Probability, Uncertainty and Quantitative Risk
Vol. 10, No. 3, 2025, 385–404

https://www.aimsciences.org/PUQR
DOI:10.3934/puqr.2025017

Asymptotic smiles for an affine jump-diffusion model

Nian Yao[1], Junfeng Lin[1], Zhiqiu Li[2,*]

[1]*School of Mathematical Sciences, Shenzhen University, Shenzhen* 518060, *China*
[2]*Department of Mathematics, Florida State University, Tallahassee, FL* 32304, *USA*

Email: yaonian@szu.edu.cn, Little_wind@163.com, zhiqiuli.edu@gmail.com

Abstract In this paper, we study the asymptotic behaviors of implied volatility in an affine jump-diffusion model. By assuming that log stock prices under the risk-neutral measure follow an affine jump-diffusion model, we show that an explicit form of the moment-generating function for log stock price can be obtained by solving a set of ordinary differential equations. A large-time large deviation principle for log stock prices is derived by applying the Gärtner–Ellis theorem. We characterize the asymptotic behaviors of implied volatility in the large-maturity and large-strike regimes using the rate function in the large deviation principle. The asymptotics of the implied volatility for fixed-maturity, large-strike and small-strike regimes are also studied. Numerical results are provided to validate the theoretical work.

Keywords Affine jump-diffusion model, Large deviations, Implied volatility, Asymptotics

2020 Mathematics Subject Classification 60F10, 60H10

1. Introduction

Option pricing problems have been extensively studied when the underlying follows a jump-diffusion process. In the 1970s, Merton [22] proposed a jump-diffusion process and assumed that the jump size follows a log-normal distribution. They showed a European option can be written as a weighted sum of Black-Scholes European option prices. Later Kou [20] assumed the jump size follows a double exponential distribution and provided a closed-form solution to the problem. However, option pricing problems in which the underlying asset follows an affine jump-diffusion point process or has Hawkes-type jumps have been far less studied. This is because a closed-form solution for option pricing is no longer available.

Implied volatility problems have been studied in the mathematical finance literature. It represents a forward-looking measure of volatility for a specific underlying asset, whereas a historical or realized volatility is backward-looking. The implied volatility of an actively traded option on an underlying asset is often used by traders, to estimate the appropriate implied volatility for other similar options on the same underlying. Forde and Jacquier [11] studied the large-time asymptotic behaviors of implied volatility for European call and put options under the

Received 7 March 2024; Accepted 4 June 2025; Early access 4 September 2025
*Corresponding author

Heston stochastic volatility model. Similar work has been extended to other stochastic volatility models, such as the SABR and CEV-Heston models (Forde and Pogudin [12]), a class of affine stochastic volatility models. (Jacquier et al. [17]) and multivariate Wishart stochastic volatility models (Alfonsi et al. [4]). Kanaya and Otsu [19] established large deviation and moderate deviation principles for realized volatility estimators based on high-frequency data, providing rigorous asymptotic results under infill asymptotics in a stochastic volatility setting. Gatheral et al. [14] obtained the first- and second-order terms in the short-time asymptotics of the European call option prices using an expansion of the transition density function of a one-dimensional time-inhomogeneous diffusion. Mijatović and Tankov [23] analyzed the behavior of the implied volatility smile for options close to expiration in the exponential Lévy class of asset price models with jumps. Jacquier and Roome [18] provided a complete characterization of the large-maturity forward implied volatility smile in the Heston model. Djellout et al. [7] derived large deviation principles for the realized (co-)volatility vector in the high-frequency framework, particularly focusing on multivariate diffusion models and associated estimators of integrated covariance. Caravenna and Corbetta [5] used the large deviation techniques to determine the asymptotic shape of the implied volatility surface in any regime of small maturity or extreme log-strike for a stochastic volatility model. Yao et al. [24] studied the short-term asymptotic behaviors of implied volatility for conditional Asian options in local volatility models using the large deviation theory. Feng et al. [10] investigated the large deviation principles for the realized Laplace transform of volatility, deriving both pointwise and functional LDPs and highlighting their applications in inference and testing problems.

Affine point process (also known as affine jump-diffusion model or affine point process driven by a jump-diffusion) is a point process whose event arrival intensity is governed by an affine jump-diffusion (Duffie et al. [8]). An affine point process can be further characterized as self-exciting or mutual-exciting. A self-exciting process means that a jump increases the likelihood of future jumps in the same component, whereas a mutual-exciting process increases the jump intensity across other components as well. Due to their computational tractability, affine point processes have been applied in various financial and economic fields, such as Errais et al. [9]; Zhang et al. [25]; Aït-Sahalia et al. [1]; Aït-Sahalia and Hurd [2]; Aït-Sahalia et al. [3]; Zhang and Glynn [26]; Gao and Zhu [13].

In this paper, we employ large deviation technique to study the asymptotic behaviors of the implied volatility of an affine jump-diffusion model. It is worth mentioning that the approach proposed in Forde and Jacquier [11] for the Heston model is not directly applicable here in terms of characterizing the corresponding large deviation rate function due to the complexity of the affine jump-diffusion model. That is the key point to characterize the asymptotic behaviors of the implied volatility in the large-maturity and large-strike regimes. In previous studies, the focus was on the large deviation of $\frac{N_t}{t}$ or $\frac{1}{t}\sum_{i=1}^{N_t} Y_i$ or $\frac{1}{t}\int_0^t \lambda_s \mathrm{d}s$, but our focus is on the logarithmic return of the underlying stock price. That is more difficult to handle because the price of the stock includes both $\frac{1}{t}\sum_{i=1}^{N_t} Y_i$ and $\frac{1}{t}\int_0^t \lambda_s \mathrm{d}s$. The main contributions of this paper are summarized as follows:

(1) We establish the asymptotics of the implied Black-Scholes volatility in the large-maturity and large-strike regimes for an affine jump-diffusion model, which does not exist in the current literature (Theorem 2). The key is to derive the corresponding large deviation principle for the log return of the underlying stock price.

(2) We derive the asymptotics of implied volatility in the fixed-maturity, large-strike, and small-

strike regimes (Theorem 3).

The remainder of the paper is organized as follows. In Section 2.1, we introduce an affine jump-diffusion model and express the exact moment-generating function of the model as the solutions of a set of ordinary differential equations (ODEs). In Section 2.2, we obtain the large-time large deviation principle of the logarithmic return of the stock price using the risk-neutral measure. In Section 2.3, we characterize the asymptotic behavior of implied volatility in the large-maturity and large-strike regimes. In Section 2.4, we study the asymptotic of implied volatility for fixed-maturity, large-strike, or small-strike regimes. Numerical examples are presented in Section 3 to validate the theoretical work. In Section 4, we provide proofs of the main results. Finally, Section 5 presents the conclusion. The framework of this paper is shown in Figure 1.

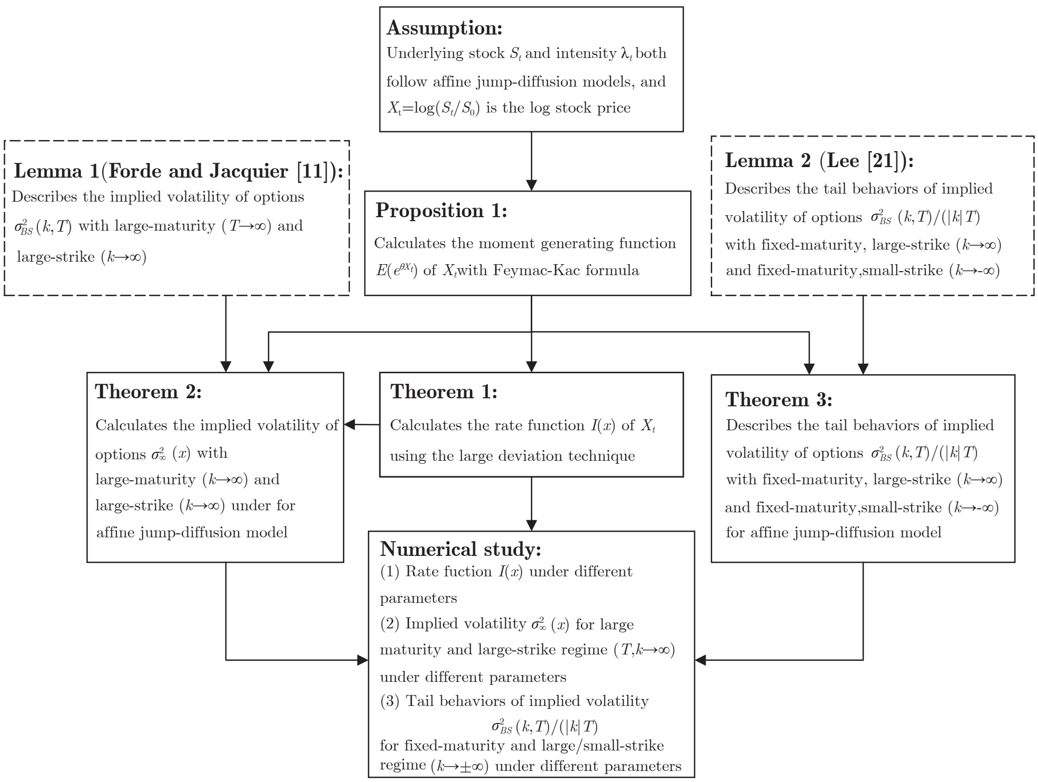

Figure 1 Framework of this paper

2. The main results

2.1 Affine jump-diffusion model

We consider a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$. The filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ is generated by the Brownian motion $W_t^{\mathbb{P}}$ and the point process N_t, i.e.,

$$\mathcal{F}_t = \sigma\left(W_s^{\mathbb{P}}, N_s : 0 \leqslant s \leqslant t\right).$$

We assume the underlying stock S_t under the risk-neutral measure $\mathbb{Q}$ follows an affine jump-diffusion model:

$$\frac{\mathrm{d}S_t}{S_{t-}} = \sigma \mathrm{d}W_t^{\mathbb{Q}} + \mathrm{d}J_t - \lambda_t^N \mu_Y \mathrm{d}t, \tag{2.1}$$

where

$$J_t = \sum_{i=1}^{N_t} (e^{Y_i} - 1),$$

where Y_i are i.i.d. random jump sizes independent of N_t and $W_t^{\mathbb{Q}}$ and $\mu_Y = \mathbb{E}[e^Y] - 1$. Y_i follows a probability distribution $\mathbb{Q}(\mathrm{d}y)$. In (2.1), $\sigma \mathrm{d}W_t^{\mathbb{Q}}$ and $\mathrm{d}J_t$ describe instantaneous returns due to normal price fluctuations and due to the event-driven jumps, respectively. The shift part $-\lambda_t^N \mu_Y \mathrm{d}t$ makes $\frac{\mathrm{d}S_t}{S_{t-}}$ a martingale.

A temporal point process is a stochastic process that plays a crucial role in the analysis of the observed patterns of points, where the points represent the locations of some underlying object of study. The Hawkes process is perhaps the most parsimonious univariate self-exciting point process whose conditional intensity function is linear and increasing. The linear Hawkes process was first introduced by Hawkes[15,16]. It naturally generalizes the Poisson process and can capture the self-exciting property and clustering effect. The Hawkes process is a versatile model that is amenable to statistical analysis, as evident by its wide application in neuroscience, genome analysis, criminology, social networks, seismology, insurance, finance, and many other related fields (e.g., see Zhu [27] and references therein).

We assume that N_t is an affine point process, which has been introduced in the Introduction and referenced in the preceding discussion that has intensity $\lambda_t^N = \alpha + \beta\lambda_t$ at $t > 0$ and λ_t satisfies the dynamics:

$$\mathrm{d}\lambda_t = b(c - \lambda_t)\mathrm{d}t + \sigma\sqrt{\lambda_t}\mathrm{d}B_t + a\mathrm{d}N_t, \tag{2.2}$$

where B_t is a standard BM. We make the following basic assumptions that are required for modeling an affine jump-diffusion model (Zhu [28]):

Assumption 1 (1) $a, b, c, \alpha, \beta, \sigma > 0$.

(2) $b > a\beta$. *This condition indicates that there exists a unique stationary process* λ^∞ *which satisfies the dynamics (2.2).*

(3) $2bc \geqslant \sigma^2$. *This condition ensures that the upward drift term is large enough that* λ_t *cannot reach 0, that is,* $\lambda_t > 0$ *with probability 1.*

In addition, we assume that B_t is independent of $W_t^{\mathbb{Q}}$. It is worth noting that the point process N_t reduces to a linear Hawkes process with an exponential decay kernel when the Brownian motion term $B_t = 0$. If $adN_t = 0$, then the process λ_t reduces to a Cox-Ingersoll-Ross process. The log stock price under the risk-neutral measure via $S_t = S_0 e^{X_t}$ is given by

$$X_t = -\frac{1}{2}\sigma^2 t + \sigma W_t^{\mathbb{Q}} - \mu_Y \int_0^t \lambda_s^N \mathrm{d}s + \sum_{i=1}^{N_t} Y_i. \tag{2.3}$$

We can write $N_t = \sum_{i=1} \mathbb{1}_{\{T_i \leqslant t\}}$ and $L_t = \sum_{i \geqslant 1} Y_i \mathbb{1}_{\{T_i \leqslant t\}}$ where T_n is the n-th jump time of N_t. The two-dimensional process (λ, L) is Markovian on $D = \mathbb{R}_+ \times \mathbb{R}$ with an infinite generator given by

$$\mathcal{L}f(\lambda, L) = b(c - \lambda)\frac{\partial f}{\partial \lambda} + \frac{1}{2}\sigma^2\lambda\frac{\partial^2 f}{\partial \lambda^2} + (\alpha + \beta\lambda)\int_{\mathbb{R}} (f(\lambda + a, L + y) - f(\lambda, L))\mathbb{Q}(\mathrm{d}y) \tag{2.4}$$

for a given function $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ with twice continuously differentiable and for all $\lambda \in \mathbb{R}_+$, $|\int_{\mathbb{R}} f(L+y, \lambda + a)\mathbb{Q}(\mathrm{d}y)| < \infty$.

We compute the moment-generating function for X_t. The result is summarized in the following Proposition 1.

Proposition 1 *The moment generating function for X_t is given by*

$$\mathbb{E}[e^{\theta X_t}] = e^{(-\frac{1}{2}\theta\sigma^2 + \frac{1}{2}\theta^2\sigma^2 - \theta\mu_Y\alpha)t + D(t;\Theta)\lambda + \theta L + F(t;\Theta)}, \tag{2.5}$$

where $\theta \in \mathbb{R}$, $\Theta = (\theta_1, \theta_2, \theta_3) = (-\theta\mu_Y\beta, 0, \theta) \in \mathbb{R}^3$ and $D(t;\Theta)$, $F(t;\Theta)$ satisfy the following ODEs,

$$\begin{cases} D'(t;\Theta) + bD(t;\Theta) - \dfrac{1}{2}\sigma^2 D^2(t;\Theta) - \beta \displaystyle\int_{\mathbb{R}} (e^{D(t;\Theta)a + \theta y} - 1)\mathbb{Q}(\mathrm{d}y) + \theta\mu_Y\beta = 0, \\ F'(t;\Theta) - bcD(t;\Theta) - \alpha \displaystyle\int_{\mathbb{R}} (e^{D(t;\Theta)a + \theta y} - 1)\mathbb{Q}(\mathrm{d}y) = 0, \\ D(0;\Theta) = 0, F(0;\Theta) = 0. \end{cases} \tag{2.6}$$

Proof A detailed proof can be found in Section 4.1. □

Remark 1 *We now consider the following system of nonlinear ODEs (2.6), where:*

- $b, c, \sigma, a, \alpha, \beta > 0$ *are constants;*
- $\theta \in \mathbb{R}$ *is a parameter;*
- $\mathbb{Q}$ *is a probability measure for the jump size* Y*;*
- $\mu_Y = \int_{\mathbb{R}} (e^y - 1)\,\mathbb{Q}(\mathrm{d}y)$.

To establish the existence and uniqueness of a solution $(D(t), F(t))$ to the above system in a finite time interval $[0, T]$. Let us define:

$$G(D) := -bD + \frac{1}{2}\sigma^2 D^2 + \beta \int_{\mathbb{R}} \left(e^{aD+\theta y} - 1\right)\mathbb{Q}(\mathrm{d}y) - \theta\mu_Y\beta,$$

$$H(D) := bcD + \alpha \int_{\mathbb{R}} \left(e^{aD+\theta y} - 1\right)\mathbb{Q}(\mathrm{d}y).$$

Then the system (2.6) *becomes:*

$$D'(t) = -G(D(t)), \quad F'(t) = H(D(t)), \quad D(0) = 0, \quad F(0) = 0.$$

We now verify the standard conditions under the Picard–Lindelöf theorem (also known as the Cauchy–Lipschitz theorem), which guarantees the local existence and uniqueness of solutions for ODE systems.

Continuity: *The functions $G(D)$ and $H(D)$ are continuous in D provided that the moment-generating function*

$$\int_{\mathbb{R}} e^{\theta y}\,\mathbb{Q}(\mathrm{d}y)$$

is finite for the given parameter $\theta \in \mathbb{R}$. This ensures that the integral terms in both equations are well defined and smooth with respect to D.

Local Lipschitz Continuity: *The mapping $D \mapsto \int_{\mathbb{R}} e^{aD+\theta y}\,\mathbb{Q}(\mathrm{d}y)$ is C^∞ in any bounded domain of D. Hence, both $G(D)$ and $H(D)$ are locally Lipschitz continuous on bounded intervals.*

By the Picard–Lindelöf theorem, under the assumption that

$$\int_{\mathbb{R}} e^{\theta y}\,\mathbb{Q}(\mathrm{d}y) < \infty,$$

the ODE system (2.6) admits a unique solution $(D(t), F(t))$ on a finite interval $[0, T]$. Furthermore, the solution continuously depends on the parameter θ and the initial conditions.

2.2 Large deviation principle for X_t

We derive the following theorem which describes the large-time large deviation asymptotic behaviors of the log stock price. This result will be used later to derive the asymptotics for option pricing and implied volatility smiles in the regime where the maturity is large and the log-moneyness is of the same order as the maturity. We refer the reader to Dembo and Zeitouni [6] for a formal definition of the large deviation principle and the applications.

Theorem 1 *(Large Deviation Principle for X_t). Under Assumption 1, $\mathbb{Q}(\frac{1}{t}X_t \in \cdot)$ satisfies a large deviation principle on $\mathbb{R}$ with the rate function:*

$$I(x) = \sup_{\theta\in\mathbb{R}} \{\theta x - \Lambda(\theta)\}, \tag{2.7}$$

where

$$\Lambda(\theta) = \left(\frac{1}{2}\sigma^2\theta^2 - \left(\frac{1}{2}\sigma^2 + \mu_Y\alpha\right)\theta + bcD(\theta) + \alpha\left(e^{aD(\theta)}\mathbb{E}[e^{\theta Y}] - 1\right)\right)$$

and $D(\theta)$ is the smaller solution of the equation

$$-bD + \frac{1}{2}\sigma^2 D^2 + \beta(\mathbb{E}[e^{aD+\theta Y}] - 1) - \theta\mu_Y\beta = 0, \tag{2.8}$$

if solution exists. Otherwise, $\Lambda(\theta) = +\infty$.

Proof We present the proof in Section 4.2. □

2.3 Asymptotics of implied volatility in the large-maturity and large-strike regime

We use the rate function in the large deviation principle for X_t, referring to Forde and Jacquier [11], to characterize the asymptotic behaviors of implied volatility in large-maturity and large-strike regimes, which is shown in Theorem 2.

Consider an European call option with maturity T and strike K is given as

$$C(K, T) := \mathrm{d}(T)\mathbb{E}\left[(S_T - K)^+\right],$$

where S_T is the underlying stock price at maturity T and $d(T)$ is the discount factor. Note that the corresponding put option price $P(K, T)$ can be found straightforwardly using call-put parity. $C(K, T)$ indicates the dependence on maturity T and strike K. Let $F_0 = \mathbb{E}S_T$ is the forward price of the underlying stock. For a given F_0, the logarithmic moneyness k is related to the strike by

$$k := \log(K/F_0), \tag{2.9}$$

so $K(k) = F_0 e^k$ is the strike at log moneyness k. The Black-Scholes implied volatility with logarithmic moneyness k and at maturity T is defined as $\sigma_{BS}(k, T)$ that uniquely solves

$$C(K(k), T) = C^{BS}(k, \sigma_{BS}(k, T)), \tag{2.10}$$

where

$$C^{BS}(k,\sigma) = D(T)\left(F_0\Phi(d_+) - K(k)\Phi(d_-)\right) \quad \text{and} \quad d_\pm = \frac{-k}{\sigma\sqrt{T}} \pm \frac{\sigma\sqrt{T}}{2},$$

and Φ is the cumulative distribution function of a standard normal distribution. Similarly, for a European put option, its implied volatility $\sigma_{BS}(k,T)$ uniquely solves

$$P(K(k),T) = P^{BS}(k,\sigma_{BS}(k,T)), \tag{2.11}$$

where

$$P^{BS}(k,\sigma) = D(T)(K(k)\Phi(-d_-) - F_0\Phi(-d_+)).$$

The following lemma describes the large-time asymptotic behavior and implied volatility in the large-maturity and large-strike regime.

Lemma 1 *(Forde and Jacquier [11]) For the model defined in (2.1) under Assumption 1, we obtain the large-time asymptotic behaviors*

$$I(x) - x = \begin{cases} -\lim\limits_{T\to\infty}\frac{1}{T}\log\mathbb{E}[(S_T - S_0e^{xT})^+], & \text{for } x \geqslant x_R, \\ -\lim\limits_{T\to\infty}\frac{1}{T}\log(S_0 - \mathbb{E}[(S_T - S_0e^{xT})^+]), & \text{for } x_L \leqslant x \leqslant x_R, \\ -\lim\limits_{T\to\infty}\frac{1}{T}\log\mathbb{E}[(S_0e^{xT} - S_T)^+], & \text{for } x \leqslant x_L, \end{cases}$$

where $x_L = \Lambda'(0), x_R = \bar{\Lambda}'(0)$, and $\bar{\Lambda}(\theta)$ is $\Lambda(\theta)$ under the share measure $\bar{\mathbb{Q}}$ as

$$\left.\frac{\mathrm{d}\bar{\mathbb{Q}}}{\mathrm{d}\mathbb{Q}}\right|_{\mathcal{F}_t} = \frac{S_t}{S_0} = e^{X_t}.$$

And we have the implied volatility in large-time and large-strike regime

$$\sigma_\infty^2(x) = \lim_{T\to\infty}\sigma_{\mathrm{BS}}^2(xT,T) = \begin{cases} 2(2I(x) - x - 2\sqrt{I^2(x) - xI(x)}), & x \in (-\infty, x_L)\cup(x_R,\infty), \\ 2(2I(x) - x + 2\sqrt{I^2(x) - xI(x)}), & x \in [x_L, x_R]. \end{cases}$$

Applying Lemma 1, we derive the following results for the proposed model.

Theorem 2 *In the joint regime of large-maturity, large-strike with $k = \log(K/S_0)$ $(T\to\infty, |k|\to\infty)$, the implied volatility $\sigma_{\mathrm{BS}}(k,T)$ approaches the limit*

$$\lim_{T\to\infty}\sigma_{\mathrm{BS}}^2(xT,T) = \sigma_\infty^2(x), \tag{2.12}$$

where

$$\sigma_\infty^2(x) = \begin{cases} 2(2I(x) - x - 2\sqrt{I^2(x) - xI(x)}), & x \in (-\infty, x_L)\cup(x_R,\infty), \\ 2(2I(x) - x + 2\sqrt{I^2(x) - xI(x)}), & x \in [x_L, x_R], \end{cases} \tag{2.13}$$

where $I(x)$ is defined in (2.7) and

$$x_L = -\left(\frac{1}{2}\sigma^2 + \mu_Y\alpha\right) + (bc + a\alpha)\frac{\beta\left(\mu_Y - \mathbb{E}[Y]\right)}{a\beta - b} + \alpha\mathbb{E}[Y], \tag{2.14}$$

and

$$x_R = \left(\frac{1}{2}\sigma^2 - \mu_Y\mathbb{E}[e^Y]\alpha\right) + \left(bc + a\mathbb{E}[e^Y]\alpha\right)\frac{\mathbb{E}[e^Y]\beta\left(\mu_Y - \mathbb{E}[\bar{Y}]\right)}{a\mathbb{E}[e^Y]\beta - b} + \mathbb{E}[e^Y]\alpha\mathbb{E}[\bar{Y}], \tag{2.15}$$

where $\bar{Y}$ follows the probability distribution $\frac{e^Y}{\mathbb{E}[e^Y]}\mathrm{d}\mathbb{Q}$.

Proof Please see Section 4.3. □

2.4 Asymptotics of implied volatility in fixed-maturity, large-strike and small-strike regimes

We apply Lee's moment formula (Lee [21]) to derive the asymptotics for the Black-Scholes implied volatility in fixed-maturity, large-strike $(K \to \infty)$ and small-strike $(K \to 0)$ regimes, which is shown in Theorem 3.

Define

$$\tilde{p} := \sup\left\{p : \mathbb{E}^{\mathbb{Q}}[S_T^{1+p}] < \infty\right\}, \tag{2.16}$$

and

$$\tilde{q} := \sup\left\{q : \mathbb{E}^{\mathbb{Q}}[S_T^{-q}] < \infty\right\}. \tag{2.17}$$

The following lemma gives an explicit formula relating the right-hand (or large-K or positive-x) tail slope and the left-hand (or small-K or negative-x) tail slope to how many finite moments the underlying possesses.

Lemma 2 *(Lee [21]) For $k = \log(K/S_0)$. Let $\beta_R := \limsup_{k\to+\infty} \frac{\sigma_{\rm BS}^2(k)}{|k|/T}$ and $\beta_L := \limsup_{k\to-\infty} \frac{\sigma_{\rm BS}^2(k)}{|k|/T}$. Then $\beta_R \in [0,2]$ and $\beta_L \in [0,2]$ and*

$$\tilde{p} = \frac{1}{2\beta_R} + \frac{\beta_R}{8} - \frac{1}{2},$$
$$\tilde{q} = \frac{1}{2\beta_L} + \frac{\beta_L}{8} - \frac{1}{2},$$

where $\frac{1}{0} := \infty$. Equivalently,

$$\beta_R = 2 - 4(\sqrt{\tilde{p}^2 + \tilde{p}} - \tilde{p}),$$
$$\beta_L = 2 - 4(\sqrt{\tilde{q}^2 + \tilde{q}} - \tilde{q}),$$

where the right-hand expression is to be read as zero, in the case $\tilde{p} = \infty$ or $\tilde{q} = \infty$.

Applying Lee's moment formula, we obtain the following results for our model:

Theorem 3 *In the joint regime of fixed-maturity, large-strike (small-strike) with $k = \log(K/S_0)$ $(|k| \to \infty)$, the implied volatility $\sigma_{\rm BS}(k,T)$ approaches the limit*

$$\begin{aligned} \limsup_{k\to+\infty} \frac{\sigma_{\rm BS}^2(k,T)}{|k|/T} &= 2 - 4(\sqrt{\tilde{p}^2 + \tilde{p}} - \tilde{p}), \quad \text{(large strike)}, \\ \limsup_{k\to-\infty} \frac{\sigma_{\rm BS}^2(k,T)}{|k|/T} &= 2 - 4(\sqrt{\tilde{q}^2 + \tilde{q}} - \tilde{q}), \quad \text{(small strike)}, \end{aligned} \tag{2.18}$$

where $\tilde{p}$ and $\tilde{q}$ are defined via

$$\int_0^\infty \frac{\mathrm{d}D}{H(D;\tilde{p}-1)} = T, \quad \int_0^\infty \frac{\mathrm{d}D}{H(D;-\tilde{q})} = T,$$

and

$$H(D;p) := -bD + \frac{1}{2}\sigma^2 D^2 + \beta(\mathbb{E}[e^{aD+pY}] - 1) - p\mu_Y\beta.$$

Proof Please see Section 4.4. □

Remark 2 *Numerical examples are provided in Section 3 to verify the existence of $\tilde{p}$ and $\tilde{q}$ values for different T's in (2.18).*

3. Numerical study

In this section, we present numerical study results. The strength of the self-exciting process is controlled by a in (2.2) and β in the intensity function λ_t^N. Hence we vary a and β values to study how these two parameters affect the rate function and the asymptotic implied volatility. a is chosen to be 0.05, 0.5 and 1 and β is chosen to be 0.1, 0.25 and 0.5. For all numerical studies, we define the jump size $Y \sim \mathcal{N}(0, \sigma^2)$. Other parameters are $b = 1$, $c = 0.05$, $\alpha = 1$, $\sigma^2 = 0.1$ and $\delta^2 = 0.1$.

Figure 2 shows the rate function of the selected a values. As shown in this figure, the growth rate of $I(x)$ increases as a increases. This trend is expected as more rare events occur when a increases; thus, the rate function $I(x)$ tends to be smaller. The right figure is the zoom-in of the left figure, and it shows that the minimums do not coincide. Figure 3 shows the rate function $\bar{I}(x)$, which has similar behaviors as $I(x)$ in Figure 2. Figure 4 shows the asymptotic behavior of implied volatility in the large-maturity and large-strike regimes for different a values. The affine point jump-diffusion model can capture the implied volatility smiles in this regime. Forde and Jacquier [11] found similar implied volatility smiles for the Heston model in the same regime. Consider the At-The-Money (ATM) cases when $x = 0$, and the ATM volatility increases as a increases. This is because as more rare events occur, the implied volatility becomes higher. Additionally, the growth rate of the implied volatility into In-The-Money/Out-The-Money

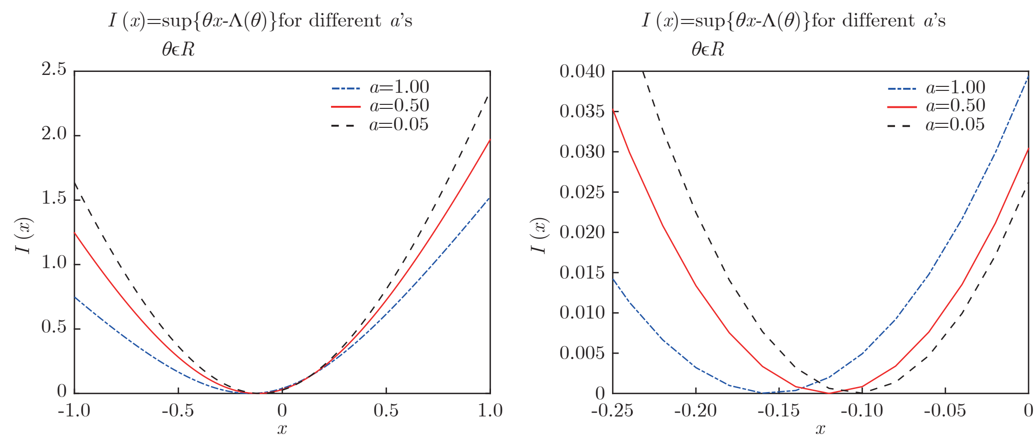

Figure 2 Left: $I(x)$ for $a = 0.05,\ 0.5$ and 1; Right: Zoom-in of left figure near $I(x) = 0$

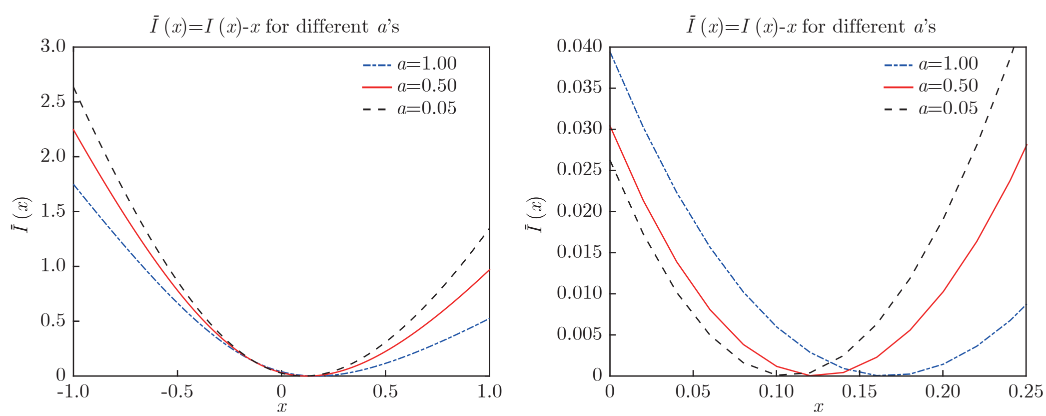

Figure 3 Left: $\bar{I}(x)$ for $a = 0.05,\ 0.50$ and 1; Right: Zoom-in of left figure near $\bar{I}(x) = 0$

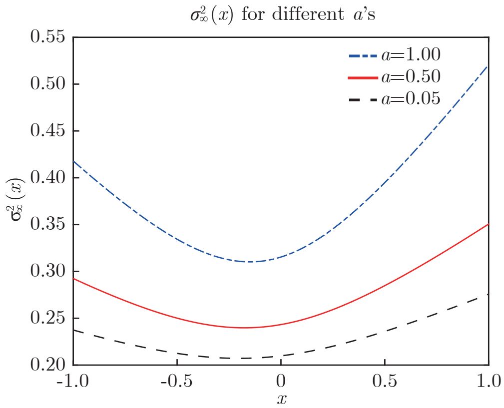

Figure 4 $\sigma^2_\infty(x)$ for $a = 0.05,\ 0.5$ and 1

increases as a increases.

Figures 5, 6, and 7 show the numerical results for different β values. Because the parameter β controls the strength of the intensity of the self-exciting process, varying β has similar effects to varying a.

Numerical examples in fixed-maturity large, small-strike and large-strike are presented. The left figure in Figure 8 shows the ratio of Black-Scholes implied volatility to log-moneyness in the

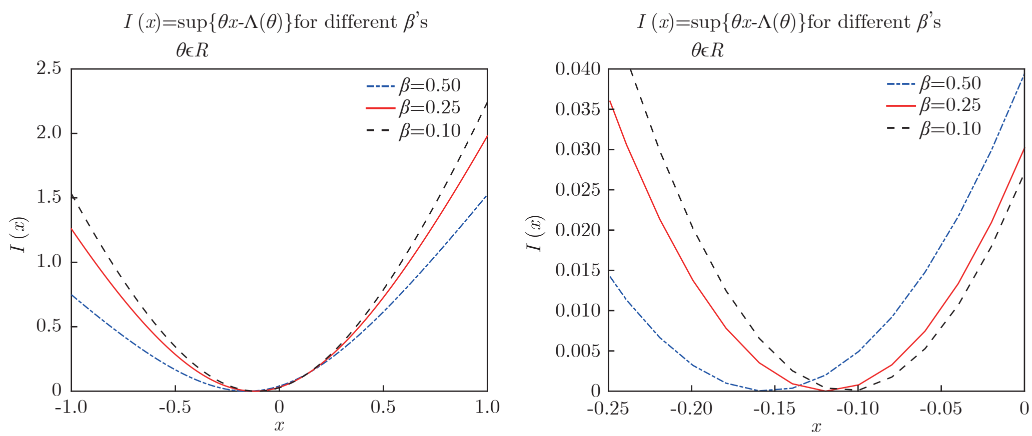

Figure 5 Left: $I(x)$ for $\beta = 0.1,\ 0.25$ and 0.5; Right: Zoom-in of left figure near $I(x) = 0$

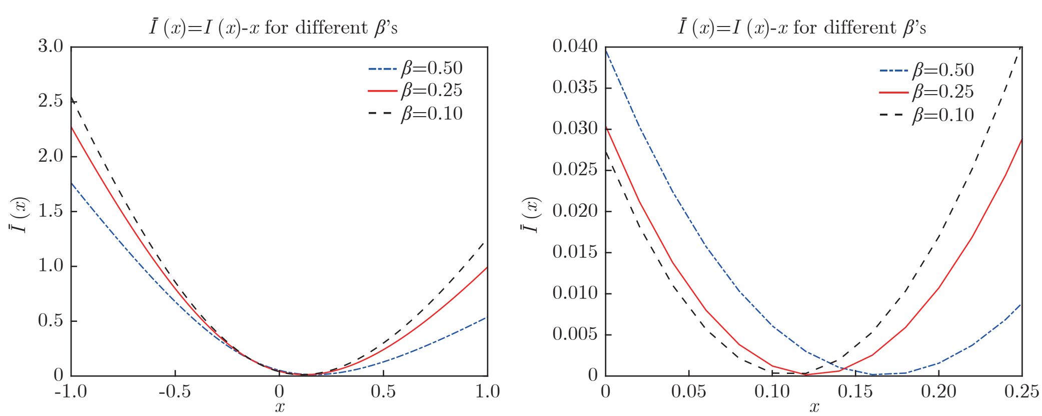

Figure 6 Left: $\bar{I}(x)$ for $\beta = 0.1,\ 0.25$ and 0.5; Right: Zoom-in of left figure near $\bar{I}(x) = 0$

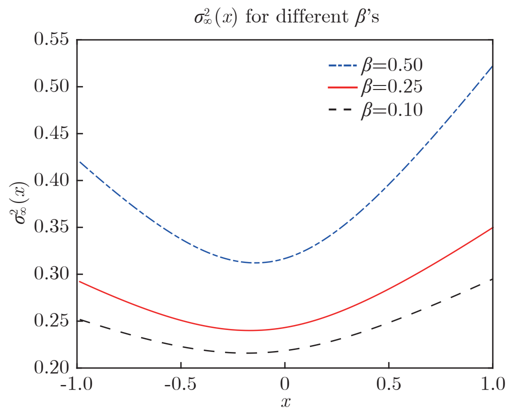

Figure 7 $\sigma^2_\infty(x)$ for $\beta = 0.1,\ 0.25$ and 0.5

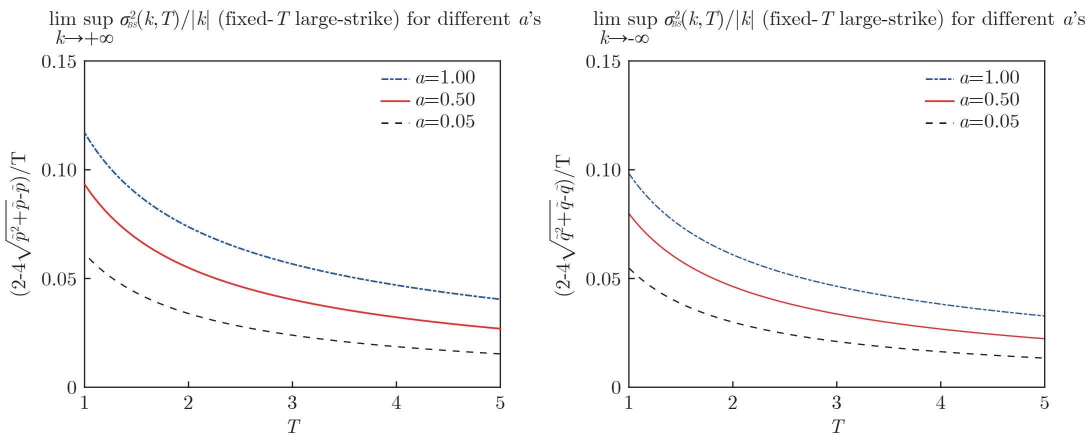

Figure 8 Left: $\limsup\limits_{k\to+\infty} \frac{\sigma^2_{\mathrm{BS}}(k,T)}{|k|}$ (fixed-maturity large-strike) for $a = 0.05,\ 0.5$ and 1;
Right: $\limsup\limits_{k\to-\infty} \frac{\sigma^2_{\mathrm{BS}}(k,T)}{|k|}$ (fixed-maturity small-strike) for $a = 0.05,\ 0.5$ and 1

fixed-maturity and large-strike regime for different a values; while right figure displays the ratio in the fixed-maturity and small-strike regime. The maturity T was selected within a reasonable range. In both figures, we observe that for a given T, the ratio of implied volatility to log-moneyness increases as the self-exciting intensity parameter a increases. It is interesting to point out that in these regimes, the ratio decreases as maturity increases. This is practically observed on an implied volatility surface. Results for different β's are provided in Figure 9. We obtain similar results because β controls the strength of the self-exciting process as well.

4. Proofs of the main results

4.1 Proof of Proposition 1

Proof Given any θ in $\mathbb{R}$, the moment generating function for X_t is

$$\begin{aligned}\mathbb{E}[e^{\theta X_t}] &= \mathbb{E}\left[e^{\theta\left(-\frac{1}{2}\sigma^2 t+\sigma W_t^{\mathbb{Q}}-\mu_Y\int_0^t \lambda_s^N ds+\sum_{i=1}^{N_t} Y_i\right)}\right] \\ &= e^{(-\frac{1}{2}\theta\sigma^2+\frac{1}{2}\theta^2\sigma^2-\theta\mu_Y\alpha)t}\mathbb{E}[e^{-\theta\mu_Y\beta\int_0^t\lambda_s ds+\theta L_t}].\end{aligned} \tag{4.1}$$

For any $\Theta = (\theta_1, \theta_2, \theta_3) \in \mathbb{R}^3$, we assume

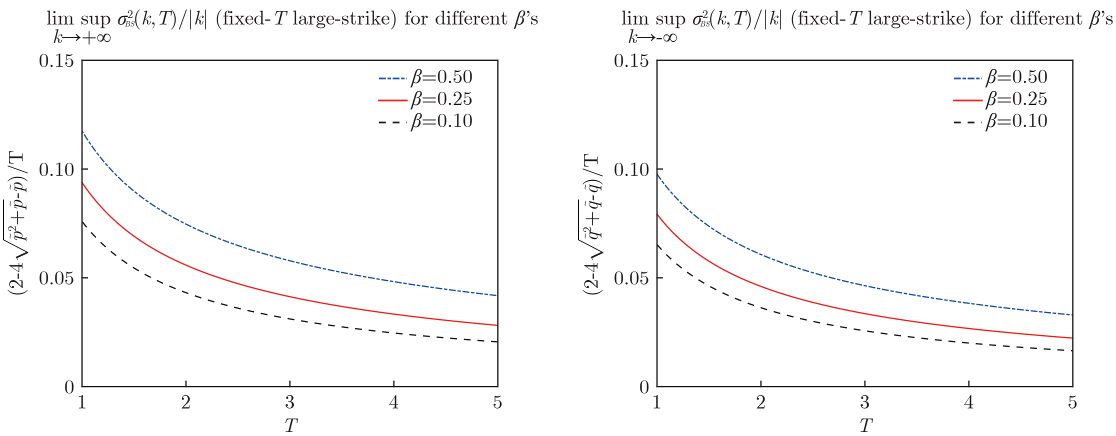

Figure 9 Left: $\limsup_{k\to+\infty}\frac{\sigma^2_{\text{BS}}(k,T)}{|k|}$ (fixed-maturity large-strike) for $\beta=0.1,\ 0.25$ and 0.5; Right: $\limsup_{k\to-\infty}\frac{\sigma^2_{\text{BS}}(k,T)}{|k|}$ (fixed-maturity small-strike) for $\beta=0.1,\ 0.25$ and 0.5

$$\mathbb{E}[e^{\theta_1\int_t^T\lambda_s\mathrm{d}s+\theta_2\lambda_T+\theta_3L_T}|\lambda_t=\lambda,L_t=L]=u(t,\lambda,L):=u(t,\lambda,L,\Theta). \tag{4.2}$$

By applying Feynman-Kac formula, we have

$$\begin{cases}\dfrac{\partial u}{\partial t}+b(c-\lambda)\dfrac{\partial u}{\partial\lambda}+\dfrac{1}{2}\sigma^2\lambda\dfrac{\partial^2u}{\partial\lambda^2}+(\alpha+\beta\lambda)\displaystyle\int_{\mathbb{R}}(u(t,\lambda+a,L+y)-u(t,\lambda,L))\mathbb{Q}(\mathrm{d}y)+\theta_1\lambda u=0,\\ u(T,\lambda,L,\Theta)=e^{\theta_2\lambda+\theta_3L}.\end{cases} \tag{4.3}$$

Let us try a solution in the form of $u(t,\lambda,L)=e^{A(t;\Theta)\lambda+B(t;\Theta)L+C(t;\Theta)}$, by substituting it into (4.3), we have

$$\begin{cases}\Big[A'(t;\Theta)-bA(t;\Theta)+\dfrac{1}{2}\sigma^2A^2(t;\Theta)+\beta\displaystyle\int_{\mathbb{R}}(e^{A(t;\Theta)a+B(t;\Theta)y}-1)\mathbb{Q}(\mathrm{d}y)+\theta_1\Big]\lambda\\ \quad+B'(t;\Theta)L+[C'(t;\Theta)+bcA(t;\Theta)+\alpha\displaystyle\int_{\mathbb{R}}(e^{A(t;\Theta)a+B(t;\Theta)y}-1)\mathbb{Q}(\mathrm{d}y)]=0,\\ A(T;\Theta)=\theta_2,\quad B(T;\Theta)=\theta_3,\quad C(T;\Theta)=0.\end{cases} \tag{4.4}$$

Furthermore we have $A(t;\Theta),B(t;\Theta),C(t;\Theta)$ satisfy the following ODEs,

$$\begin{cases}A'(t;\Theta)-bA(t;\Theta)+\dfrac{1}{2}\sigma^2A^2(t;\Theta)+\beta\displaystyle\int_{\mathbb{R}}(e^{A(t;\Theta)a+B(t;\Theta)y}-1)\mathbb{Q}(\mathrm{d}y)+\theta_1=0,\\ B'(t;\Theta)=0,\\ C'+bcA(t;\Theta)+\alpha\displaystyle\int_{\mathbb{R}}(e^{A(t;\Theta)a+B(t;\Theta)y}-1)\mathbb{Q}(\mathrm{d}y)=0,\\ A(T;\Theta)=\theta_2,\quad B(T;\Theta)=\theta_3,\quad C(T;\Theta)=0.\end{cases} \tag{4.5}$$

With $B'(t;\Theta)=0,B(T;\Theta)=\theta_3$, we have $B(t;\Theta)=\theta_3$ and thus $u(t,\lambda,L)=e^{A(t;\Theta)\lambda+\theta_3L+C(t;\Theta)}$, then $A(t;\Theta),C(t;\Theta)$ satisfy the following ODEs:

$$\begin{cases}A'(t;\Theta)-bA(t;\Theta)+\dfrac{1}{2}\sigma^2A^2(t;\Theta)+\beta\displaystyle\int_{\mathbb{R}}(e^{A(t;\Theta)a+\theta_3y}-1)\mathbb{Q}(\mathrm{d}y)+\theta_1=0,\\ C'+bcA(t;\Theta)+\alpha\displaystyle\int_{\mathbb{R}}(e^{A(t;\Theta)a+\theta_3y}-1)\mathbb{Q}(\mathrm{d}y)=0,\\ A(T;\Theta)=\theta_2,\quad C(T;\Theta)=0.\end{cases} \tag{4.6}$$

Define $f(t,\lambda,L):=f(t,\lambda,L,\Theta):=\mathbb{E}[e^{\theta_1\int_0^t\lambda_s\mathrm{d}s+\theta_2\lambda_t+\theta_3 L_t}|\lambda_0=\lambda,L_0=L]$. Let $u(t,\lambda,L)=f(T-t,\lambda,L)$ and make the time change $t\mapsto T-t$ to change the backward equation to the forward equation, we obtain

$$\begin{cases} -\dfrac{\partial f}{\partial s}+b(c-\lambda)\dfrac{\partial f}{\partial \lambda} \\ \quad +\dfrac{1}{2}\sigma^2\lambda\dfrac{\partial^2 f}{\partial\lambda^2}+(\alpha+\beta\lambda)\displaystyle\int_{\mathbb{R}}(f(s,\lambda+a,L+y)-f(s,\lambda,L))\mathbb{Q}(\mathrm{d}y)+\theta_1\lambda f=0, \\ f(0,\lambda,L,\Theta)=e^{\theta_2\lambda+\theta_3 L}. \end{cases} \tag{4.7}$$

We try $f(t,\lambda,L)=e^{D(t;\Theta)\lambda+E(t;\Theta)L+F(t;\Theta)}$, then we have $D(t;\Theta),E(t;\Theta),F(t;\Theta)$ satisfy the following ordinary differential equations

$$\begin{cases} D'(t;\Theta)+bD(t;\Theta)-\dfrac{1}{2}\sigma^2D^2(t;\Theta)-\beta\displaystyle\int_{\mathbb{R}}(e^{D(t;\Theta)a+E(t;\Theta)y}-1)\mathbb{Q}(\mathrm{d}y)-\theta_1=0, \\ E'(t;\Theta)=0, \\ F'-bcD(t;\Theta)-\alpha\displaystyle\int_{\mathbb{R}}(e^{D(t;\Theta)a+E(t;\Theta)y}-1)\mathbb{Q}(\mathrm{d}y)=0, \\ D(0;\Theta)=\theta_2,\quad E(0;\Theta)=\theta_3,\quad F(0;\Theta)=0. \end{cases} \tag{4.8}$$

Finally, by substituting $\Theta=(\theta_1,\theta_2,\theta_3)=(-\theta\mu_Y\beta,0,\theta)$ into (4.8), we have $f(s,\lambda,L)=e^{D(s;\Theta)\lambda+\theta L+F(s;\Theta)}$ and $D(s;\Theta)$, $F(s;\Theta)$ satisfy the following ODEs,

$$\begin{cases} D'(s;\Theta)+bD(s;\Theta)-\dfrac{1}{2}\sigma^2D^2(s;\Theta)-\beta\displaystyle\int_{\mathbb{R}}(e^{D(s;\Theta)a+\theta y}-1)\mathbb{Q}(\mathrm{d}y)+\theta\mu_Y\beta=0, \\ F'(s;\Theta)-bcD(s;\Theta)-\alpha\displaystyle\int_{\mathbb{R}}(e^{D(s;\Theta)a+\theta y}-1)\mathbb{Q}(\mathrm{d}y)=0, \\ D(0;\Theta)=0,\quad F(0;\Theta)=0. \end{cases} \tag{4.9}$$

This completes the proof. □

4.2 Proof of Theorem 1

Proof From (4.1) and (4.9), for any $\theta\in\mathbb{R}$, we have

$$\begin{aligned} \mathbb{E}[e^{\theta X_t}] &= e^{(-\frac{1}{2}\theta\sigma^2+\frac{1}{2}\theta^2\sigma^2-\theta\mu_Y\alpha)t}\mathbb{E}[e^{-\theta\mu_Y\beta\int_0^t\lambda_s ds+\theta L_t}] \\ &= e^{(-\frac{1}{2}\theta\sigma^2+\frac{1}{2}\theta^2\sigma^2-\theta\mu_Y\alpha)t+D(t,\theta)\lambda+\theta L+F(t,\theta)}, \end{aligned} \tag{4.10}$$

where $D(t;\theta)$ and $F(t;\theta)$ satisfy the following ODEs,

$$\begin{cases} D'(t;\theta)+bD(t;\theta)-\dfrac{1}{2}\sigma^2D^2(t;\theta)-\beta\displaystyle\int_{\mathbb{R}}(e^{\bar{D}(t;\theta)a+\theta y}-1)\mathbb{Q}(\mathrm{d}y)+\theta\mu_Y\beta=0, \\ F'(t;\theta)-bcD(t;\theta)-\alpha\displaystyle\int_{\mathbb{R}}(e^{D(t;\theta)a+\theta y}-1)\mathbb{Q}(\mathrm{d}y)=0, \\ D(0;\theta)=0,\quad F(0;\theta)=0. \end{cases} \tag{4.11}$$

Thus, from (4.10) we have

$$\begin{aligned} \Lambda(\theta) &: = \lim_{t\to\infty}\frac{1}{t}\log\mathbb{E}[e^{\theta X_t}] \\ &= \frac{1}{2}\sigma^2\theta^2-\left(\frac{1}{2}\sigma^2+\mu_Y\alpha\right)\theta+\lambda\lim_{t\to\infty}\frac{D(t;\theta)}{t}+\lim_{t\to\infty}\frac{F(t;\theta)}{t}. \end{aligned}$$

We want the conditions that $\Lambda(\theta)$ exists and is finite. Consider the following two scenarios of $\lim\limits_{t\to\infty}D(t;\theta)$:

(1) When $\lim_{t\to\infty} D(t;\theta)$ is infinite, we know $F'(t;\theta)$ is also infinite from (4.11), and hence $\lim_{t\to\infty} \frac{F(t;\theta)}{t} = \lim_{t\to\infty} F'(t;\theta) = +\infty$, which does not satisfies the conditions.

(2) When $\lim_{t\to\infty} D(t;\theta)$ is finite, we know that $\lim_{t\to\infty} F'(t;\theta)$ is also finite from (4.11), and hence $\lim_{t\to\infty} \frac{D(t;\theta)}{t}=0$, $\lim_{t\to\infty} \frac{F(t;\theta)}{t} = \lim_{t\to\infty} F'(t;\theta) = \lim_{t\to\infty} \left[bcD(\theta)+\alpha\left(e^{aD(\theta)}\mathbb{E}[e^{\theta Y}]-1\right)\right] < \infty$, which exactly satisfies the conditions.

Since $\lim_{t\to\infty} D(t;\theta)$ is finite, $\lim_{t\to\infty} D'(t;\theta)$ must be 0. From (4.11), one can see that

$$\begin{aligned}\lim_{t\to\infty} D'(t;\theta) = \Gamma(D,\theta) :&= -bD + \frac{1}{2}\sigma^2 D^2 + \beta\int_{\mathbb{R}}(e^{aD+\theta y}-1)\mathbb{Q}(\mathrm{d}y) - \theta\mu_Y\beta \\ &= -bD + \frac{1}{2}\sigma^2 D^2 + \beta(\mathbb{E}[e^{aD+\theta Y}]-1) - \theta\mu_Y\beta.\end{aligned}$$

Next we find the range of θ such that $\lim_{t\to\infty} D'(t;\theta) = 0$, that is,

$$\Gamma(D,\theta) = -bD + \frac{1}{2}\sigma^2 D^2 + \beta(\mathbb{E}[e^{aD+\theta Y}]-1) - \theta\mu_Y\beta = 0 \tag{4.12}$$

has a solution of $D(\theta)$. We know that

$$\begin{aligned}\Gamma'_D(D,\theta) &= -b + \sigma^2 D + a\beta e^{aD}\mathbb{E}[e^{\theta Y}], \\ \Gamma''_D(D,\theta) &= \sigma^2 + a^2\beta e^{aD}\mathbb{E}[e^{\theta Y}],\end{aligned}$$

and we find that $\Gamma''_D(D,\theta) > 0$, so $\Gamma(D,\theta)$ is convex and $\Gamma'_D(D,\theta)$ is increasing in D. Clearly we have $\lim_{D\to-\infty} \Gamma'_D(D,\theta) = -\infty$ and $\lim_{D\to+\infty} \Gamma'_D(D,\theta) = +\infty$, so there exists a unique $D_c(\theta)$ that satisfies the following equation and that $\Gamma(D_c,\theta)$ is the minimum of $\Gamma(D,\theta)$.

$$\Gamma'_D(D_c,\theta) = -b + \sigma^2 D_c + a\beta e^{aD_c}\mathbb{E}[e^{\theta Y}] = 0. \tag{4.13}$$

We take the derivative of $D_c(\theta)$ on θ,

$$D'_c(\theta) = -\frac{a\beta e^{aD_c(\theta)}\mathbb{E}[Ye^{\theta Y}]}{\sigma^2 + a^2\beta e^{aD_c(\theta)}\mathbb{E}[e^{\theta Y}]}. \tag{4.14}$$

And we can rewrite this minimum $\Gamma(D_c(\theta),\theta)$ as

$$\Gamma(D_c(\theta),\theta) = G(\theta) := -bD_c(\theta) + \frac{\sigma^2}{2}D_c^2(\theta) + \beta(e^{aD_c(\theta)}(\mathbb{E}[e^{\theta Y}]-1) - \theta\beta\mu_Y. \tag{4.15}$$

Since $\lim_{D\to+\infty} \Gamma'_D(D,\theta) = +\infty$, if $G(\theta) \leqslant 0$, then (4.12) must have a solution of $D(\theta)$. As a result, we now arrive at finding the scope of θ such that $G(\theta) \leqslant 0$. Take the derivative of $G(\theta)$ on θ,

$$G'(\theta) = \beta\left(e^{aD_c(\theta)}\mathbb{E}[Ye^{\theta Y}] - \mu_Y\right), \tag{4.16}$$

$$G''(\theta) = \frac{\sigma^2\beta e^{aD_c(\theta)}\mathbb{E}[Y^2e^{\theta Y}] + a^2\beta^2e^{2aD_c(\theta)}(\mathbb{E}[Y^2e^{\theta Y}]\mathbb{E}[e^{\theta Y}] - \mathbb{E}[Ye^{\theta Y}]^2)}{\sigma^2 + a^2\beta e^{aD_c(\theta)}\mathbb{E}[e^{\theta Y}]}. \tag{4.17}$$

By Cauchy–Schwarz inequality and (4.14), we can get $G''(\theta) > 0$, so $G(\theta)$ is convex and $G'(\theta)$ is increasing. Further, with the fact that $\lim_{\theta\to-\infty} D_c(\theta) = \frac{b}{\sigma^2}$ from (4.13), we can easily see that $\lim_{\theta\to-\infty} G'(\theta) < 0$.

When $\theta = 0$, the equation becomes

$$\Gamma(D,0) = -bD + \frac{1}{2}\sigma^2 D^2 + \beta e^{aD} - \beta = 0. \tag{4.18}$$

It is straightforward to see that $D = 0$ is the solution, which means $G(0) \leqslant 0$. Furthermore, it can be proved that $G(0) < 0$ by contradiction. If $G(0) = 0$, then $D_c(0) = 0$, and from (4.13) we must have

$$-b + \sigma^2 D_c(\theta) + a\beta e^{aD_c(\theta)}\mathbb{E}[e^{\theta Y}]\,\Big|_{\theta=0} = 0, \tag{4.19}$$

however,

$$-b + \sigma^2 D_c(0) + a\beta e^{aD_c(0)}\mathbb{E}[e^{0Y}] = -b + a\beta < 0, \tag{4.20}$$

which is inconsistent with (4.19).

Then, the discussion of $G(\theta) \leqslant 0$ can be summarized in the following two cases.

Case one: $\lim_{\theta\to+\infty} G'(\theta) > 0$, in this case, $G(\theta) = 0$ has two solutions $\theta_{\min}, \theta_{\max}$, thus when $\theta \in [\theta_{\min}, \theta_{\max}]$, $G(\theta) \leqslant 0$(as in Figure 10).

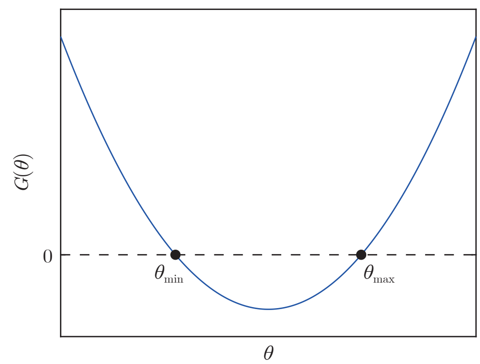

Figure 10 Case one

Case two: $\lim_{\theta\to+\infty} G'(\theta) \leqslant 0$, in this case, $G(\theta) = 0$ has the unique solution $\theta_{\min}$, thus when $\theta \geqslant \theta_{\min}$, $G(\theta) \leqslant 0$(as in Figure 11).

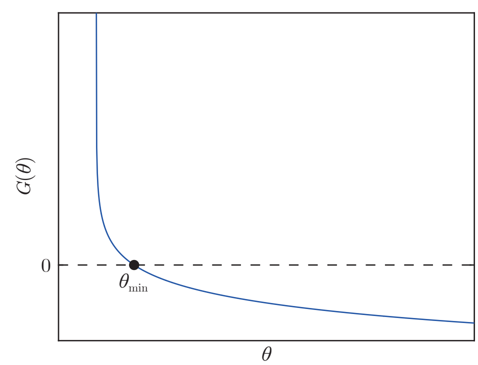

Figure 11 Case two

Therefore for $\theta \in [\theta_{min}, \theta_{max}]$ (in **Case two**, $\theta_{max} \longrightarrow +\infty$), we have

$$\Lambda(\theta) = \lim_{t\to\infty} \frac{1}{t}\log\mathbb{E}[e^{\theta X_t}] = \frac{1}{2}\sigma^2\theta^2 - \left(\frac{1}{2}\sigma^2 + \mu_Y\alpha\right)\theta + bcD(\theta) + \alpha\left(e^{aD(\theta)}\mathbb{E}[e^{\theta Y}] - 1\right).$$

When $\theta \notin [\theta_{min}, \theta_{max}]$, this limit is ∞. We are to check two conditions for G ä rtner–Ellis theorem. The first condition is essential smoothness. By differentiating the equation (4.12) with respect to θ, that is when $\theta \to \theta_{\min(\max)}$, then $D \to D_c$, and

$$\frac{\partial D}{\partial \theta}=\frac{\beta(\mu_Y-e^{aD}\mathbb{E}[Ye^{\theta Y}])}{-b+\sigma^2 D+a\beta e^{aD}\mathbb{E}[e^{\theta Y}]}\to+\infty.$$

The second is $0\in(\theta_{min},\theta_{\max})$, which is obviously satisfied since $G(0)<0$.

Upon applying Gärtner–Ellis theorem (refer to Dembo and Zeitouni [6] for the definition of essential smoothness and statement of G ä rtner–Ellis theorem), $\mathbb{Q}(\frac{1}{t}X_t\in\cdot)$ satisfies a large deviation principle with rate function

$$I(x)=\sup_{\theta\in\mathbb{R}}\left\{\theta x-\left(\frac{1}{2}\sigma^2\theta^2-\left(\frac{1}{2}\sigma^2+\mu_Y\alpha\right)\theta+bcD(\theta)+\alpha\left(e^{aD(\theta)}\mathbb{E}[e^{\theta Y}]-1\right)\right)\right\}.$$

□

4.3 Proof of Theorem 2

Proof First, let us give a more explicit expression for $I(x)$ in (2.7). Note that

$$I(x)=\theta^*x-\Lambda(\theta^*),$$

let $\frac{\mathrm{d}}{\mathrm{d}\theta}I(x)=0$, where $x=\Lambda'(\theta^*)$ so that

$$\sigma^2\theta^*-\left(\frac{1}{2}\sigma^2+\mu_Y\alpha\right)+bcD'(\theta^*)+\alpha D'(\theta^*)e^{aD}\mathbb{E}[e^{\theta^*Y}]+\alpha\mathbb{E}[Ye^{aD+\theta^*Y}]=x,$$

which gives that

$$D'(\theta^*)=\frac{x+\frac{1}{2}\sigma^2+\mu_Y\alpha-\theta^*\sigma^2-\alpha\mathbb{E}[Ye^{aD+\theta^*Y}]}{bc+\alpha e^{aD}\mathbb{E}[e^{\theta^*Y}]}.$$

On the other hand, take the derivative of equation $\Gamma(D(\theta),\theta)=0$ on θ,

$$-bD'(\theta)+\sigma^2D(\theta)D'(\theta)+\beta\mathbb{E}\left[(aD'(\theta)+Y)e^{aD(\theta)+\theta Y}\right]-\mu_Y\beta=0,$$

that is

$$D'(\theta)\left(\sigma^2D(\theta)-b+a\beta\mathbb{E}[e^{aD(\theta)+\theta Y}]\right)=\mu_Y\beta-\beta\mathbb{E}[Ye^{aD(\theta)+\theta Y}].$$

Therefore we can solve for θ^* and $D(\theta^*)$ from the following equations:

$$\begin{cases}\dfrac{x+\frac{1}{2}\sigma^2+\mu_Y\alpha-\theta^*\sigma^2-\alpha\mathbb{E}[Ye^{aD+\theta^*Y}]}{bc+\alpha e^{aD}\mathbb{E}[e^{\theta^*Y}]}\left(\sigma^2D(\theta^*)-b+a\beta\mathbb{E}[e^{aD(\theta^*)+\theta^*Y}]\right)\\ \quad=\beta\left(\mu_Y-\mathbb{E}[Ye^{aD(\theta^*)+\theta^*Y}]\right),\\ -bD(\theta^*)+\dfrac{1}{2}\sigma^2D(\theta^*)^2+\beta\left(\mathbb{E}[e^{aD(\theta^*)+\theta^*Y}]-1\right)-\theta^*\mu_Y\beta=0.\end{cases}\tag{4.21}$$

Second, let us define the share measure $\bar{\mathbb{Q}}$ as

$$\left.\frac{\mathrm{d}\bar{\mathbb{Q}}}{\mathrm{d}\mathbb{Q}}\right|_{\mathcal{F}_t}=\frac{S_t}{S_0}=e^{X_t}.\tag{4.22}$$

Note that

$$\begin{aligned}\frac{S_t}{S_0}&=e^{-\frac{1}{2}\sigma^2t+\sigma W_t^{\mathbb{Q}}-\mu_Y\int_0^t\lambda_s^N\mathrm{d}s+\sum_{i=1}^{N_t}Y_i}\\&=e^{-\frac{1}{2}\sigma^2t+\sigma W_t^{\mathbb{Q}}}\cdot\prod_{i=1}^{N_t}\frac{e^{Y_i}}{\mathbb{E}[e^Y]}\cdot e^{\log\mathbb{E}[e^Y]N_t-\mu_Y\int_0^t\lambda_s^N\mathrm{d}s}.\end{aligned}$$

Thus, under the share measure $\bar{\mathbb{Q}}$,

$$\bar{X}_t = \frac{1}{2}\sigma^2 t + \sigma W_t^{\bar{\mathbb{Q}}} - \mu_Y \int_0^t \bar{\lambda}_s^{\bar{N}} \mathrm{d}s + \sum_{i=1}^{\bar{N}_t} \bar{Y}_i, \tag{4.23}$$

where $\bar{Y}_i$ are i.i.d. and according to $\bar{\mathbb{Q}}$ so that it has the probability distribution

$$\frac{e^Y}{\mathbb{E}[e^Y]} \mathrm{d}\mathbb{Q}$$

and $\bar{N}_t$ is an affine point process with intensity

$$\bar{\lambda}_t^{\bar{N}} = \mathbb{E}[e^Y]\lambda_t^N.$$

Thus, $\bar{\mathbb{Q}}(\frac{1}{t}\bar{X}_t \in \cdot)$ satisfies a large deviation principle with

$$\bar{I}(x) := \sup_{\theta\in\mathbb{R}}\{\theta x - \bar{\Lambda}(\theta)\},$$

where

$$\bar{\Lambda}(\theta) := \lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}[e^{\theta \bar{X}_t}] = \frac{1}{2}\sigma^2\theta^2 + \left(\frac{1}{2}\sigma^2 - \mu_Y \mathbb{E}[e^Y]\alpha\right)\theta + bc\bar{D}(\theta) + \mathbb{E}[e^Y]\alpha\left(e^{a\bar{D}(\theta)}\mathbb{E}[e^{\theta\bar{Y}}] - 1\right),$$

where $\bar{D}(\theta)$ is the smaller solution of the equation

$$-b\bar{D}(\theta) + \frac{1}{2}\sigma^2\bar{D}(\theta)^2 + \mathbb{E}[e^Y]\beta\left(\mathbb{E}[e^{a\bar{D}(\theta)+\theta\bar{Y}}] - 1\right) - \theta\mu_Y\mathbb{E}[e^Y]\beta = 0. \tag{4.24}$$

As a corollary, $\bar{\mathbb{Q}}(-\frac{1}{t}\bar{X}_t \in \cdot)$ satisfies a large deviation principle with the rate function $\bar{I}(-x)$. Moreover, for any $x \in \mathbb{R}$ and for any sufficiently small $\delta > 0$,

$$\bar{\mathbb{Q}}\left(x - \delta < \frac{\bar{X}_t}{t} < x + \delta\right) = \mathbb{E}\left[e^{X_t} 1_{x-\delta<\frac{X_t}{t}<x+\delta}\right],$$

which implies that

$$\bar{I}(x) = I(x) - x.$$

Third, according to Lemma 1, we have

$$I(x) - x = \begin{cases} -\lim_{T\to\infty} \frac{1}{T}\log\mathbb{E}[(S_T - S_0e^{xT})^+], & \text{for } x \geqslant x_R, \\ -\lim_{T\to\infty} \frac{1}{T}\log(S_0 - \mathbb{E}[(S_T - S_0e^{xT})^+]), & \text{for } x_L \leqslant x \leqslant x_R, \\ -\lim_{T\to\infty} \frac{1}{T}\log\mathbb{E}[(S_0e^{xT} - S_T)^+], & \text{for } x \leqslant x_L, \end{cases} \tag{4.25}$$

from which we can compute that

$$x_L = \Lambda'(0), \quad x_R = \bar{\Lambda}'(0). \tag{4.26}$$

Differentiating $\Lambda(\theta)$ with respect to θ, we get

$$\Lambda'(\theta) = \sigma^2\theta - \left(\frac{1}{2}\sigma^2 + \mu_Y\alpha\right) + bcD'(\theta) + \alpha e^{aD(\theta)}\left(aD'(\theta)\mathbb{E}[e^{\theta Y}] + \mathbb{E}[Ye^{\theta Y}]\right). \tag{4.27}$$

From equation (4.21), we have

$$D'(\theta) = \frac{\beta\left(\mu_Y - \mathbb{E}[Ye^{aD(\theta)+\theta Y}]\right)}{\sigma^2 D(\theta) - b + a\beta\mathbb{E}[e^{aD(\theta)+\theta Y}]},$$

and $D(0) = 0$ from (4.18), so

$$D'(0) = \frac{\beta\left(\mu_Y - \mathbb{E}[Y]\right)}{a\beta - b}. \tag{4.28}$$

Plugging equation (4.28) into equation (4.27), we have

$$x_L = \Lambda'(0) = -\left(\frac{1}{2}\sigma^2 + \mu_Y\alpha\right) + (bc + a\alpha)\,\frac{\beta\,(\mu_Y - \mathbb{E}[Y])}{a\beta - b} + \alpha\mathbb{E}[Y].$$

Similarly, differentiating $\bar{\Lambda}(\theta)$ w.r.t. θ,

$$\bar{\Lambda}'(\theta) = \sigma^2\theta + \left(\frac{1}{2}\sigma^2 - \mu_Y\mathbb{E}[e^Y]\alpha\right) + bc\bar{D}'(\theta) + \mathbb{E}[e^Y]\alpha e^{a\bar{D}(\theta)}\left(a\bar{D}'(\theta)\mathbb{E}[e^{\theta\bar{Y}}] + \mathbb{E}[\bar{Y}e^{\theta\bar{Y}}]\right). \quad (4.29)$$

In addition, from equation (4.24) we have

$$\bar{D}'(\theta) = \frac{\beta\mathbb{E}[e^Y]\left(\mu_Y - \mathbb{E}[\bar{Y}e^{a\bar{D}(\theta)+\theta\bar{Y}}]\right)}{\sigma^2\bar{D}(\theta) - b + a\beta\mathbb{E}[e^Y]\mathbb{E}[e^{a\bar{D}(\theta)+\theta\bar{Y}}]},$$

and $\bar{D}(0) = 0$ from (4.24), so

$$\bar{D}'(0) = \frac{\beta\mathbb{E}[e^Y]\left(\mu_Y - \mathbb{E}[\bar{Y}]\right)}{a\beta\mathbb{E}[e^Y] - b}. \quad (4.30)$$

Plugging equation (4.30) into equation (4.29), we have

$$x_R = \bar{\Lambda}'(0) = \left(\frac{1}{2}\sigma^2 - \mu_Y\mathbb{E}[e^Y]\alpha\right) + \left(bc + a\mathbb{E}[e^Y]\alpha\right)\frac{\mathbb{E}[e^Y]\beta\left(\mu_Y - \mathbb{E}[\bar{Y}]\right)}{a\mathbb{E}[e^Y]\beta - b} + \mathbb{E}[e^Y]\alpha\mathbb{E}[\bar{Y}].$$

In summary,

$$x_L = \Lambda'(0) = -\left(\frac{1}{2}\sigma^2 + \mu_Y\alpha\right) + (bc + a\alpha)\,\frac{\beta\,(\mu_Y - \mathbb{E}[Y])}{a\beta - b} + \alpha\mathbb{E}[Y]$$

and

$$x_R = \bar{\Lambda}'(0) = \left(\frac{1}{2}\sigma^2 - \mu_Y\mathbb{E}[e^Y]\alpha\right) + \left(bc + a\mathbb{E}[e^Y]\alpha\right)\frac{\mathbb{E}[e^Y]\beta\left(\mu_Y - \mathbb{E}[\bar{Y}]\right)}{a\mathbb{E}[e^Y]\beta - b} + \mathbb{E}[e^Y]\alpha\mathbb{E}[\bar{Y}].$$

Then in the joint regime of large-maturity, large-strike with $k = \log(K/S_0)$ $(T \to \infty,\ |k| \to \infty)$, according to Lemma 1, the implied volatility $\sigma_{\mathrm{BS}}(k, T)$ approaches the limit

$$\lim_{T\to\infty} \sigma^2_{\mathrm{BS}}(xT, T) = \sigma^2_\infty(x),$$

where

$$\sigma^2_\infty(x) = \begin{cases} 2(2I(x) - x - 2\sqrt{I^2(x) - xI(x)}), & x \in (-\infty, x_L) \cup (x_R, \infty), \\ 2(2I(x) - x + 2\sqrt{I^2(x) - xI(x)}), & x \in [x_L, x_R]. \end{cases}$$

□

4.4 Proof of Theorem 3

Proof Let us determine the $\tilde{p}$ and $\tilde{q}$ in (2.16) and (2.17) for S_T in (2.1). Recall that $\tilde{p} + 1$ is the largest p such that $\mathbb{E}[e^{pX_T}] < \infty$. From (4.10), we know

$$\mathbb{E}[e^{pX_T}] = e^{(-\frac{1}{2}p\sigma^2 + \frac{1}{2}p^2\sigma^2 - p\mu_Y\alpha)T + D(T;p)\lambda + pL + F(T;p)},$$

where $D(T;p)$ and $F(T;p)$ solve a set of ODEs. According to the ODEs (4.11), we see $F(T;p)$ is determined by $D(T;p)$, so $\mathbb{E}[e^{pX_T}] < \infty \Longleftrightarrow D(T;p) < \infty$ and the critical $\tilde{p}$ is the value of p such that $D(T;p) = \infty$. Recall that $D(t;p)$ solves the ODE in (4.11)

$$\begin{cases} D'(t;p) = -bD(t;p) + \dfrac{1}{2}\sigma^2 D^2(t;p) + \beta(\mathbb{E}[e^{aD+pY}] - 1) - p\mu_Y\beta, \\ D(0;p) = 0. \end{cases}$$

Define $H(D;p) := D'(t;p)$,

$$\int_{D(0;p)}^{D(T;p)} \frac{\mathrm{d}D}{H(D;p)} = \int_0^T \mathrm{d}t = T.$$

Therefore the critical $p = \tilde{p} - 1$ satisfies $\int_0^\infty \mathrm{d}D/H(D,p) = T$ as $D(T;p) = \infty$. For a given maturity T, we can find p which satisfies

$$\int_0^\infty \frac{\mathrm{d}D}{-bD + \frac{1}{2}\sigma^2 D^2 + \beta(\mathbb{E}[e^{aD+pY}] - 1) - p\mu_Y\beta} = T.$$

Similarly, the critical $\tilde{q} = -q$ satisfies $\int_0^\infty \mathrm{d}D/H(D,q) = T$. Therefore, according to Lemma 2, in the joint regime of fixed-maturity, large-strike (small-strike) with $k = \log(K/S_0)$ $(|k| \to \infty)$, the implied volatility $\sigma_{\mathrm{BS}}(k,T)$ approaches the limit

$$\limsup_{k\to+\infty} \frac{\sigma_{\mathrm{BS}}^2(k,T)}{|k|/T} = 2 - 4(\sqrt{\tilde{p}^2 + \tilde{p}} - \tilde{p}), \quad \text{(large strike)},$$

$$\limsup_{k\to-\infty} \frac{\sigma_{\mathrm{BS}}^2(k,T)}{|k|/T} = 2 - 4(\sqrt{\tilde{q}^2 + \tilde{q}} - \tilde{q}), \quad \text{(small strike)}.$$

□

5. Concluding remarks

In this paper, we investigate the asymptotic behavior of implied volatility in an affine jump-diffusion model. We assume that $X_t = \log(S_t/S_0)$ and S_t follow an affine jump-diffusion model under a risk-neutral measure. Applying the Feynman–Kac formula, we compute the moment-generating function for X_t. An explicit form of the moment-generating function was obtained by solving a set of ODEs. A large-maturity large deviation principle for X_t was obtained using the Gärtner–Ellis Theorem. We characterized the asymptotic behavior of implied volatility for X_t in the joint large-maturity and large-strike regimes. We used Lee's moment formula to derive the asymptotic behavior for Black–Scholes implied volatility in the fixed-maturity, large-strike, and fixed-maturity, small-strike regimes. Numerical studies were also presented to validate the theoretical work. We observed volatility smiles in the joint regime of large maturity and strike. The results showed that as the self-exciting intensity parameter (a or β) increases, which means more rare events tend to occur, the ATM volatility increases, and volatility smile tends to be more convex. Ratios of Black–Scholes implied volatility to log-moneyness in fixed-maturity large, small-strike, and large-strike regimes were also demonstrated. For a given maturity T, as the self-exciting parameter (a or β) increases, the ratio of implied volatility to log-moneyness increases. In these two regimes, we observed that the ratio of implied volatility to log-moneyness decreases as maturity increases, and this is usually detected on an implied volatility surface in practice.

Acknowledgements

Nian Yao was supported in part by the Natural Science Foundation of China(Grant No. 12071361), the Natural Science Foundation of Guangdong Province(Grant No. 2020A1515010822) and Shenzhen Natural Science Fund (the Stable Support Plan Program 20220810152104001).

References

[1] Aït-Sahalia, Y., Cacho-Diaz, J. and Laeven, R. J., Modeling financial contagion using mutually exciting jump processes, Journal of Financial Economics, 2015, 117(3): 585–606.

[2] Aït-Sahalia, Y. and Hurd, T. R., Portfolio choice in markets with contagion, Journal of Financial Economics, 2016, 14(1): 1–28.

[3] Aït-Sahalia, Y., Li, C. and Li, C. X., Closed-form implied volatility surfaces for stochastic volatility models with jumps, Journal of Econometrics, 2021, 222: 364–392.

[4] Alfonsi, A., Krief, D. and Tankov, P., Long-time large deviations for the multi-asset Wishart stochastic volatility model and option pricing, SIAM Journal on Financial Mathematics, 2019, 10(4): 942–976.

[5] Caravenna, F. and Corbetta, J., The asymptotic smile of a multiscaling stochastic volatility model, Stochastic Processes and their Applications, 2018, 128(3): 1034–1071.

[6] Dembo, A. and Zeitouni, O., Large Deviations Techniques and Applications, 2nd ed., Springer, New York, 1998.

[7] Djellout, H., Guillin, A. and Samoura, Y., Estimation of the realized (co-)volatility vector: Large deviations approach, Stochastic Processes and their Applications, 2017, 127(9): 2926–2960.

[8] Duffie, D., Pan, J. and Singleton, K., Transform analysis and asset pricing for affine jump-diffusions, Econometrica, 2000, 68(6): 1343–1376.

[9] Errais, E., Giesecke, K. and Goldberg, L. R., Affine point processes and portfolio credit risk, SIAM Journal on Financial Mathematics, 2010, 1(1): 642–665.

[10] Feng, X., He, L. and Liu, Z., Large deviation principles of realized Laplace transform of volatility, Journal of Theoretical Probability, 2022, 35: 186–208.

[11] Forde, M. and Jacquier, A., The large-maturity smile for the Heston model, Finance and Stochastics, 2011, 15: 755-–780.

[12] Forde, M. and Pogudin, A., The large-maturity smile for the Sabr and Cev-Heston models, International Journal of Theoretical and Applied Finance, 2013, 16(08): 1350047.

[13] Gao, X. and Zhu, L., Affine point processes: Refinements to large-time asymptotics, arXiv: 1903.06371, 2019.

[14] Gatheral, J., Hsu, E. P., Laurence, P., Ouyang, C. and Wang, T.-H., Asymptotics of implied volatility in local volatility models, Mathematical Finance, 2012, 22(4): 591–620.

[15] Hawkes, A. G., Point spectra of some mutually exciting point processes, Journal of the Royal Statistical Society Series B (Methodological), 1971, 33(3): 438–443.

[16] Hawkes, A. G., Spectra of some self-exciting and mutually exciting point processes, Biometrika, 1971, 58: 83–90.

[17] Jacquier, A., Keller-Ressel, M. and Mijatović, A., Large deviations and stochastic volatility with jumps: Asymptotic implied volatility for affine models, Stochastics, 2013, 85(2): 321–345.

[18] Jacquier, A. and Roome, P., Large-maturity regimes of the heston forward smile, Stochastic Processes and their Applications, 2016, 126(4): 1087–1123.

[19] Kanaya, S. and Otsu, T., Large deviations of realized volatility, Stochastic Processes and their Applications, 2012, 122(2): 546–581.

[20] Kou, S. G., A jump-diffusion model for option pricing, Management Science, 2002, 48(8): 1086–1101.

[21] Lee, R. W., The moment formula for implied volatility at extreme strikes, Mathematical Finance, 2004, 14(3): 469–480.

[22] Merton, R. C., Option pricing when underlying stock returns are discontinuous, Journal of Financial Economics, 1976, 3(1–2): 125–144.

[23] Mijatović, A. and Tankov, P., A new look at short-term implied volatility in asset price models with jumps, Mathematical Finance, 2016, 26(1): 149–183.

[24] Yao, N., Ling, Z., Zhang, J. and Xiao, M., Short maturity conditional Asian options in local volatility models, Mathematics and Financial Economics, 2020, 14: 307–328.

[25] Zhang, X., Blanchet, J., Giesecke, K. and Glynn, P. W., Affine point processes: Approximation and efficient simulation, Mathematics of Operations Research, 2015, 40(4): 797–819.

[26] Zhang, X. and Glynn, P. W., Affine jump-diffusions: Stochastic stability and limit theorems, arXiv: 1811.00122, 2018.

[27] Zhu, L., Nonlinear Hawkes Processes, PhD thesis, New York University, New York, 2013.

[28] Zhu, L., Limit theorems for a Cox-Ingersoll-Ross process with Hawkes jumps, Journal of Applied Probability, 2014, 51(3): 699–712.