Willow tree algorithms for pricing VIX derivatives under jump-diffusion dynamics of index and variance processes

Changfu Ma\textsuperscript{1}, Wei Xu\textsuperscript{2†} and Yue Kuen Kwok\textsuperscript{3}

\textsuperscript{1,2} School of Mathematical Sciences, Tongji University, Shanghai 200092, China
\textsuperscript{3} Department of Mathematics, Hong Kong University of Science and Technology, Hong Kong, China

Abstract

VIX futures and option 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. To price VIX derivatives, one may adopt either the direct modeling of VIX dynamics or the consistent model approach with joint modeling of the dynamics of S&P index and its instantaneous variance. Under certain class of the joint process of the S&P index and its instantaneous variance, analytic integral formulas can be derived to price VIX futures and option. However, these integral formulas invariably involve Fourier inversion integrals with cumbersome hyper-geometric functions, thus posing various challenges in numerical evaluation. In this paper, we propose a unified numerical approach based on the willow tree algorithms to price VIX derivatives under 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 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 to approximate 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 the European and American options under the consistent model approach. We examine the impact of inclusion of jumps in the instantaneous variance process and other parameter values in pricing VIX derivatives. We also examine the implied volatility smirks and the term structures of the implied skewness of the model prices of VIX options.

\textsuperscript{†}Corresponding author. E-mails: \textsuperscript{1} 2014mcf@tongji.edu.cn; \textsuperscript{2} wxu@tongji.edu.cn; \textsuperscript{3} maykwok@ust.hk.
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 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). They present an interesting discussion on 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 option. Indeed, before the introduction of VIX, Brenner and Galai (1989) propose a volatility index (Sigma index) and discuss various derivatives on this index. In addition, Whaley (1993) discusses the role of derivatives on market volatility as effective hedging tools. Market practitioners use VIX derivatives to hedge the risks of investments 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 option was started in 2006. The contract multiplier for each VIX futures contract is $1,000 while that of VIX option contract is $100. The popularity of trading on VIX futures and option 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 volume was 644,892 in reaction to the surprise outcome of the US presidential election. For VIX option at 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 approaches for pricing VIX derivatives. In the consistent model approach, one considers the stochastic dynamics of SPX under stochastic volatility and derives the dynamics for VIX. The other approach directly models the dynamics for VIX (Psychoyios et al., 2010; Goard and Mazur, 2013; Park, 2017; 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 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. The direct modeling approach may ignore the linkage between SPX and VIX, where VIX is derived from the prices of options on SPX.

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 Heston stochastic volatility model with simultaneous jumps, Luo and Zhang (2012) analyze the term structure of VIX and show that the square of VIX can be expressed as the weighted average of the long term mean of the instantaneous variance process and the instantaneous variance as well. In general, nice analytic tractability exists under the affine Heston model with jumps and various forms of price formulas of VIX futures and option 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 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 non-affine diffusion and variance jumps are required to capture both the short-term and long-term implied volatility distribution. Baldeaux and Badran (2014) manage to obtain integral price formulas for VIX futures and option under the pure diffusion 3/2-model. Based on the 4/2 stochastic volatility model proposed by Grasselli (2017), Lin et al. (2017) derive integral price formulas for VIX derivatives under the 4/2 stochastic volatility plus jump model. Later, Lin et al. (2018) derive price formulas of VIX derivatives under their free stochastic volatility model, with an arbitrary power parameter of the instantaneous variance in the dynamic equation of SPX and its instantaneous variance. This class of models reduce to the usual Heston 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 various types of consistent models, numerical valuations 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 saddle-point approximation methods to evaluate the integral price formulas of VIX derivatives under the Heston model. However, numerical accuracy may deteriorate when the jump sizes are significant, options are deep-out-of-the-money and at times close to expiration. Therefore, it is necessary 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). The willow tree method has been demonstrated in these research papers 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 employs 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, together 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 processes.

This paper is organized as follows. In the next section, we discuss the model formulation
of VIX under the 1/2-model (Heston model) and 3/2-model with jumps on both the index and its
instantaneous variance. We show how $VIX_t^2$ can be expressed in terms of the instantaneous variance
$v_t$ and other relevant model parameters. In Section 3, we present the construction of the willow
tree algorithm that price VIX futures and option under the Heston 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. It is effective to compute
these probabilities using the fast Fourier cosine algorithm based on the knowledge of the conditional
characteristic function of the instantaneous variance. In Section 4, we present the numerical tests
that were performed to assess accuracy, efficiency and reliability of the willow tree algorithm. With
availability of the effective willow tree pricing algorithm, we examine various pricing behaviors of
VIX futures and option, 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 in CBOE. Following Zhang and Zhu (2006), VIX is defined via the relation:

$$VIX_t^2(\tau) = -\frac{2}{\tau} E_t^Q [ \ln \frac{S_{t+\tau}}{S_t e^{r\tau}} ] \times 100^2, \tag{2.1}$$

where $r$ is the risk free interest rate and $E_t^Q$ is the expectation under the risk neutral measure
$Q$ conditional on 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 in traded VIX options. In our dynamic joint process of $S_t$ and $v_t$, we assume that
simultaneous jumps on $S_t$ and $v_t$ are modeled by the common Poisson process $N_t$ with constant
intensity \( \lambda \). 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

\[
J^v \sim \exp(\mu_v); \quad (2.2a)
\]

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

\[
J^S|J^v \sim N(\mu_S + \rho_J J^v, \sigma_S^2). \quad (2.2b)
\]

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 \( \rho_J \mu_v < 1 \).

The joint process of \( S_t \) and \( v_t \) under the consistent model is specified by the dynamic equations:

\[
\frac{dS_t}{S_t} = (r - \bar{\mu} \lambda) \ dt + \sqrt{v_t} \ dW_t^1 + (e^{J^S} - 1) \ dN_t \quad (2.3a)
\]

\[
dv_t = \eta v_t^{\frac{\alpha - 1}{2}}(\theta - v_t) \ dt + \sigma_v v_t^\alpha \ dW_t^2 + J^v \ dN_t, \quad (2.3b)
\]

where the Brownian motions \( W_t^1 \) and \( W_t^2 \) observe \( \mathbb{E}^Q[dW_t^1 dW_t^2] = \rho \ dt \), \( \rho \) is the constant correlation coefficient between \( dW_t^1 \) and \( dW_t^2 \), \( \theta \) 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 consider the two special cases of \( \alpha \): (i) \( \alpha = \frac{1}{2} \) corresponds to the 1/2-model, commonly called the Heston jump-diffusion model (Heston, 1993); (ii) \( \alpha = \frac{3}{2} \) corresponds to the 3/2-model. Based on the above consistent model of the joint process of \( S_t \) and \( v_t \), we manage to express \( VIX_t^2(\tau) \) in terms of the instantaneous variance \( v_t \) and other relevant model parameters.

### 2.1 Heston model

For the 1/2-model, we deduce from Itô’s lemma that

\[
d\ln S_t = (r - \bar{\mu} \lambda - \frac{v_t}{2}) \ dt + \sqrt{v_t} \ dW_t^1 + J^S \ dN_t.
\]

Substituting the above equation into (2.1), we obtain

\[
VIX_t^2(\tau) = \frac{d}{t} \mathbb{E}_t^Q \left[ \int_t^{t+\tau} (\frac{v_u}{2} + \bar{\mu} \lambda) \ du - J^S \ dN_u \right] \times 100^2
\]

\[
= \left\{ \frac{1}{2} \mathbb{E}_t^Q \left[ \int_t^{t+\tau} v_u \ du \right] + 2\lambda[\bar{\mu} - (\mu_S + \rho_J \mu_v)] \right\} \times 100^2.
\]
By solving the dynamic equation for \( v_t \) in (2.3b) under \( \alpha = \frac{1}{2} \), we obtain
\[
v_u = e^{-\eta(u-t)}v_t + \theta[1 - e^{-\eta(u-t)}] + \int_t^u \sigma_v e^{-\eta(u-s)}\sqrt{s} \, dW_s + \int_t^u e^{-\eta(u-s)}J^v \, dN_s, \quad u > t.
\]
It remains to evaluate the integration of \( v_u \) over the time interval \([t, t+\tau]\). Taking expectation \( \mathbb{E}^Q_t \) on both sides of the above equation, we have
\[
\mathbb{E}^Q_t [v_u] = e^{-\eta(u-t)}v_t + \theta[1 - e^{-\eta(u-t)}] + \int_t^u e^{-\eta(u-s)}\mathbb{E}^Q_t [J^v] \mathbb{E}^Q_t [dN_s]
\]
so that
\[
VIX^2_t(\tau) = \frac{1}{\tau} \int_t^{t+\tau} \mathbb{E}^Q_t [v_u] \, du + 2\lambda[\tilde{\mu} - (\mu_S + \rho J \mu_v)] \bigg|_{\tau} \times 100^2
\]
This result provides nice tractability since \( VIX^2_t(\tau) \) can be expressed as a linear function in \( v_t \) under the 1/2-model in the form
\[
VIX^2_t(\tau) = [a_0(\tau) + a_1(\tau)v_t] \times 100^2, \quad (2.4)
\]
where the \( \tau \)-dependent parameters are given by
\[
a_0(\tau) = (1 - \frac{1-e^{-\eta \tau}}{\eta \tau})(\theta + \frac{\lambda \mu_v}{\eta}) + 2\lambda[\tilde{\mu} - (\mu_S + \rho J \mu_v)],
\]
\[
a_1(\tau) = \frac{1-e^{-\eta \tau}}{\eta \tau}.
\]

### 2.2 3/2-model

For the 3/2-model, thanks to the reciprocal transformation \( \hat{v}_t = \frac{1}{v_t} \), the dynamic equation of \( \hat{v}_t \) can be reduced to
\[
d\hat{v}_t = \hat{\eta}(\hat{\theta} - \hat{v}_t) \, dt + \hat{\sigma}_v \hat{v}_t^{\frac{1}{2}} \, dW_t^2,
\]
where \( \hat{\eta} = \eta \theta, \hat{\theta} = \frac{\eta + \sigma_v^2}{\eta \tau} \) 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, \( VIX^2_t(\tau) \) under the 3/2-model can be expressed as
\[
VIX^2_t(\tau) = \frac{1}{\tau} \int_t^{t+\tau} \mathbb{E}^Q_t [\frac{1}{\hat{v}_u}] \, du + 2\lambda[\tilde{\mu} - (\mu_S + \rho J \mu_v)] \bigg|_{\tau} \times 100^2. \quad (2.5)
\]
Unlike the 1/2-model case, we cannot obtain closed form formula for \( \mathbb{E}^Q_t [\frac{1}{\hat{v}_u}] \).

### 2.3 VIX futures and options

The exchange traded VIX has fixed \( \tau \) at the value of 30 days. For notational convenience, we drop \( \tau \) in \( VIX_t(\tau) \) if there is no ambiguity on \( \tau \). For VIX futures maturing at \( T \), the time-\( t \) futures price is given by
\[
F_t = \mathbb{E}^Q_t [VIX_T], \quad (2.6)
\]
Similarly, the time-$t$ prices of the European and American VIX call options maturing at $T$ with strike price $K$ are given by

$$c_t = e^{-r(T-t)} \mathbb{E}_t^Q[(\text{VIX}_T - K)^+], \quad (2.7a)$$

$$C_t = \sup_{u \in [t,T]} \mathbb{E}_t^Q[e^{-r(u-t)}(\text{VIX}_u - K)^+], \quad (2.7b)$$

respectively.

3 Pricing VIX derivatives by willow tree algorithm

In this section, we discuss how the willow tree structure can be constructed based on the approximation of the distribution of 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 computing the intrinsic value and continuation value at each node. For the 1/2-model, $\text{VIX}_t^2$ 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 [see (2.5)].

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 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 points 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 $t_n = n\Delta t$, $n = 0,1,2,\cdots,N$. On the willow tree structure, we assume that there are $m$ nodes at each time step. At time $t_n$, we employ $m$ discrete values of the instantaneous variance $v^n_i$, $i = 1,2,\cdots,m$, 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.
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.

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, $v^n_i$ can be obtained by the inverse Johnson curve transform according to the sample points $Z_i$. Using the inverse Johnson curve transform, we have

$$v_{tn} = a_3 + a_4 G^{-1}(\frac{Z - a_1}{a_2}),$$  \hspace{1cm} (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 $v_{tn}$ (Hill and Holder, 1976). Accordingly, the discrete values of the instantaneous variance are given by

$$v^n_i = a_3 + a_4 G^{-1}(\frac{Z_i - a_1}{a_2}), \hspace{1cm} \text{for } i = 1, 2, \cdots, m.$$  \hspace{1cm} (3.1b)

First, we present the construction of the willow tree under the Heston dynamics of $v_t$. The procedure takes advantage of availability of the analytic formula of the conditional characteristic function of $v_t$. Let $f(\phi; t, \tau, v_t)$ denote the conditional characteristic function of the instantaneous variance $v_{t+\tau}$ conditional on filtration $\mathcal{F}_t$ 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 a complex-valued Fourier transform variable and

\[
\begin{align*}
\Gamma(\phi, \tau) &= \frac{2\mu_{\lambda}}{2\mu_{\lambda} - \sigma^2} \ln \left( 1 + \frac{(\sigma^2 - 2\mu_{\lambda}\eta)}{2\eta(1-\mu_{\lambda})} (e^{-\eta \tau} - 1) \right), \\
E(\phi, \tau) &= \frac{2\eta\phi}{\sigma^2} \ln \left( 1 + \frac{\sigma^2 \eta}{2\eta} (e^{-\eta \tau} - 1) \right), \\
D(\phi, \tau) &= \frac{2\eta\phi}{\sigma^2 (\eta^2 - \sigma^2 \phi)} e^{\eta \tau}.
\end{align*}
\]

By taking the inverse Fourier transform, the probability density function of \(v_t\) conditional on filtration \(\mathcal{F}_0\) is

\[
g(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ixw} f(iw; 0, t, v_0) \, dw. \tag{3.3}
\]

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

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

where

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

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 \(n\)th order moment of the instantaneous variance \(v_t\) is given by

\[
\mathbb{E}_0^Q [(v_t)^n] = \int_0^\infty x^n g(x) \, dx, \tag{3.5a}
\]

which can be approximated by

\[
\begin{align*}
\mathbb{E}_0^Q [(v_t)^n] &\approx \sum_{k=0}^N F_k \int_A^B x^n \cos(k \frac{x - A}{B - A} \pi) \, dx, \\
&= \sum_{k=0}^N \frac{2}{B - A} \text{Re} \left\{ f(i \frac{k\pi}{B - A}; 0, t, v_0) \exp(-i \frac{k A \pi}{B - A}) \right\} \varphi_n(k), \tag{3.5b}
\end{align*}
\]

where

\[
\varphi_n(k) = \int_A^B x^n \cos(k \frac{x - A}{B - A} \pi) \, dx.
\]

When \(n = 1, 2, 3, 4\), the explicit expressions of these integrals are given by

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

\[
\varphi_2(k) = \frac{4}{k\pi^2} \left[ (\frac{B-A}{k\pi})^2 [(-1)^k - 1] \right], \quad \text{for } k \neq 0,
\]

\[
\varphi_3(k) = \frac{4}{k\pi^2} \left[ (\frac{B-A}{k\pi})^2 \left\{ (\frac{B-A}{k\pi})^2 [(-1)^k - 1] \right\} \right], \quad \text{for } k \neq 0,
\]

\[
\varphi_4(k) = \frac{4}{k\pi^2} \left[ (\frac{B-A}{k\pi})^2 \left\{ (\frac{B-A}{k\pi})^2 \left\{ (\frac{B-A}{k\pi})^2 [(-1)^k - 1] \right\} \right\} \right], \quad \text{for } k \neq 0.
\]
\[ \varphi_2(k) = \begin{cases} 2\left(\frac{B-A}{k\pi}\right)^2[B(-1)^k - A], & \text{for } k \neq 0, \\ \frac{B^3-A^3}{3}, & \text{for } k = 0; \end{cases} \]

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

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

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 \( Z_i, i = 1, 2, 3, \ldots, m \), as the approximation of the standard normal distribution.

2. According to (3.5b), we obtain the approximate values of the first four moments of \( v_t^n \). We then determine the parameters \( a_1, a_2, a_3, a_4 \) and function \( G^{-1}(\cdot) \) using the Johnson curve transform.

3. Using the inverse Johnson curve transform [see (3.1b)], we obtain

\[ v^n_i = a_3 + a_4 G^{-1}\left(\frac{Z_i - a_1}{a_2}\right). \]

4. Steps 2 and 3 are repeated to obtain \( v^n_i \), \( i = 1, 2, \ldots, m \) and \( n = 1, 2, \ldots, N \).

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 \( v^n_t = v^n_i \), the conditional probability density function of \( v^n_{t+1} \) is given by

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

Using the fast Fourier cosine series algorithm [see (3.4)], \( g(x|v^n_i) \) can be approximated by

\[ g(x|v^n_i) \approx \sum_{k=0}^{N} F_k \cos\left(\frac{k x - A}{B-A} \pi\right), \quad \text{for } x \in [A, B], \]
where the discrete Fourier cosine series coefficients are given by

\[ F_k = \frac{2}{B - A} \Re \left\{ f(i \frac{k\pi}{B - A}; t_n, \Delta t, v^n_i) \exp(-i \frac{kA\pi}{B - A}) \right\}. \]

The corresponding conditional cumulative function over the interval \([C, D]\) is given by

\[ P[C < v_{n+1} \leq D|v^n_i] = \int_C^D g(x|v^n_i) \, dx \]

\[ \approx \sum_{k=0}^{N'} F_k \int_C^D \cos(k \frac{x - A}{B - A}) \, dx \]

\[ = \sum_{k=0}^{N'} \frac{2}{B - A} \Re \left\{ f(i \frac{k\pi}{B - A}; t_n, \Delta t, v^n_i) \exp(-i \frac{kA\pi}{B - A}) \right\} \varphi_0(k), \]

(3.6)

where

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

In summary, 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 \( C^{n+1} \) denote the \((m+1)\)-dimension column vector \((C^{n+1}_1, C^{n+1}_2, \ldots, C^{n+1}_{m+1})^T\), where

\[ C^{n+1}_i = \begin{cases} 0, & \text{for } i = 1, \\ \frac{v^{n+1}_{i+1} + v^{n+1}_i}{2}, & \text{for } i = 2, 3, \ldots, m, \\ \infty, & \text{for } i = m+1. \end{cases} \]

The transition probability from node \( v^n_i \) at \( t_n \) to \( v^{n+1}_j \) at \( t_{n+1} \) is given by

\[ p^n_{ij} = P[C^{n+1}_j < v_{n+1} \leq C^{n+1}_{j+1}|v^n_i], \quad \text{for } i, j = 1, 2, \ldots, m, \]

(3.7)

where \( P[C^{n+1}_j < v_{n+1} \leq C^{n+1}_{j+1}|v^n_i] \) is computed using (3.6).

2. By repeating step 1 for each \( t_n, \, n = 1, 2, \ldots, N \), we obtain the transition probability matrix

\[ P^n = [p^n_{ij}]_{m \times m} \]

of the willow tree for the instantaneous variance at \( t_n \).

3. At \( t_0 \), the transition probability matrix \( P^0 \) is reduced to the row vector \([p^0_j]\)_{1 \times m}, where

\[ p^0_j = P[C^1_j < v_1 \leq C^1_{j+1}|v^1_0], \quad \text{for } j = 1, 2, \ldots, m. \]

Once the willow tree structure of the instantaneous variance has been constructed, the corresponding value of \( \text{VIX}_t \) can be computed by [see (2.4)]

\[ \text{VIX}^n_i = \sqrt{a_0(\tau) + a_1(\tau)v^{n}_i} \times 100, \quad \text{for } i = 1, 2, \ldots, m \quad \text{and} \quad n = 1, 2, \ldots, N. \]

(3.8)
The transition probability from node VIX$^n_i$ at $t_n$ to node VIX$^{n+1}_j$ at $t_{n+1}$ is $p^n_{ij}$ [see (3.7)], which should be same as that of the instantaneous variance.

As a remark, under the assumption of zero jump in the instantaneous variance process, 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 (Wang and Xu, 2018) without the necessity of finding numerical approximation values using the Fourier cosine series expansion.

Next, we show how to price VIX derivatives under the 1/2-model and 3/2-model based on the willow tree structure constructed using Algorithms 1 and 2.

3.2 Pricing VIX derivatives under the 1/2-model

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 VIX$_{t_N} = VIX^N_i$ at maturity date $T$, the terminal value of the VIX futures is

$$F^n_i = VIX^N_i,$$

and the terminal value of the European VIX call option is

$$c^n_i = \max(VIX^N_i - K, 0).$$

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

$$F^n_i = \sum_{j=1}^m p^n_{ij} F^{n+1}_j.$$

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

$$c^n_i = e^{-r\Delta t} \sum_{j=1}^m p^n_{ij} c^{n+1}_j.$$
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

\[
c_0 = e^{-r \Delta t} \sum_{j=1}^{m} p_j^0 c_j^1.
\]

For the American VIX call option, the early exercise feature is incorporated by following the 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:

\[
C^n_i = \max \{ e^{-r \Delta t} \sum_{j=1}^{m} p^n_{ij} C^{n+1}_j, (\text{VIX}_i^n - K)^+ \}.
\]

3.3 Pricing VIX derivatives under the 3/2-model

As mentioned earlier, the dynamic equation of \( v_t \) in the 3/2-model can be reduced to that of the 1/2-model using the reciprocal transformation \( \hat{v} = \frac{1}{v_t} \). However, since closed form formula for \( \mathbb{E}^Q_t[\frac{1}{v_t}] \) does not exist, so we employ numerical integration to estimate the value of \( \mathbb{E}^Q_t[\frac{1}{v_t}] \).

In order to estimate the value of VIX under the 3/2-model, we first construct a willow tree structure with \( m \) spacial nodes at each time point of the reciprocal of the instantaneous variance \( \hat{v}_t \) over the extended time interval \([0, T + \tau]\), where \( T \) is the maturity date of VIX derivative. The time interval \([0, T + \tau]\) is divided into \( N_1 \) subintervals with uniform time step \( \Delta t = \frac{T + \tau}{N_1} \), where \( t_0 = 0 < t_1 < t_2 < \cdots < t_N = T < t_{N+1} < \cdots < t_{N_1} = T + \tau \), and \( t_i = i \Delta t \).

According to (2.5), we employ the trapezoidal rule in computing the conditional expectation by taking \( \frac{1}{2} \left( \frac{1}{\hat{v}_{t_k}} + \frac{1}{\hat{v}_{t_{k+1}}} \right) \) to approximate the value of \( \frac{1}{\hat{v}_t} \) over the integral interval \( u \in [t_k, t_{k+1}] \). The approximation of \( \text{VIX}_t^n \) under the 3/2-model is given by

\[
\text{VIX}_t^n \approx \left\{ \frac{1}{2} \sum_{k=n}^{n+N} \mathbb{E}^{Q}_{t_n} \left[ \frac{1}{\hat{v}_{t_k}} \right] + \mathbb{E}^{Q}_{t_n} \left[ \frac{1}{\hat{v}_{t_{k+1}}} \right] \right\} \Delta t + 2 \lambda [\bar{\mu} - (\mu S + \rho J \mu_v)] \times 100^2,
\]

where \( \bar{N} = N_1 - N - 1 \). Given \( \hat{v}_{t_n} = \hat{v}_i^n \) at \( t_n \), the corresponding value of VIX\(_{t_n} \) can be obtained by

\[
(\text{VIX}_t^n)^2 = \left\{ \frac{1}{2} \sum_{k=n}^{n+N} \left[ \mathbb{E}^{Q}_{t_n} \left[ \frac{1}{\hat{v}_{t_k}} \right] + \mathbb{E}^{Q}_{t_n} \left[ \frac{1}{\hat{v}_{t_{k+1}}} \right] \right] \Delta t + 2 \lambda [\bar{\mu} - (\mu S + \rho J \mu_v)] \right\} \times 100^2. \tag{3.9}
\]

According to (3.9), there are \( \bar{N} + 1 \) expectations, \( \mathbb{E}^{Q}_{t_n} \left[ \frac{1}{\hat{v}_{t_k}} \right], \ k = n+1, n+2, \cdots, n + \bar{N} + 1 \), to be computed.
We define the $m$-dimensional column vector $\hat{v}_n$, $n = 1, 2, \cdots, N_1$, where $\hat{v}_n = (\frac{1}{\sqrt{1}}, \frac{1}{\sqrt{2}}, \cdots, \frac{1}{\sqrt{m}})^T$.

Given $\hat{v}_{t_n} = \hat{v}_i^n$, the transition probability from node $\hat{v}_i^n$ to $m$ nodes at the next time step is the $i^{th}$ row of the transition matrix $\hat{P}_n$. Let $\hat{p}_i^n = (p_{i1}^n, p_{i2}^n, \cdots, p_{im}^n)$, we then have

$$E^Q[\frac{1}{\hat{v}_{t_{n+1}}} | \hat{v}_i^n] \approx \hat{p}_i^n \hat{v}^{n+1}.$$  

The transition probabilities from node $\hat{v}_i^n$ to $m$ nodes at time $t = t_{n+2}$ is

$$\hat{p}_i^{n+1} = \hat{p}_i^n \hat{P}_{n+1}^{n+1},$$

so that

$$E^Q[\frac{1}{\hat{v}_{t_{n+2}}} | \hat{v}_i^n] = \hat{p}_i^{n+1} \hat{v}^{n+2}. $$

By applying the forward iteration, we have

$$\hat{p}_i^{n+k} = \hat{p}_i^{n+k-1} \hat{P}_{n+k}, \quad \text{for } k = 2, 3, \cdots, N + 1.$$

$$E^Q[\frac{1}{\hat{v}_{t_{n+k}}} | \hat{v}_i^n] = \hat{p}_i^{n+k-1} \hat{v}^{n+k}, \quad \text{for } k = 2, 3, \cdots, N + 1.$$

With the estimated values of $E^Q[\frac{1}{\hat{v}_{t_{n+k}}} | \hat{v}_i^n]$, $k = 1, 2, \cdots, N + 1$ and $E^Q[\frac{1}{\hat{v}_i^n}] = \frac{1}{\hat{v}_i^n}$, we can obtain the corresponding value of VIX$_i^n$ at node $\hat{v}_i^n$ under the 3/2-model. Subsequently, we construct the willow tree structure of VIX under the 3/2-model. By following similar procedures in Algorithm 3, we can price VIX futures and option based on the willow tree structure of VIX under the 3/2-model.

4 Numerical tests

This section presents numerical results to 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 $\tau = \frac{1}{12}$, same as convention adopted in CBOE. In our construction of the willow tree structure, the number of spacial 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 \( \text{VIX}^2 \) 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 (2.5). 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 algorithm (WT) and Monte Carlo simulation method (MC) 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 VIX futures and the European call option as an additional benchmark. 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 with those by MC. In addition to 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.

<table>
<thead>
<tr>
<th>$r$</th>
<th>$v_0$</th>
<th>$\eta$</th>
<th>$\theta$</th>
<th>$\sigma_v$</th>
<th>$\mu_S$</th>
<th>$\sigma_S$</th>
<th>$\lambda$</th>
<th>$\rho_J$</th>
<th>$\mu_v$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0319</td>
<td>0.0076</td>
<td>3.46</td>
<td>0.008</td>
<td>0.14</td>
<td>-0.0865</td>
<td>0.0001</td>
<td>0.47</td>
<td>-0.38</td>
<td>0.05</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Maturity (month)</th>
<th>Price</th>
<th>CPU time (second)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WT</td>
<td>MC (standard error)</td>
</tr>
<tr>
<td>1</td>
<td>12.0765</td>
<td>12.0895 (0.0269)</td>
</tr>
<tr>
<td>2</td>
<td>12.3793</td>
<td>12.3816 (0.0339)</td>
</tr>
<tr>
<td>3</td>
<td>12.6240</td>
<td>12.6426 (0.0379)</td>
</tr>
<tr>
<td>4</td>
<td>12.8185</td>
<td>12.8209 (0.0396)</td>
</tr>
<tr>
<td>5</td>
<td>12.9712</td>
<td>12.9778 (0.0411)</td>
</tr>
<tr>
<td>6</td>
<td>13.0899</td>
<td>13.0938 (0.0417)</td>
</tr>
<tr>
<td>7</td>
<td>13.1816</td>
<td>13.1766 (0.0420)</td>
</tr>
<tr>
<td>8</td>
<td>13.2520</td>
<td>13.2582 (0.0424)</td>
</tr>
<tr>
<td>9</td>
<td>13.3058</td>
<td>13.3221 (0.0426)</td>
</tr>
<tr>
<td>10</td>
<td>13.3467</td>
<td>13.3615 (0.0427)</td>
</tr>
</tbody>
</table>

In our numerical experiments, we used the same set of parameters of the 1/2-model showed 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 simulation (MC) method and direct numerical integration (NI). VIX futures values obtained by WT are very close to those obtained using NI, and the difference between VIX futures prices by WT and MC is typically less than the standard error of MC. Comparing the CPU times required by WT and MC, the former is typically much less, even for long-maturity futures. This reveals WT to be accurate and more computationally efficient than MC when pricing VIX futures. Table 3 presents
Table 3: Pricing results of 3-month European VIX call option under the 1/2-model.

<table>
<thead>
<tr>
<th>Strike price</th>
<th>Price</th>
<th>Price (standard error)</th>
<th>CPU time (second)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WT</td>
<td>MC</td>
<td>WT</td>
</tr>
<tr>
<td>10.5</td>
<td>2.2310</td>
<td>2.2391 (0.0364)</td>
<td>2.2353</td>
</tr>
<tr>
<td>11.0</td>
<td>1.8646</td>
<td>1.8659 (0.0356)</td>
<td>1.8690</td>
</tr>
<tr>
<td>11.5</td>
<td>1.5572</td>
<td>1.5527 (0.0345)</td>
<td>1.5615</td>
</tr>
<tr>
<td>12.0</td>
<td>1.3082</td>
<td>1.3057 (0.0335)</td>
<td>1.3125</td>
</tr>
<tr>
<td>12.5</td>
<td>1.1121</td>
<td>1.1218 (0.0328)</td>
<td>1.1164</td>
</tr>
<tr>
<td>13.0</td>
<td>0.9607</td>
<td>0.9574 (0.0312)</td>
<td>0.9649</td>
</tr>
<tr>
<td>13.5</td>
<td>0.8446</td>
<td>0.8428 (0.0304)</td>
<td>0.8487</td>
</tr>
<tr>
<td>14.0</td>
<td>0.7549</td>
<td>0.7590 (0.0294)</td>
<td>0.7590</td>
</tr>
</tbody>
</table>

the pricing results of 3-month European VIX call option with varying strike prices. Similar high level of numerical performance 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). Since their set of parameters, the jump in the index process is too small and there is no jump in the instantaneous variance process, so we add the jump components 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.

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

<table>
<thead>
<tr>
<th>r</th>
<th>v_0</th>
<th>η</th>
<th>θ</th>
<th>σ_v</th>
<th>μ_S</th>
<th>σ_S</th>
<th>λ</th>
<th>ρ_J</th>
<th>μ_v</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.0076</td>
<td>26.3189</td>
<td>0.0935</td>
<td>9.2499</td>
<td>-0.0865</td>
<td>0.0001</td>
<td>0.47</td>
<td>-0.38</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Tables 5 and 6 present the pricing results of VIX futures and 3-month European VIX call 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
Table 5: Pricing results of VIX futures under the 3/2-model.

<table>
<thead>
<tr>
<th>Maturity (month)</th>
<th>Price WT</th>
<th>Price MC (standard error)</th>
<th>CPU time (second) WT</th>
<th>CPU time (second) MC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.3001</td>
<td>12.2853 (0.0098)</td>
<td>3.924</td>
<td>3904.024</td>
</tr>
<tr>
<td>2</td>
<td>13.0004</td>
<td>13.0029 (0.0166)</td>
<td>7.265</td>
<td>3891.672</td>
</tr>
<tr>
<td>3</td>
<td>13.7153</td>
<td>13.7137 (0.0231)</td>
<td>10.472</td>
<td>3899.843</td>
</tr>
<tr>
<td>4</td>
<td>14.4272</td>
<td>14.4311 (0.0296)</td>
<td>14.041</td>
<td>3913.363</td>
</tr>
<tr>
<td>5</td>
<td>15.1376</td>
<td>15.1821 (0.0341)</td>
<td>16.994</td>
<td>3925.915</td>
</tr>
<tr>
<td>6</td>
<td>15.8232</td>
<td>15.7861 (0.0407)</td>
<td>20.259</td>
<td>3982.429</td>
</tr>
<tr>
<td>7</td>
<td>16.4479</td>
<td>16.3506 (0.0439)</td>
<td>23.737</td>
<td>3998.580</td>
</tr>
<tr>
<td>8</td>
<td>17.0000</td>
<td>16.9615 (0.0475)</td>
<td>27.032</td>
<td>4012.579</td>
</tr>
<tr>
<td>9</td>
<td>17.4756</td>
<td>17.4092 (0.0509)</td>
<td>30.488</td>
<td>4013.147</td>
</tr>
<tr>
<td>10</td>
<td>17.8673</td>
<td>17.7810 (0.0510)</td>
<td>34.239</td>
<td>4042.311</td>
</tr>
</tbody>
</table>

Table 6: Pricing results of 3-month European VIX call option under the 3/2-model.

<table>
<thead>
<tr>
<th>Strike price</th>
<th>Price WT</th>
<th>Price MC (standard error)</th>
<th>CPU time (second) WT</th>
<th>CPU time (second) MC</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.5</td>
<td>3.1780</td>
<td>3.1982 (0.0231)</td>
<td>10.552</td>
<td>4115.228</td>
</tr>
<tr>
<td>11.0</td>
<td>2.6968</td>
<td>2.6880 (0.0227)</td>
<td>10.752</td>
<td>4101.472</td>
</tr>
<tr>
<td>11.5</td>
<td>2.2404</td>
<td>2.2491 (0.0219)</td>
<td>10.682</td>
<td>4084.472</td>
</tr>
<tr>
<td>12.0</td>
<td>1.8253</td>
<td>1.8126 (0.0217)</td>
<td>10.817</td>
<td>4098.624</td>
</tr>
<tr>
<td>12.5</td>
<td>1.4631</td>
<td>1.4771 (0.0204)</td>
<td>11.183</td>
<td>4083.217</td>
</tr>
<tr>
<td>13.0</td>
<td>1.1585</td>
<td>1.1503 (0.0192)</td>
<td>10.724</td>
<td>4091.092</td>
</tr>
<tr>
<td>13.5</td>
<td>0.9095</td>
<td>0.9024 (0.0174)</td>
<td>10.695</td>
<td>4078.079</td>
</tr>
<tr>
<td>14.0</td>
<td>0.7103</td>
<td>0.7216 (0.0156)</td>
<td>10.713</td>
<td>4092.423</td>
</tr>
</tbody>
</table>
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 \( \sigma_v \).](image)

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. These phenomena are similar to those reported in Yuen et al. (2015) for variance swaps and the same explanation of the pricing behavior 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 the 1/2-model and 3/2-model. The other model parameters used in our numerical tests remain unchanged.

![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](image)

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 has positive correlation with the jump intensity under the 1/2-model and 3/2-model. The price of the European VIX call option increases quite significantly with increasing jump intensity.

### 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.

### 4.6 American VIX option

In Figure 6, we show the plots of the prices of European and American VIX call option with varying values of $VIX_0$ under the 1/2-model and 3/2-model. 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 4: Implied volatility of 3-month European VIX call option with respect to strike price (left) and with respect to maturity with strike price $K = 13$ (right) under the 1/2-model.

Figure 5: Implied volatility of 3-month European VIX call option with respect to strike price (left) and with respect to maturity with strike price $K = 13$ (right) under the 3/2-model.
Figure 6: Plots of the European and American VIX call option against VIX₀ with strike price \( K = 13 \) and maturity \( T = 3 \) months under the 1/2-model and 3/2-model.

5 Conclusion

We present the construction of the willow tree algorithms for pricing VIX derivatives under the 1/2-model and 3/2-model with jumps in the S&P index and its instantaneous variance processes. In the construction of the willow tree for the instantaneous variance, it is necessary to find the positions of the nodes on the willow tree structure and the transition probability matrices between nodes of successive time steps using the information of the first four order moments of the instantaneous variance. Since the characteristic function of the instantaneous variance under the CIR process is available in closed form, 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 computational 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 to those observations reported in other papers on related variance and VIX products.
References


Appendix A —— Johnson curve transform

According to the Johnson curve transform (Johnson, 1949), 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 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 $S_U$): $Z = a + b \sinh^{-1}\left(\frac{v - c}{d}\right)$
- bounded system (or $S_B$): $Z = a + b \ln\left(\frac{v - c}{c + d - v}\right), \quad c < v < c + d$

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 to decide 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).