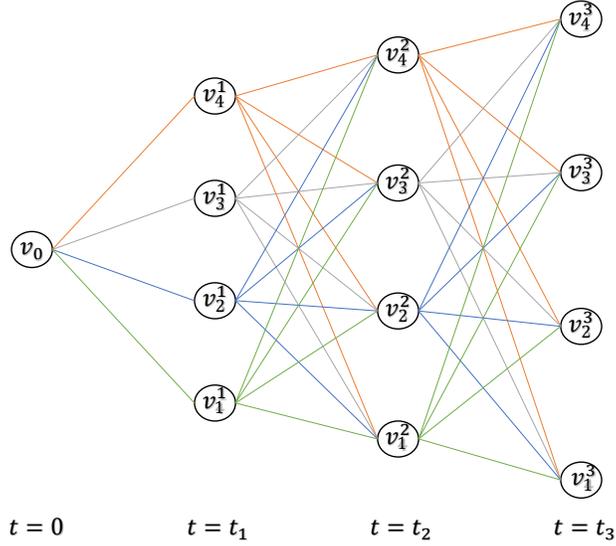
Figure 1: A pictorial representation of the willow tree structure with 3 time steps and 4 discrete values of the instantaneous variance at each time step.

are $m$ nodes at each time step. At time $t_n$, we employ $m$ discrete values of the instantaneous variance $v_i^n$, $i = 1, 2, \cdots, m$, to approximate the distribution of $v_{t_n}$ with initial value $v_0$ at $t = 0$. According to the Johnson curve transform (Johnson, 1949), the distribution of $v_{t_n}$ conditional on $v_0$ can be transformed into a standard normal distribution by an appropriate choice of the Johnson transform function $G$. More details on the Johnson curve transform can be found in Appendix A. We sample $m$ points $Z_i$, $i = 1, 2, \cdots, m$, to approximate the standard normal distribution $N(0,1)$. More details on the sampling method can be found in Xu *et al.* (2013). Conversely, $v_i^n$ can be obtained by the inverse Johnson curve transform according to the sample points $Z_i$. Using the inverse Johnson curve transform, we have

$$v_{t_n} = a_3 + a_4 G^{-1}\left(\frac{Z - a_1}{a_2}\right), \tag{3.1a}$$

where $Z \sim N(0,1)$. The constant parameters $a_1$, $a_2$, $a_3$, $a_4$ and function $G^{-1}(\cdot)$ can be determined from the first four moments of $v_{t_n}$ (Hill and Holder, 1976). Accordingly, the discrete values of the instantaneous variance are given by

$$v_i^n = a_3 + a_4 G^{-1}\left(\frac{Z_i - a_1}{a_2}\right), \quad \text{for } i = 1, 2, \cdots, m. \tag{3.1b}$$

10

Firstly, we present the construction of the willow tree under the affine dynamics of $v_t$ as depicted in (2.2b). The procedure takes advantage of availability of the analytic formula of the conditional characteristic function of $v_t$ (Heston, 1993). Let $f(\phi; t, \tau, v_t)$ denote the conditional moment generating function of the instantaneous variance $v_{t+\tau}$ conditional on $v_t$ under the 1/2-model. Lian and Zhu (2013) manage to obtain the following analytic exponential affine form:

$$f(\phi; t, \tau, v_t) = \mathbb{E}_t^Q[e^{\phi v_{t+\tau}}] = e^{E(\phi, \tau) + D(\phi, \tau) v_t + \Gamma(\phi, \tau)}, \tag{3.2}$$

where $\phi$ is the complex-valued variable and

$$\Gamma(\phi, \tau) = \frac{2\mu_v \lambda}{2\mu_v \eta - \sigma_v^2} \ln\left(1 + \frac{(\sigma_v^2 - 2\mu_v \eta)\phi}{2\eta(1 - \mu_v \phi)}(e^{-\eta\tau} - 1)\right),$$
$$E(\phi, \tau) = -\frac{2\eta\theta}{\sigma_v^2} \ln\left(1 + \frac{\sigma_v^2 \phi}{2\eta}(e^{-\eta\tau} - 1)\right),$$
$$D(\phi, \tau) = \frac{2\eta\phi}{\sigma_v^2 \phi + (2\eta - \sigma_v^2 \phi)e^{\eta\tau}}.$$

We set $\phi = iw$, then $f(iw; 0, t, v_0)$ becomes the characteristic function of $v_t$ conditional on $v_0$. By taking the inverse Fourier transform, the probability density function of $v_t$ conditional on $v_0$ is given by

$$g(x|v_0) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iwx} f(iw; 0, t, v_0) \, \mathrm{d}w. \tag{3.3}$$

Using the fast Fourier cosine series algorithm (Fang and Oosterlee, 2008), together with the proper choice of the finite integration interval $[A, B]$, $g(x|v_0)$ can be approximated by

$$g(x|v_0) \approx {\sum_{k=0}^{N}}' F_k \cos\left(k\frac{x - A}{B - A}\pi\right), \quad \text{for } x \in [A, B], \tag{3.4}$$

where

$$F_k = \frac{2}{B - A}\text{Re}\left\{f(i\frac{k\pi}{B - A}; 0, t, v_0)\exp(-i\frac{kA\pi}{B - A})\right\}.$$

Here, $\sum'$ indicates that the first term in the summation is weighted by 1/2 and $\text{Re}\{\bullet\}$ denotes taking the real part of the argument.

The $j^{\text{th}}$ order moment of the instantaneous variance $v_t$ can be computed by

$$\mathbb{E}_0^Q[(v_t)^j] = \int_0^{\infty} x^j g(x|v_0) \, \mathrm{d}x, \tag{3.5a}$$

which can be approximated by

$$\begin{aligned}
\mathbb{E}_0^Q[(v_t)^j] &\approx {\sum_{k=0}^{N}}' F_k \int_A^B x^j \cos\left(k\frac{x - A}{B - A}\pi\right) \mathrm{d}x, \\
&= {\sum_{k=0}^{N}}' \frac{2}{B - A}\text{Re}\left\{f(i\frac{k\pi}{B - A}; 0, t, v_0)\exp(-i\frac{kA\pi}{B - A})\right\}\varphi_j(k),
\end{aligned} \tag{3.5b}$$

where

$$\varphi_j(k) = \int_A^B x^j \cos\left(k\frac{x - A}{B - A}\pi\right) \mathrm{d}x.$$

11

For $j = 1, 2, 3, 4$, the explicit expressions of these integrals are given by

$$\varphi_1(k) = \begin{cases} (\frac{B-A}{k\pi})^2[(-1)^k - 1], & \text{for } k \neq 0, \\ \frac{B^2 - A^2}{2}, & \text{for } k = 0; \end{cases}$$

$$\varphi_2(k) = \begin{cases} 2(\frac{B-A}{k\pi})^2[B(-1)^k - A], & \text{for } k \neq 0, \\ \frac{B^3 - A^3}{3}, & \text{for } k = 0; \end{cases}$$

$$\varphi_3(k) = \begin{cases} 3(\frac{B-A}{k\pi})^2 \left\{ B^2(-1)^k - A^2 - 2(\frac{B-A}{k\pi})^2[(-1)^k - 1] \right\}, & \text{for } k \neq 0, \\ \frac{B^4 - A^4}{4}, & \text{for } k = 0; \end{cases}$$

$$\varphi_4(k) = \begin{cases} 4(\frac{B-A}{k\pi})^2 \left\{ B^3(-1)^k - A^3 - 6(\frac{B-A}{k\pi})^2[B(-1)^k - A] \right\}, & \text{for } k \neq 0, \\ \frac{B^5 - A^5}{5}, & \text{for } k = 0. \end{cases}$$

According to the inverse Johnson curve transform, the values of nodes in the willow tree of the instantaneous variance $v_t$ conditional on the filtration $\mathcal{F}_0$ can be obtained by the following algorithm:

**Algorithm 1**

1. *Select $m$ sample points $Z_i$, $i = 1, 2, \cdots, m$, as the approximation of the standard normal distribution.*

2. *According to (3.5b), calculate the approximate values of the first four moments of $v_{t_n}$. The parameters $a_1$, $a_2$, $a_3$, $a_4$ and function $G^{-1}(\cdot)$ using the Johnson curve transform are then determined.*

3. *Use the inverse Johnson curve transform [see (3.1b)] to obtain*

$$v_i^n = a_3 + a_4 G^{-1}(\frac{Z_i - a_1}{a_2}).$$

4. *Steps 2 and 3 are repeated to obtain $v_i^n$, $i = 1, 2, \cdots, m$ and $n = 1, 2, \cdots, N$.*

Once the node values on the willow tree structure of the instantaneous variance have been determined, we then compute the transition probability between any two nodes across consecutive time points in the willow tree. According to (3.3), given $v_{t_n} = v_i^n$, the conditional probability density function of $v_{t_{n+1}}$ is given by

$$g(x|v_i^n) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iwx} f(iw; t_n, \Delta t, v_i^n) \, dw.$$

Using the fast Fourier cosine series algorithm [ see (3.4)], $g(x|v_i^n)$ can be approximated by

$$g(x|v_i^n) \approx \sum_{k=0}^{\prime N} F_k \cos(k \frac{x - A}{B - A}\pi), \quad x \in [A, B],$$

where the discrete Fourier cosine series coefficients are given by

$$F_k = \frac{2}{B-A}\text{Re}\left\{f(i\frac{k\pi}{B-A}; t_n, \Delta t, v_i^n)\exp(-i\frac{kA\pi}{B-A})\right\}.$$

The corresponding conditional cumulative distribution function over the interval $[C, D]$ is given by

$$
\begin{aligned}
P[C < v_{t_{n+1}} \le D|v_i^n] &= \int_C^D g(x|v_i^n)\,\mathrm{d}x \\
&\approx {\sum_{k=0}^{N}}' F_k \int_C^D \cos(k\frac{x-A}{B-A}\pi)\,\mathrm{d}x \\
&= {\sum_{k=0}^{N}}' \frac{2}{B-A}\text{Re}\left\{f(i\frac{k\pi}{B-A}; t_n, \Delta t, v_i^n)\exp(-i\frac{kA\pi}{B-A})\right\}\varphi_0(k),
\end{aligned}
\tag{3.6}
$$

where

$$\varphi_0(k) = \int_C^D \cos(k\frac{x-A}{B-A}\pi)\,\mathrm{d}x = \begin{cases} \frac{B-A}{k\pi}[\sin(k\pi\frac{C-A}{B-A}) - \sin(k\pi\frac{D-A}{B-A})], & \text{for } k \neq 0, \\ D - C, & \text{for } k = 0. \end{cases}$$

Next, the transition probability between any two nodes across two consecutive time points in the willow tree can be computed systematically by the following numerical procedures.

**Algorithm 2**

1. *Let $\boldsymbol{E}^{n+1}$ denote the $(m+1)$-dimension column vector $(E_1^{n+1}, E_2^{n+1}, \cdots, E_{m+1}^{n+1})^T$, where*

$$E_i^{n+1} = \begin{cases} 0, & \text{for } i = 1, \\ \frac{v_{i-1}^{n+1} + v_i^{n+1}}{2}, & \text{for } i = 2, 3, \cdots, m, \\ \infty, & \text{for } i = m+1. \end{cases}$$

*The transition probability from node $v_i^n$ at $t_n$ to $v_j^{n+1}$ at $t_{n+1}$ is given by*

$$p_{ij}^n = P[E_j^{n+1} < v_{t_{n+1}} \le E_{j+1}^{n+1}|v_i^n], \quad \text{for } i, j = 1, 2, \cdots, m, \tag{3.7}$$

*where $P[E_j^{n+1} < v_{t_{n+1}} \le E_{j+1}^{n+1}|v_i^n]$ is computed using (3.6).*

2. *By repeating step 1 for each $t_n$, $n = 1, 2, \cdots, N$, we obtain the transition probability matrix $\boldsymbol{P}^n = [p_{ij}^n]_{m\times m}$ of the willow tree for the instantaneous variance at $t_n$.*

3. *At the initial time $t_0$, the transition probability matrix $\boldsymbol{P}^0$ is reduced to the row vector $[p_j^0]_{1\times m}$, where*

$$p_j^0 = P[E_j^1 < v_{t_1} \le E_{j+1}^1|v_0], \quad \text{for } j = 1, 2, \cdots, m.$$

As a remark, when the instantaneous variance process has no jump, Algorithm 2 can be simplified since the first four order moments of the instantaneous variance and the transition matrices of the

willow tree of the instantaneous variance can be evaluated directly [similar to the order moments calculation performed for the CIR process of the short rate in Wang and Xu (2018)] without the necessity of finding numerical approximation values using the Fourier cosine series expansion.

Next, we show how to relate the corresponding dynamics of VIX to the willow tree of instantaneous variance for 1/2- and 3/2-models in order to price the VIX derivatives.

## 3.2 Pricing VIX derivatives under the 1/2-model

Once the willow tree structure of the instantaneous variance has been constructed, the corresponding value of $\text{VIX}_t$ whose underlying SPX follows the 1/2-model as depicted in (2.2) can be computed by [see (2.4)]

$$\text{VIX}_i^n = \sqrt{a_0(\tau) + a_1(\tau)v_i^n} \times 100, \qquad \text{for } i = 1, 2, \cdots, m \quad \text{and} \quad n = 1, 2, \cdots, N. \qquad (3.8)$$

The transition probability from node $\text{VIX}_i^n$ at $t_n$ to node $\text{VIX}_j^{n+1}$ at $t_{n+1}$ is $p_{ij}^n$ [see (3.7)], which is the same as that of the instantaneous variance. In other words, the willow tree of the VIX for the affine model (1/2-model) is almost the same as the one for instantaneous variance, except for updating the tree node values of VIX using (3.8).

The prices of VIX futures and the European VIX call option with strike price $K$ can be computed by the standard backward induction procedure based on the available willow tree structure of VIX. The algorithm is summarized as follows.

**Algorithm 3**

1. *Given $\text{VIX}_{t_N} = \text{VIX}_i^N$ on the maturity date $T$, the terminal value of the VIX futures is*

$$F_i^N = \text{VIX}_i^N,$$

   *and the terminal value of the European VIX call option is*

$$c_i^N = \max(\text{VIX}_i^N - K, 0).$$

2. *The usual backward induction procedure is adopted. At $t_n$, $n = N-1, N-2, \cdots, 1$, the numerical value of VIX futures $F_i^n$ is computed by the expectation of VIX futures prices $F_j^{n+1}$ at $t_{n+1}$, where*

$$F_i^n = \sum_{j=1}^{m} p_{ij}^n F_j^{n+1}.$$

14

For the European VIX call option, its numerical value $c_i^n$ is computed by the discounted expectation of the European VIX call option numerical values $c_j^{n+1}$ at $t_{n+1}$, where

$$c_i^n = e^{-r\Delta t} \sum_{j=1}^{m} p_{ij}^n c_j^{n+1}.$$

3. At the initial time $t_0$, the numerical value of VIX futures and the European VIX call option are computed by

$$F_0 = \sum_{j=1}^{m} p_j^0 F_j^1,$$

and

$$c_0 = e^{-r\Delta t} \sum_{j=1}^{m} p_j^0 c_j^1,$$

respectively.

For the American VIX call option, the early exercise feature is incorporated by following the usual dynamic programming procedure of taking the maximum of the intrinsic value and continuation value at each node. It suffices to modify step 2 in Algorithm 3 by the dynamic programming procedure as follows:

$$C_i^n = \max\{e^{-r\Delta t} \sum_{j=1}^{m} p_{ij}^n C_j^{n+1}, (\text{VIX}_i^n - K)^+\}.$$

## 3.3 Pricing VIX derivatives under the 3/2-model

When the index value dynamics follows the 3/2-model, $\text{VIX}_t^2(\tau)$ cannot be expressed explicitly in terms of $v_t$ [see (2.6)]. As a result, numerical integration procedure is required to estimate the conditional expectation $\int_t^{t+\tau} \mathbb{E}_t^Q[\frac{1}{\hat{v}_u}]\,\mathrm{d}u$ based on the willow tree of $\hat{v}_t$. Unlike the 1/2-model, we construct the willow tree for the dynamics of the reciprocal of the instantaneous variance $\hat{v}_t$ over the extended time interval $[0,\ T+\tau]$ for the computation of the conditional expectation. The extended time interval $[0,\ T+\tau]$ is divided into $N_1$ uniform subintervals with $t_0 = 0 < t_1 < t_2 < \cdots < t_N = T < t_{N+1} < \cdots < t_{N_1} = T + \tau$, $t_n = n\Delta t$, where $N_1 = \frac{T+\tau}{\Delta t}$ and $l = \frac{\tau}{\Delta t}$. The nodes on the willow tree at $t_n$ are denoted by $\hat{v}_i^n$, $i = 1, 2, \cdots, m$, while the transition probability matrix between $t_n$ and $t_{n+1}$ on the willow tree is denoted by $\mathbf{P}^n = [p_{ij}^n]$, $i, j = 1, 2, \cdots, m$. We compute VIX at time $t_n$ as follows:

$$\begin{aligned}
\text{VIX}_{t_n}^2 &= \{\tfrac{1}{\tau} \int_{t_n}^{t_n+\tau} \mathbb{E}_{t_n}^Q[\tfrac{1}{\hat{v}_u}]\,\mathrm{d}u + 2\lambda(\bar{\mu} - \mu_S)\} \times 100^2, \\
&\approx \{\tfrac{1}{\tau} \sum_{k=n}^{l+n} \mathbb{E}_{t_n}^Q[\tfrac{1}{2}(\tfrac{1}{\hat{v}_{t_k}} + \tfrac{1}{\hat{v}_{t_{k+1}}})]\Delta t + 2\lambda(\bar{\mu} - \mu_S)\} \times 100^2, \quad\quad (3.9) \\
&= \{\tfrac{1}{\tau}(\tfrac{1}{2}\tfrac{1}{\hat{v}_{t_n}} + \sum_{k=n+1}^{l+n} \mathbb{E}^Q[\tfrac{1}{\hat{v}_{t_k}}|\hat{v}_{t_n}] + \tfrac{1}{2}\mathbb{E}^Q[\tfrac{1}{\hat{v}_{t_{l+n+1}}}|\hat{v}_{t_n}])\Delta t + 2\lambda(\bar{\mu} - \mu_S)\} \times 100^2.
\end{aligned}$$

The next step is to perform numerical computation of the conditional expectation $\mathbb{E}^Q[\frac{1}{\hat{v}_{t_k}}|\hat{v}_{t_n}]$, $k = n+1, n+2, \cdots, l+n+1$. Given $\hat{v}_{t_n} = \hat{v}_i^n$ on the willow tree of $\hat{v}_t$ at $t_n$, we first compute

$$\mathbb{E}^Q[\frac{1}{\hat{v}_{t_{n+1}}}|\hat{v}_i^n] \approx \sum_{j=1}^m p_{ij}^n \frac{1}{\hat{v}_j^{n+1}},$$

where $p_{ij}^n$ is the transition probability between $\hat{v}_i^n$ and $\hat{v}_j^{n+1}$ on the willow tree. Next, we compute

$$
\begin{aligned}
\mathbb{E}^Q[\frac{1}{\hat{v}_{t_{n+2}}}|\hat{v}_i^n] &\approx \sum_{j=1}^m P[\hat{v}_j^{n+1}|\hat{v}_i^n]\mathbb{E}^Q[\frac{1}{\hat{v}_{t_{n+2}}}|\hat{v}_i^n, \hat{v}_j^{n+1}], \\
&= \sum_{j=1}^m p_{ij}^n \mathbb{E}^Q[\frac{1}{\hat{v}_{t_{n+2}}}|\hat{v}_j^{n+1}], \\
&= \sum_{j=1}^m p_{ij}^n \sum_{k=1}^m p_{jk}^{n+1} \frac{1}{\hat{v}_k^{n+2}}.
\end{aligned}
\tag{3.10}
$$

Writing $\mathbf{p}_i^n = [p_{i1}^n, p_{i2}^n, \cdots, p_{im}^n]^{\mathrm{T}}$ and $\mathbf{v}^{n+2} = [\frac{1}{\hat{v}_1^{n+2}}, \frac{1}{\hat{v}_2^{n+2}}, \cdots, \frac{1}{\hat{v}_m^{n+2}}]^{\mathrm{T}}$, we rewrite (3.10) in the matrix form as follows:

$$\mathbb{E}^Q[\frac{1}{\hat{v}_{t_{n+2}}}|\hat{v}_i^n] = (\mathbf{p}_i^n)^{\mathrm{T}}\mathbf{P}^{n+1}\mathbf{v}^{n+2}.$$

Recursively, the conditional expectation of $\frac{1}{\hat{v}_{t_k}}$ given $\hat{v}_i^n$ can be expressed in the form

$$\mathbb{E}^Q[\frac{1}{\hat{v}_{t_k}}|\hat{v}_i^n] = (\mathbf{p}_i^n)^{\mathrm{T}}\mathbf{P}^{n+1}\mathbf{P}^{n+2}\cdots\mathbf{P}^{k-1}\mathbf{v}^k, \qquad k = n+1, n+2, \cdots, l+n+1.$$

In summary, the conditional expectation in (3.10) can be calculated as

$$\mathbb{E}^Q[\frac{1}{\hat{v}_{t_k}}|\hat{v}_i^n] = (\mathbf{p}_i^{k-1})^{\mathrm{T}}\mathbf{v}^k, \qquad k = n+1, n+2, \cdots, l+n+1,$$

where $\mathbf{p}_i^n = [p_{i1}^n, p_{i2}^n, \cdots, p_{in}^n]^{\mathrm{T}}$ and $(\mathbf{p}_i^k)^{\mathrm{T}} = (\mathbf{p}_i^{k-1})^{\mathrm{T}}\mathbf{P}^k$.

Therefore, we manage to obtain the value of $\mathrm{VIX}_i^n$ given $\hat{v}_{t_n} = \hat{v}_i^n$. The willow tree of VIX is also constructed after we repeat the above procedure for all $\hat{v}_i^n$ for $i = 1, 2, \cdots, m$ and $n = 1, 2, \cdots, N$, using the transition probability matrices $\mathbf{P}^n, \mathbf{P}^{n+1}, \cdots, \mathbf{P}^{n+l}$.

Once we have obtained $\mathrm{VIX}_{t_N} = \mathrm{VIX}_i^N$, the pricing schemes for VIX futures and options are essentially the same as those performed for the 1/2-model in the previous subsection.

## 4  Numerical tests

This section presents numerical results that assess accuracy, efficiency and reliability of the willow tree algorithms for pricing VIX derivatives under the 1/2-model and 3/2-model. In our numerical tests, we take $\tau = \frac{1}{12}$, the convention of 30 calendar days adopted in CBOE. In our construction of the willow tree structure, the number of spatial nodes $m$ and time step $\Delta t$ are chosen to be 200 and 1/12, respectively. In order to obtain the values of the first four order moments of the instantaneous

variance and determine the transition probability matrices of the willow tree of the instantaneous variance, the Fourier-cosine series algorithm is used based on the knowledge of the conditional characteristic function of the instantaneous variance under the CIR process in the Fourier domain. In our numerical tests, the interval of integration [A,B] of the Fourier-cosine series expansion is determined by

$$[A, B] := \left[\max(c_1 - L\sqrt{c_2 + \sqrt{c_4}}, 0), c_1 + L\sqrt{c_2 + \sqrt{c_4}}\right] \quad \text{with} \quad L = 10,$$

where $c_1$, $c_2$ and $c_4$ are the respective first, second and fourth order cumulants of instantaneous variance $v_t$ [see (3.5b)]. More details on the discussion of the interval of integration can be found in Fang and Oosterlee (2008). Since the value of variance should not be negative, the lower bound of the integration is chosen to be the maximum value between $c_1 - L\sqrt{c_2 + \sqrt{c_4}}$ and 0. We set the number of summation terms $N$ in (3.5b) to be 1,000, which is sufficiently large for $L = 10$.

To provide the benchmark comparison for our willow tree calculation results, we also performed the Monte Carlo simulation to price VIX derivatives. We adopt the Euler-Maruyama discretization for the instantaneous variance dynamics. Under the 1/2-model, since $VIX^2$ can be expressed as a linear function in $v_t$ [see (2.4)], we simulate the path of $v_t$ from $t = 0$ to $t = T$. As a result, the value of VIX at $t = T$ can be obtained using (2.4). The prices of VIX futures and the European VIX call option can be obtained based on the Monte Carlo simulation procedure. Here, we set the number of simulation paths to be 10,000. However, under the 3/2-model, we cannot obtain the closed form formula for VIX. We have to employ the nested Monte Carlo simulation to price VIX derivatives under the 3/2-model. Similar to the numerical calculations under the 1/2-model, we first simulate 10,000 paths of $v_t$ from $t = 0$ to $t = T$. For each simulated path, as an additional numerical expectation procedure, we simulate 100,000 paths of $v_t$ from $t = T$ to $t = T + \tau$ and estimate the value of VIX at time $T$ using (2.6). Once the numerical estimate value of VIX under the 3/2-model at time $T$ is available, VIX derivatives can be priced in a similar manner.

In this section, we present the numerical pricing results of VIX derivatives using the willow tree (WT) algorithm and Monte Carlo (MC) simulation method under the 1/2-model and 3/2-model. For the 1/2-model, we also give the pricing results using the numerical integration (NI) of the Fourier integral price formulas for the VIX futures and the European call option as an additional benchmark calculations. As a remark, the saddlepoint approximation methods proposed by Kwok and Zheng (2018) can also be used to find accurate approximation values of VIX derivatives based on the available integral price formulas. However, the saddlepoint approximation methods may fail

to give sufficiently accurate results when the option are sufficiently deep-out-of-money or very near to maturity. Since there is no corresponding analytic integral price formulas for VIX derivatives under the 3/2-model, we can only compare the pricing results of VIX derivatives by WT algorithm with those of MC simulation calculations. In addition to the assessment of numerical accuracy, we examine the impact of various model parameters, like $\sigma_v$ and $\lambda$, on the pricing behaviors of the European VIX call options. We also examine the implied volatility smirks and the term structures of the implied skews of the European VIX call options under both the 1/2 model and 3/2-model. Lastly, we compute the American VIX call option prices and observe significant early exercise premium when compared with those of the European vanilla options.

## 4.1 1/2-model

Table 1: Parameter values for the 1/2-model.

| $r$ | $v_0$ | $\eta$ | $\theta$ | $\sigma_v$ | $\mu_S$ | $\sigma_S$ | $\lambda$ | $\rho_J$ | $\mu_v$ |
|---|---|---|---|---|---|---|---|---|---|
| 0.0319 | 0.0076 | 3.46 | 0.008 | 0.14 | -0.0865 | 0.0001 | 0.47 | -0.38 | 0.05 |

Table 2: Pricing results of VIX futures with varying maturities under the 1/2-model.

| Maturity | Price | | | CPU time (second) | |
|---|---|---|---|---|---|
| (month) | WT | MC (standard error) | NI | WT | MC |
| 1 | 12.0765 | 12.0895 (0.0269) | 12.0747 | 0.663 | 7.641 |
| 2 | 12.3793 | 12.3816 (0.0339) | 12.3788 | 3.942 | 15.177 |
| 3 | 12.6240 | 12.6426 (0.0379) | 12.6246 | 7.197 | 22.701 |
| 4 | 12.8185 | 12.8209 (0.0396) | 12.8199 | 10.459 | 30.289 |
| 5 | 12.9712 | 12.9778 (0.0411) | 12.9733 | 13.641 | 37.928 |
| 6 | 13.0899 | 13.0938 (0.0417) | 13.0926 | 17.179 | 45.501 |
| 7 | 13.1816 | 13.1766 (0.0420) | 13.1847 | 20.534 | 53.028 |
| 8 | 13.2520 | 13.2582 (0.0424) | 13.2555 | 23.443 | 60.531 |
| 9 | 13.3058 | 13.3221 (0.0426) | 13.3095 | 26.875 | 68.239 |
| 10 | 13.3467 | 13.3615 (0.0427) | 13.3507 | 30.390 | 75.632 |

In our numerical experiments, we used the same set of parameters of the 1/2-model showed

Table 3: Pricing results of 3-month European VIX call options with varying strike prices under the 1/2-model.

| Strike price | Price | | | CPU time (second) | |
|---|---|---|---|---|---|
| | WT | MC (standard error) | NI | WT | MC |
| 10.5 | 2.2310 | 2.2391 (0.0364) | 2.2353 | 6.723 | 22.856 |
| 11.0 | 1.8646 | 1.8659 (0.0356) | 1.8690 | 6.689 | 22.890 |
| 11.5 | 1.5572 | 1.5527 (0.0345) | 1.5615 | 6.826 | 22.841 |
| 12.0 | 1.3082 | 1.3057 (0.0335) | 1.3125 | 6.688 | 22.867 |
| 12.5 | 1.1121 | 1.1218 (0.0328) | 1.1164 | 6.698 | 22.856 |
| 13.0 | 0.9607 | 0.9574 (0.0312) | 0.9649 | 6.863 | 22.871 |
| 13.5 | 0.8446 | 0.8428 (0.0304) | 0.8487 | 6.831 | 22.842 |
| 14.0 | 0.7549 | 0.7590 (0.0294) | 0.7590 | 6.656 | 22.847 |

in Table 1 (unless otherwise stated) as reported in Lian and Zhu (2013). In Table 2, we present the numerical results of VIX futures using the willow tree (WT) algorithm, Monte Carlo (MC) simulation method with standard errors and direct numerical integration (NI). The VIX futures values obtained by WT algorithm are seen to be very close to those obtained using NI, and the difference between VIX futures prices obtained by the WT algorithm and MC simulation is typically less than the standard error of MC simulation. Comparing the CPU times required by the WT algorithm and MC simulation, the former is typically much less, even for long-maturity futures. This reveals that the WT algorithm is accurate and more computationally efficient than MC simulation when pricing VIX futures. Table 3 presents the pricing results of the 3-month European VIX call option with varying strike prices. Similar high level of numerical performance of the WT algorithm in terms of accuracy and computational efficiency is observed.

## 4.2   3/2-model

For the 3/2-model, we choose some of the parameters of the 3/2-model in Lin *et al.* (2017). In their set of parameters, the jump in the index process is quite small, so we modify the index jump component to be the same magnitude as those in the 1/2-model. The values of parameters for the 3/2-model used in our numerical tests are listed in Table 4.

Tables 5 and 6 present the pricing results of VIX futures and 3-month European VIX call

Table 4: Parameter values for the 3/2-model.

| $r$ | $v_0$ | $\eta$ | $\theta$ | $\sigma_v$ | $\mu_S$ | $\sigma_S$ | $\lambda$ |
|---|---|---|---|---|---|---|---|
| 0.05 | 0.0076 | 26.3189 | 0.0935 | 9.2499 | -0.0865 | 0.0001 | 0.47 |

Table 5: Pricing results of VIX futures with varying maturities under the 3/2-model.

| Maturity | Price | | CPU time (second) | |
|---|---|---|---|---|
| (month) | WT | MC (standard error) | WT | MC |
| 1 | 11.5464 | 11.5306 (0.0098) | 3.924 | 3904.024 |
| 2 | 12.2835 | 12.2857 (0.0166) | 7.265 | 3891.672 |
| 3 | 13.0314 | 13.0289 (0.0231) | 10.472 | 3899.843 |
| 4 | 13.7724 | 13.7754 (0.0296) | 14.041 | 3891.363 |
| 5 | 14.5091 | 14.5556 (0.0341) | 16.994 | 3925.915 |
| 6 | 15.2182 | 15.1781 (0.0407) | 20.259 | 3982.429 |
| 7 | 15.8632 | 15.7306 (0.0439) | 23.737 | 3998.580 |
| 8 | 16.4325 | 16.4850 (0.0475) | 27.032 | 4012.579 |
| 9 | 16.9226 | 16.8488 (0.0509) | 30.488 | 4013.147 |
| 10 | 17.3262 | 17.4298 (0.0510) | 34.239 | 4042.311 |

Table 6: Pricing results of 3-month European VIX call options with varying strike prices under the 3/2-model.

| Strike price | Price | | CPU time (second) | |
|---|---|---|---|---|
| | WT | MC (standard error) | WT | MC |
| 10.5 | 2.5380 | 2.5477 (0.0231) | 10.552 | 4115.228 |
| 11.0 | 2.1045 | 2.1107 (0.0227) | 10.752 | 4101.472 |
| 11.5 | 1.7166 | 1.7015 (0.0219) | 10.682 | 4084.472 |
| 12.0 | 1.3816 | 1.3880 (0.0217) | 10.817 | 4098.624 |
| 12.5 | 1.1007 | 1.0975 (0.0204) | 11.183 | 4083.217 |
| 13.0 | 0.8705 | 0.8743 (0.0192) | 10.724 | 4091.092 |
| 13.5 | 0.6855 | 0.6904 (0.0174) | 10.695 | 4078.079 |
| 14.0 | 0.5385 | 0.5364 (0.0156) | 10.713 | 4092.423 |

option under the 3/2-model, respectively, obtained by WT and MC. We also compare the CPU times required by these two methods. Similar to the 1/2-model, we observe good agreement of the numerical values obtained using WT and MC, while the CPU time required by the former is typically less than 1% of that of MC method. The advantage of the willow tree algorithms over the Monte Carlo simulation is more obvious under the 3/2-model since there is no closed form formula for $VIX^2$ in terms of the instantaneous variance as in the 1/2-model.

## 4.3 Impact of the volatility of instantaneous variance

In this subsection, we analyse the impact of the parameter $\sigma_v$ in the instantaneous variance process on the prices of the European VIX call option with strike price equals 13, while the other model parameters remain unchanged.
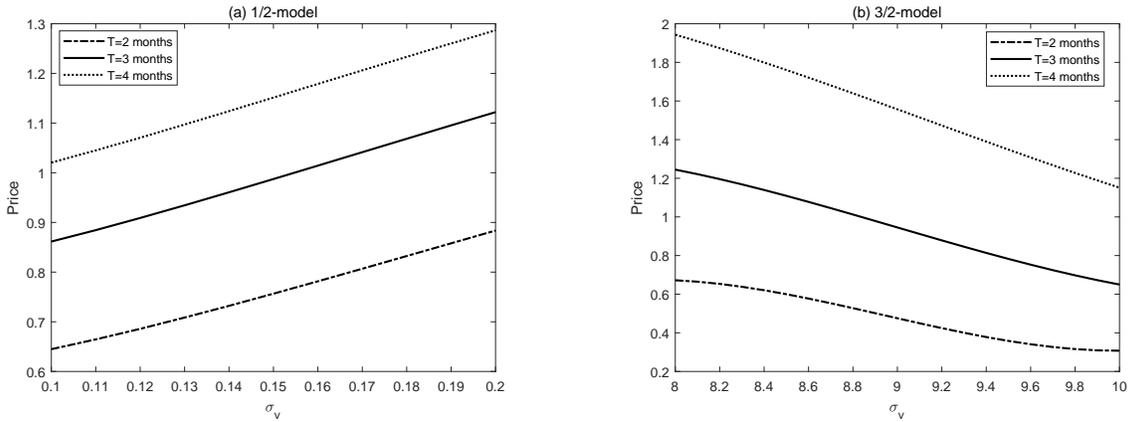


Figure 2: Pricing results of the European VIX call option with strike price $K = 13$ under different values of $\sigma_v$.

Figure 2 shows the plots of the values of the European VIX call option with varying values of $\sigma_v$ under the 1/2-model and 3/2-model. The plots show that the value of the European VIX call option increases with the increase of $\sigma_v$ for the 1/2-model while the option value is a decreasing function of $\sigma_v$ for the 3/2-model. The phenomena under the 3/2-model are similar to those reported in Yuen *et al.* (2015) for variance swaps and the same explanation of the pricing behavior that VIX option price is decreasing with respect to $\sigma_v$ under the 3/2-model can be applied.

## 4.4 Impact of jumps in index value and its instantaneous variance

We also performed numerical tests to study the impact of jumps on the European VIX call option prices with different levels of jump intensity under both 1/2-model and 3/2-model. The other model parameters used in our numerical tests remain unchanged.
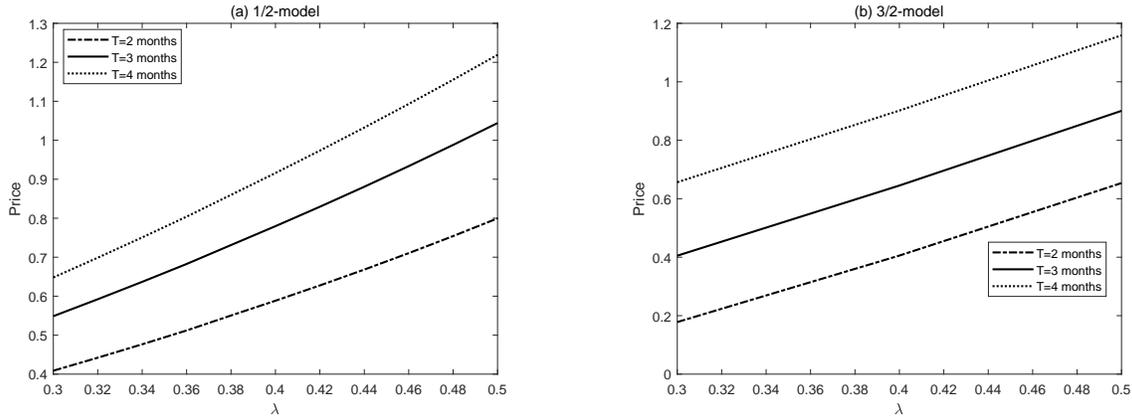


Figure 3: Pricing results of European VIX call option with strike price $K = 13$ under different values of $\lambda$.

Figure 3 shows the plots of the prices of the European VIX call options against the jump intensity $\lambda$ with varying maturities under the 1/2-model and 3/2-model. We observe that the price of the European VIX call option increases quite significantly with increasing jump intensity under both 1/2-model and 3/2-model.

## 4.5 Implied volatility

We performed numerical tests to examine the behaviors of implied volatility of the European VIX call options with respect to strike price and maturity under the 1/2-model and 3/2-model. Figures 4 and 5 show the plots of the implied volatility of the European VIX call option with respect to strike price and maturity under both stochastic volatility models with different values of $v_0$. The implied volatility values exhibit the smile pattern under the 1/2-model but downward sloping pattern under the 3/2-model. On the other hand, the implied volatility values are decreasing with respect to maturity under both stochastic volatility models.
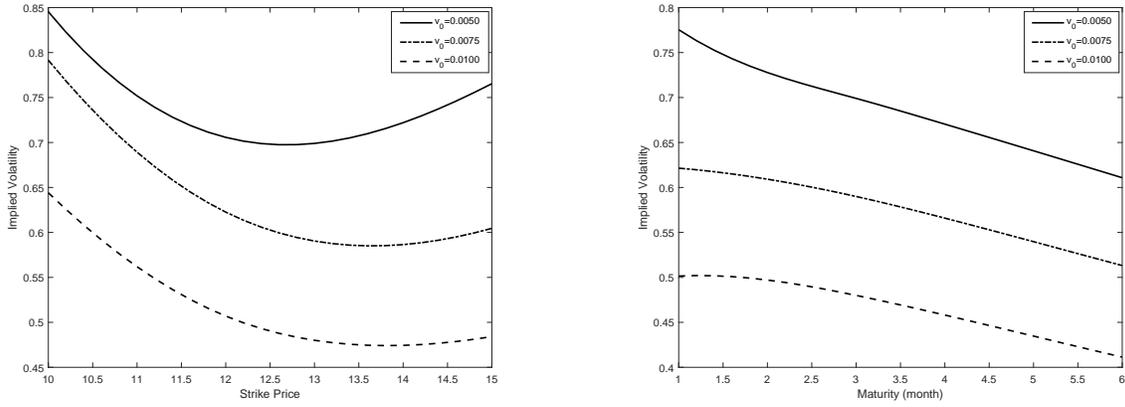
Figure 4: Plot of the implied volatility values of 3-month European VIX call option with respect to strike price (left figure) and maturity with strike price $K = 13$ (right figure) under varying values of instantaneous variance $v_0$ in the 1/2-model.
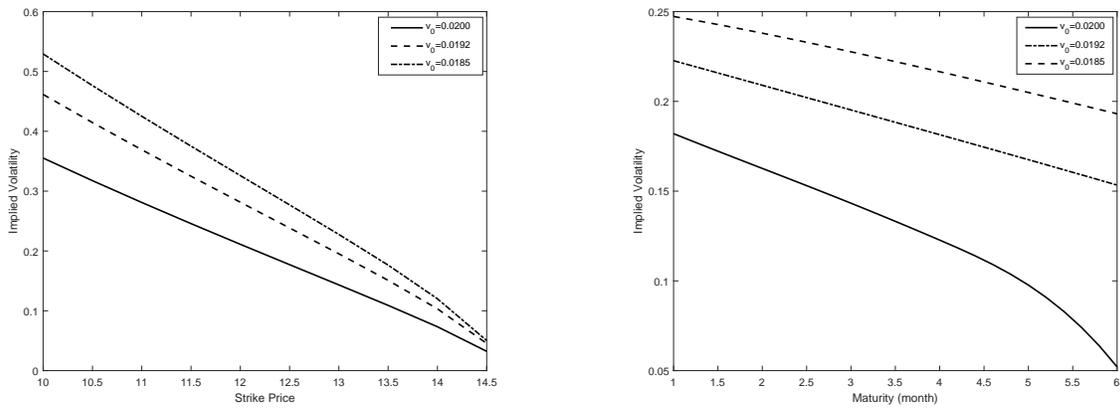


Figure 5: Plot of the implied volatility values of 3-month European VIX call option with respect to strike price (left figure) and maturity with strike price $K = 13$ (right figure) under varying values of instantaneous variance $v_0$ in the 3/2-model.

## 4.6 American VIX option

In Figure 6, we show the plots of the time-0 prices of the European and American VIX call option with varying values of $VIX_0$ under the 1/2-model and 3/2-model, where $VIX_0$ is the prevailing VIX value at the current time at $t = 0$. The intrinsic value is the exercise payoff $(VIX_0 - K)^+$, where the strike $K$ equals 13. We observe that the early exercise premium of the American VIX call option can be more significant compared with that of vanilla call option counterpart. These observations are consistent with those reported in Detemple and Kitapbayev (2018).
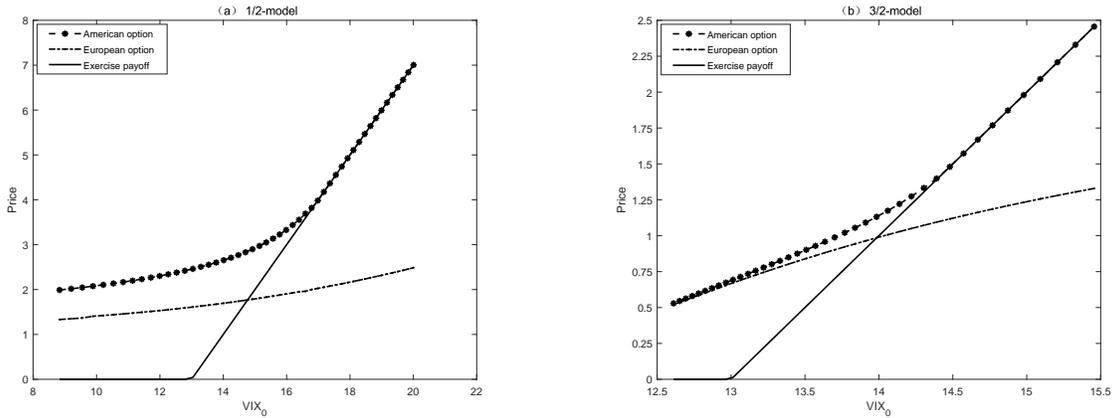


Figure 6: Plots of the European and American VIX call option values against $VIX_0$ with strike price $K = 13$ and maturity $T = 3$ months under the 1/2-model and 3/2-model.

## 5 Conclusion

We present the willow tree algorithms for pricing VIX derivatives under the 1/2-model and 3/2-model under the joint dynamics of the S&P index and its instantaneous variance processes. The construction of the willow tree for the instantaneous variance involves finding the positions of the nodes on the willow tree structure and the transition probability matrices between the nodes of successive time steps using the information of the first four order moments of the instantaneous variance. Based on the closed form representation of the characteristic function of the instantaneous variance under the CIR process, we can employ the Fourier cosine transform algorithm to compute these higher order moments effectively. We have performed comprehensive numerical tests to demonstrate that the willow tree algorithms compete well with the Monte Carlo simulation in terms of accuracy, efficiency and reliability. In particular, significant improvement on computation-

al efficiency is seen when we consider pricing VIX derivatives under the 3/2-model. While pricing VIX derivatives under the 1/2-model enjoys nice analytical tractability, there is no closed form representation of VIX in terms of the instantaneous variance under the 3/2-model. This lack of analytical tractability under the 3/2-model only poses minor additional computational efforts under our willow tree algorithms while much added efforts are required in other numerical schemes. In our numerical tests, we also examine the pricing behaviors of VIX futures and VIX call options under the 1/2-model and 3/2-model. We observe quite different pricing behaviors under the two stochastic volatility models. Similar discrepancies of pricing behaviors are also reported in other papers on related variance and VIX products.

## References

Baldeaux, J. and Badran, A. (2014). Consistent modeling of VIX and equity derivatives using a 3/2 plus jumps model. *Applied Mathematical Finance*, **21(4)**, 299-312.

Bakshi, G., Ju, N and Yang, N. (2006). Estimation of continuous time models with an application to equity volatility. *Journal of Financial Economics*, **82**, 227-249.

Barletta, A. and Nicolato, E. (2018). Orthogonal expansions for VIX options under affine jump diffusion. *Quantitative Finance*, **18(6)**, 951-967.

Branger, N., Kraftschik, A. and Völkert, C. (2017). The variance process implied in VIX options: Affine vs. non-affine models. *Working paper of University of Münster*.

Carr, P. and Wu, L.R. (2006). A tale of two indices. *Journal of Derivatives*, **13(3)**, 13-29.

Cont, R. and Kokholm, T. (2013). A consistent pricing model for index options and volatility derivatives. *Mathematical Finance*, **23(2)**, 248-274.

Curran, M. (2001). Willow power: Optimizing derivative pricing trees. *Algo Research Quarterly*, **4(4)**, 15-24.

Detemple, J. and Kitapbayev, Y. (2018). On American VIX options under the generalized 3/2 and 1/2 models. *Mathematical Finance*, **28(2)**, 550-581.

Dong, B., Xu, W. and Kwok, Y.K. (2019). Willow tree algorithms for pricing Guaranteed Minimum Withdrawal Benefit under jump-diffusion and CEV models. *Quantitative Finance*, **19(10)**, 1741-1761.

Duan, J.C. and Yeh, C.Y. (2010). Jump and volatility risk premiums implied by VIX. *Journal of Economic Dynamics and Control*, **34**, 2232-2244.

Fang, F. and Oosterlee, C.W. (2008). A novel pricing method for European options based on Fourier-cosine series expansions. *SIAM Journal on Scientific Computing*, **31(2)**, 826-848.

Goard, J. and Mazur, M. (2013). Stochastic volatility models and the pricing of VIX options. *Mathematical Finance*, **23(3)**, 439-458.

Gonzalez-Perez, M.T. (2015). Model-free volatility indexes in the financial literature. *International Review of Economics and Finance*, **40**, 141-159.

Grasselli, M. (2017). The 4/2 stochastic volatility model: A unified approach for the Heston and

the 3/2 model. *Mathematical Finance*, **27(4)**, 1013-1034.

Heston, S.L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, **6(2)**, 327-343.

Hill, I. and Holder, R. (1976). Algorithm as 99: Fitting Johnson curves by moments. *Applied Statistics*, **25**, 180-189.

Johnson, N. (1949). System of frequency curves generated by methods of translation. *Biometrika*, **36**, 149-176.

Kaeck, A. and Alexander, C. (2012). Volatility dynamics for the S&P 500: Further evidence from non-affine, multi-factor jump diffusions. *Journal of Banking and Finance*, **36**, 3110-3121.

Kwok, Y.K. and Zheng, W.D. (2018). *Saddlepoint Approximation Methods in Financial Engineering.* Springer, Berlin.

Li, J., Li, L.F. and Zhang, G.Q. (2017). Pure jump models for pricing and hedging VIX derivatives. *Journal of Economic Dynamics and Control*, **74**, 28-55.

Lian, G.H. and Zhu, S.P. (2013). Pricing VIX options with stochastic volatility and random jumps. *Decisions in Economics and Finance*, **36**, 71-88.

Lin, W., Li, S.H., Chern, S. and Zhang, J.E. (2018). Pricing VIX derivatives with free stochastic volatility model. *Review of Derivatives Research,* **22(1)**, 41–75.

Lin, W., Li, S.H., Luo, X.G. and Chern, S. (2017). Consistent pricing of VIX and equity derivatives with the 4/2 stochastic volatility plus jumps model. *Journal of Mathematical Analysis*, **447**, 778-797.

Lin, Y.N. (2007). Pricing VIX futures: Evidence from integrated physical and risk-neutral probability measures. *Journal of Futures Markets*, **27(12)**, 1175-1217.

Luo, X.G. and Zhang, J.E. (2012). The term structure of VIX. *Journal of Futures Markets*, **32(12)**, 1092-1123.

Mencía, J. and Sentana, E. (2013). Valuation of VIX derivatives. *Journal of Financial Economics*, **108**, 367-391.

Park, Y.H. (2016). The effects of asymmetric volatility and jumps on the pricing of VIX derivatives. *Journal of Econometrics*, **192(1)**, 313-328.

Psychoyios, D., Dotsis, G. and Markellos, R. (2010). A jump diffusion model for VIX volatility options and futures. *Review of Quantitative Finance and Accounting*, **35(3)**, 245-269.

Sepp, A. (2008). VIX option pricing in a jump-diffusion model. *Risk*, 84-89.

Wang, G.G. and Xu, W. (2018). A unified willow tree framework for one-factor short rate models. *Journal of Derivatives*, **25(3)**, 33-54.

Wang, T.Y., Shen, Y.W., Jiang, Y.T. and Huang, Z. (2017). Pricing the CBOE VIX futures with the Heston-Nandi GARCH model. *Journal of Futures Markets*, **37(7)**, 641-659.

Xu, W., Hong, Z.W. and Qin, C.X. (2013). A new sampling strategy willow tree method with application to path-dependent option pricing. *Quantitative Finance*, **13(6)**, 861-872.

Xu, W. and Yin, Y.F. (2014). Pricing American options by willow tree method under jump-diffusion process. *Journal of Derivatives*, **22**, 46-56.

Yuen, C.H., Zheng, W.D. and Kwok, Y.K. (2015). Pricing exotic discrete variance swaps under the 3/2-stochastic volatility models. *Applied Mathematical Finance*, **22(5)**, 421-449.

Zhang, J.E. and Zhu, Y.Z. (2006). VIX futures. *Journal of Futures Markets*, **26(6)**, 521-531.

Zhu, S.P. and Lian, G.H. (2012). An analytical formula for VIX futures and its applications. *Journal of Futures Markets*, **32(2)**, 166-190.

Zhu, Y.Z. and Zhang, J.E. (2007). Variance term structure and VIX futures pricing. *International Journal of Theoretical and Applied Finance*, **10(1)**, 111-127.

# Appendix A  Johnson curve transform

The Johnson curve transform (Johnson, 1949) states that any probability distribution can be transformed into a standard normal distribution. By inverting the transformation, we can obtain the approximate distribution that matches the first four order moments of the original distribution. There are three choices of the Johnson curves:

- lognormal system (or $S_L$): $Z = a + b \ln(v - c), \quad v > c,$

- unbounded system (or $S_U$): $Z = a + b \sinh^{-1}(\frac{v-c}{d}),$

- bounded system (or $S_B$): $Z = a + b \ln(\frac{v-c}{c+d-v}), \quad c < v < c + d,$

where $Z \sim N(0,1)$, $v$ can be any distribution and the parameters $a, b, c, d$, are determined by the first four order moments of $v$.

Hill and Holder (1976) propose a method that decides which transformation in the above to be employed for a specified distribution. Let

$$\gamma = \omega^4 + 2\omega^3 + 3\omega^2 - 3,$$

where $\omega$ is the solution of following equation:

$$(\omega - 1)(\omega + 2)^2 = \kappa_3^2.$$

Here, $\kappa_3$ is the skewness of the specified distribution. Let $\kappa_4$ be the kurtosis of the specified distribution. We choose $S_B$ or $S_U$ according to $\gamma < \kappa_4$ or $\gamma > \kappa_4$, respectively. Specifically, when $\gamma = \kappa_4$, $S_L$ is the appropriate choice. More details can be found in Hill and Holder (1976).