Finance Stoch (2007) 11: 571–589
DOI 10.1007/s00780-007-0049-1

On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility

Elisa Alòs · Jorge A. León · Josep Vives

Received: 30 June 2006 / Accepted: 23 May 2007 / Published online: 8 August 2007

Abstract In this paper we use Malliavin calculus techniques to obtain an expression for the short-time behavior of the at-the-money implied volatility skew for a generalization of the Bates model, where the volatility does not need to be a diffusion or a Markov process, as the examples in Sect. 7 show. This expression depends on the derivative of the volatility in the sense of Malliavin calculus.

Keywords Black-Scholes formula · Derivative operator · Itô's formula for the Skorohod integral · Jump-diffusion stochastic volatility model

JEL Classification G12 · G13

Mathematics Subject Classification (2000) 91B28 · 91B70 · 60H07

E. Alòs' research is supported by grants MEC FEDER MTM 2006 06427 and SEJ2006-13537.
J.A. León's research is partially supported by the CONACyT grant 45684-F.
J. Vives' research is supported by grant MEC FEDER MTM 2006 06427.

E. Alòs (✉)
Dpt. d'Economia i Empresa, Universitat Pompeu Fabra, c/Ramon Trias Fargas, 25–27, 08005 Barcelona, Spain
e-mail: elisa.alos@upf.edu

J.A. León
Control Automático, CINVESTAV-IPN, Apartado Postal 14-740, 07000 México, D.F., Mexico

J. Vives
Dpt. de Matemàtiques, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain

J. Vives
Dpt. Probabilitat, Lògica i Estadística, Universitat de Barcelona, Gran Via, 585, 08007 Barcelona, Spain

1 Introduction

In the last years several authors have studied different extensions of the classical Black-Scholes model in order to explain the current market behavior. Among these extensions, one of the most popular allows the volatility to be a stochastic process (see, for example, [5, 14, 15, 23, 24], among others).

It is well known that classical stochastic volatility diffusion models, where the volatility also follows a diffusion process, capture some important features of the implied volatility—for example, its variation with respect to the strike price, described graphically as a *smile* or *skew* (see [21]). But the observed implied volatility exhibits dependence not only on the strike price, but also on the time to maturity (*term structure*). Unfortunately, the term structure is not easily explained by classical stochastic volatility models. For instance, a popular rule of thumb for the short-time behavior with respect to time to maturity, based on empirical observations, states that the skew slope is approximately $O((T-t)^{-\frac{1}{2}})$, while the rate for these stochastic volatility models is $O(1)$; see [18, 19], or [17]. Note that in these models, for reasonable coefficients in their dynamics, volatility behaves almost as a constant, on a very short time-scale. Consequently, returns are roughly normally distributed and the skew becomes quite flat. This problem has motivated the introduction of jumps in the asset price dynamic models. Although the rate of the skew slope for models with jumps is still $O(1)$, as is shown by Medvedev and Scaillet [19], they allow flexible modeling and generate skews and smiles similar to those observed in market data (see [6–8], or [9]). Recently, Fouque et al. [13] have introduced continuous diffusion models again to describe the empirical short-time skew. Their idea is to include suitable coefficients that depend on the time till the next maturity date and that guarantee the variability to be large enough near the maturity time.

The difficulties in fitting classical stochastic volatility models or models with jumps to observed marked prices have motivated, as an alternative approach, to model directly the implied volatility surfaces. Some recent research in modeling and existence issues for stochastic implied volatility models can be found in [16, 22], and the references therein.

The main goal of this paper is to provide a method based on the techniques of Malliavin calculus to estimate the rate of the short-dated behavior of the implied volatility (see Theorem 7 below) for general jump-diffusion stochastic volatility models, where the volatility does not need to be a diffusion or a Markov process. It is well known that the Malliavin calculus is a powerful tool to deal with anticipating processes. Since the future volatility is not adapted, this theory becomes a natural tool to analyze this problem. Hence, now it is possible to deal with a volatility in a class that includes fractional processes with parameter in $(0, 1)$, Markov processes and processes with time-varying coefficients, among others.

The paper is organized as follows. In Sect. 2 we introduce the framework and the notation that we utilize in this paper. In Sect. 3 we state our basic tool, namely, an anticipating Itô formula for the Skorohod integral. As a consequence, in Sect. 4, we obtain an extended Hull and White formula for a general class of jump-diffusion models with stochastic volatility. An expression for the derivative of the implied volatility is given in Sect. 5. Section 6 is devoted to the main result of this article. This means that we figure out the short-time limit behavior. Finally, in Sect. 7, we give some examples in order to show that we can not only extend some known results, but also consider new volatility models so that we are able to capture the short-time behavior of skew slopes of order $(T-t)^{\delta}$, for $\delta > -1/2$.

2 Statement of the model and notation

In this paper we consider the following model for the log-price of a stock under a risk-neutral probability measure Q:

$$X_t = x + (r - \lambda k)t - \frac{1}{2}\int_0^t \sigma_s^2\, ds + \int_0^t \sigma_s\big(\rho\, dW_s + \sqrt{1-\rho^2}\, dB_s\big) + Z_t, \quad t \in [0,T]. \tag{2.1}$$

Here x is the current log-price, r is the instantaneous interest rate, W and B are independent standard Brownian motions, $\rho \in (-1,1)$ and Z is a compound Poisson process with intensity λ, Lévy measure ν, independent of W and B, and with $k = \frac{1}{\lambda}\int_{\mathbb{R}}(e^y - 1)\nu(dy) < \infty$. The volatility process σ is a square-integrable stochastic process with right-continuous trajectories and adapted to the filtration generated by W. In some parts of the paper we shall assume, in addition, that its trajectories are bounded below by a positive constant and that the process satisfies some suitable conditions in the Malliavin calculus sense (see hypotheses (H1)–(H5) below).

Notice that this model is a generalization of the classical Bates model introduced in [8], in the sense that we do not assume the volatility to be a diffusion process.

It is well known that any stopping time with respect to a Brownian filtration is predictable. So, an extension of our results below allows the volatility to have non-predictable jump times as advocated by Bakshi et al. [4] and Duffie et al. [11], among others. In this case we have to use Malliavin calculus for Lévy processes. The details of this extension are in preparation and will appear elsewhere.

In the following we denote by $\mathcal{F}^W$, $\mathcal{F}^B$ and $\mathcal{F}^Z$ the filtrations generated by W, B and Z, respectively. Moreover we define $\mathcal{F} := \mathcal{F}^W \vee \mathcal{F}^B \vee \mathcal{F}^Z$.

It is well known that if we price a European call with strike price K by the formula

$$V_t = e^{-r(T-t)} E\big((e^{X_T} - K)_+ | \mathcal{F}_t\big), \tag{2.2}$$

where E is the expectation with respect to Q, there is no arbitrage opportunity. Thus, V_t is a possible price for this derivative. Notice that any allowable choice of Q leads to an equivalent martingale measure and to a different no arbitrage price. The approach that we follow here is the same as in [12], where it is assumed that the market selects a unique equivalent martingale measure under which derivative contracts are priced.

In the sequel we use the following notation:

- $v_t := (\frac{Y_t}{T-t})^{\frac{1}{2}}$, with $Y_t := \int_t^T \sigma_s^2\, ds$, will denote the future average volatility.
- For any $\tau > 0$, $p(x,\tau)$ will denote the centered Gaussian kernel with variance τ^2. If $\tau = 1$ we write $p(x)$.

- $\mathrm{BS}(t, x, \sigma)$ will denote the price of a European call option under the classical Black-Scholes model with constant volatility σ, current log stock price x, time to maturity $T-t$, strike price K and interest rate r. Remember that in this case:

$$\mathrm{BS}(t, x, \sigma) = e^x N(d_+) - K e^{-r(T-t)} N(d_-),$$

where N denotes the cumulative probability function of the standard normal law and

$$d_\pm := \frac{x - x_t^*}{\sigma\sqrt{T-t}} \pm \frac{\sigma}{2}\sqrt{T-t},$$

with $x_t^* := \ln K - r(T-t)$.
- $\mathcal{L}_{\mathrm{BS}}(\sigma)$ will denote the Black-Scholes differential operator, in the log variable, with volatility σ:

$$\mathcal{L}_{\mathrm{BS}}(\sigma) = \partial_t + \frac{1}{2}\sigma^2\partial_{xx}^2 + \left(r - \frac{1}{2}\sigma^2\right)\partial_x - r.$$

It is well known that $\mathcal{L}_{\mathrm{BS}}(\sigma)\,\mathrm{BS}(\cdot, \cdot, \sigma) = 0$.
- $G(t, x, \sigma) := (\partial_{xx}^2 - \partial_x)\,\mathrm{BS}(t, x, \sigma)$.

3 An anticipating Itô formula

First of all, we describe the basic notation that is used in this article. For this, we assume that the reader is familiar with the elementary results of Malliavin calculus, as given for instance in [20].

Let us consider a standard Brownian motion $W = \{W_t,\ t \in [0, T]\}$ defined in a complete probability space $(\Omega, \mathcal{F}, P)$. The set $\mathbb{D}_W^{1,2}$ will denote the domain of the derivative operator D^W. It is well known that $\mathbb{D}_W^{1,2}$ is a dense subset of $L^2(\Omega)$ and that D^W is a closed and unbounded operator from $L^2(\Omega)$ to $L^2([0, T] \times \Omega)$. We also consider the iterated derivatives $D^{W,n}$, for $n > 1$, whose domains will be denoted by $\mathbb{D}_W^{n,2}$.

The adjoint of the derivative operator D^W, denoted by δ^W, is an extension of the Itô integral in the sense that the set $L_a^2([0, T] \times \Omega)$ of square integrable and adapted processes (with respect to the filtration generated by W) is included in $\mathrm{Dom}\,\delta^W$ and the operator δ^W restricted to $L_a^2([0, T] \times \Omega)$ coincides with the Itô integral. We shall use the notation $\delta^W(u) = \int_0^T u_t\,dW_t$. We recall that $\mathbb{L}^{n,2} := L^2([0, T]; \mathbb{D}_W^{n,2})$ is included in the domain of δ^W for all $n \geq 1$.

Now we can establish the following Itô formula, which follows from [1, 2], and is the main tool of this paper. In the sequel we use the notation $D = D^W$ to simplify the exposition.

Proposition 3.1 *Assume the model (2.1) and $\sigma \in \mathbb{L}^{1,2}$. Let $F : [0, T] \times \mathbb{R}^2 \to \mathbb{R}$ be a function in $C^{1,2}([0, T] \times \mathbb{R}^2)$ such that there exists a positive constant C such that,* *for all $t \in [0,T]$, F and its partial derivatives evaluated in (t, X_t, Y_t) are bounded by C. Then it follows that*

$$\begin{aligned}
F(t, X_t, Y_t) = F(0, X_0, Y_0) &+ \int_0^t \partial_s F(s, X_s, Y_s)\, ds \\
&+ \int_0^t \partial_x F(s, X_s, Y_s)\left(r - \lambda k - \frac{\sigma_s^2}{2}\right) ds \\
&+ \int_0^t \partial_x F(s, X_s, Y_s)\sigma_s\left(\rho\, dW_s + \sqrt{1-\rho^2}\, dB_s\right) \\
&- \int_0^t \partial_y F(s, X_s, Y_s)\sigma_s^2\, ds + \rho \int_0^t \partial_{xy}^2 F(s, X_s, Y_s)\Lambda_s\, ds \\
&+ \frac{1}{2}\int_0^t \partial_{xx}^2 F(s, X_s, Y_s)\sigma_s^2\, ds \\
&+ \int_0^t \int_{\mathbb{R}} \left(F(s, X_{s-} + y, Y_s) - F(s, X_{s-}, Y_s)\right)\tilde{J}_X(ds, dy) \\
&+ \int_0^t \int_{\mathbb{R}} \left(F(s, X_{s-} + y, Y_s) - F(s, X_{s-}, Y_s)\right) ds\, \nu(dy),
\end{aligned}$$

where $\Lambda_s := (\int_s^T D_s\sigma_r^2\, dr)\sigma_s$, J_X is the Poisson random measure such that $Z_t = \int_{[0,t]\times\mathbb{R}} y J_X(ds, dy)$ and $\tilde{J}_X(ds, dy) := J_X(ds, dy) - ds\, \nu(dy)$.

Proof Denote by T_i, $i = 1, \dots, N_T$, the jump instants of X. On $[T_i, T_{i+1})$, X evolves according to

$$dX_t^c = \left(r - \lambda k - \frac{\sigma_t^2}{2}\right) dt + \sigma_t\left(\rho\, dW_t + \sqrt{1-\rho^2}\, dB_t\right).$$

Then, applying Theorem 1 in [1], and using the fact that Z is independent of W and B, we have that

$$\begin{aligned}
&F(T_{i+1-}, X_{T_{i+1-}}, Y_{T_{i+1}-}) - F(T_i, X_{T_i}, Y_{T_i}) \\
&= \int_{T_i}^{T_{i+1}-} \partial_s F(s, X_s, Y_s)\, ds + \int_{T_i}^{T_{i+1}-} \partial_x F(s, X_s, Y_s)\, dX_s^c \\
&\quad - \int_{T_i}^{T_{i+1}-} \partial_y F(s, X_s, Y_s)\sigma_s^2\, ds + \rho \int_{T_i}^{T_{i+1}-} \partial_{xy}^2 F(s, X_s, Y_s)\Lambda_s\, ds \\
&\quad + \frac{1}{2}\int_{T_i}^{T_{i+1}-} \partial_{xx}^2 F(s, X_s, Y_s)\sigma_s^2\, ds.
\end{aligned}$$

Note that if a jump of size ΔX_t occurs then the resulting change in $F(t, X_t, Y_t)$ is given by $F(t, X_{t-} + \Delta X_t, Y_t) - F(t, X_{t-}, Y_t)$. Therefore, the total change in

$F(t, X_t, Y_t)$ can be written as the sum of these two contributions. Thus, we deduce the desired result. □

4 An extension of the Hull and White formula

In this section, using the Itô formula and the arguments developed in [1], we prove an extension of the Hull and White formula that gives the price of a European call option as a sum of the price when the model has no jumps and no correlation plus three terms: one describing the impact of the correlation on option prices and two of them, which can be presented jointly, describing the impact of jumps on these prices. Hence, this formula will be a useful tool to compare the effect of correlation and jumps (see Sect. 5).

We need the following result, inspired by Lemma 5.2 in [12].

Lemma 4.1 *Let* $0 \le t \le s < T$ *and* $\mathcal{G}_t := \mathcal{F}_t \vee \mathcal{F}_T^W \vee \mathcal{F}_T^Z$. *Then for every* $n \ge 0$, *there exists* $C = C(n, \rho)$ *such that*

$$\left| E\left(\partial_x^n G(s, X_s, v_s) | \mathcal{G}_t\right)\right| \le C \left(\int_t^T \sigma_s^2 \, ds\right)^{-\frac{1}{2}(n+1)}.$$

Proof A simple calculation gives

$$G(s, X_s, v_s) = K e^{-r(T-s)} p\left(X_s - \mu, v_s \sqrt{T-s}\right),$$

where $\mu = \ln K - (r - v_s^2/2)(T-s)$. This allows us to write

$$E\left(\partial_x^n G(s, X_s, v_s) | \mathcal{G}_t\right) = (-1)^n K e^{-r(T-s)} \partial_\mu^n E\left(p\left(X_s - \mu, v_s\sqrt{T-s}\right) | \mathcal{G}_t\right). \quad (4.1)$$

Since the conditional expectation of X_s given $\mathcal{G}_t$ is a normal random variable with mean

$$\phi = X_t + \int_t^s \left(r - \sigma_\theta^2/2\right) d\theta + Z_s - Z_t - \lambda k(s-t) + \rho \int_t^s \sigma_\theta \, dW_\theta$$

and variance $(1 - \rho^2) \int_t^s \sigma_\theta^2 \, d\theta$, and using the semigroup property of the Gaussian density function, it follows that

$$E\left(p\left(X_s - \mu, v_s\sqrt{T-s}\right) | \mathcal{G}_t\right) = p\left(\phi - \mu, \sqrt{(1-\rho^2) \int_t^T \sigma_\theta^2 \, d\theta + \rho^2 \int_s^T \sigma_\theta^2 \, d\theta}\right).$$

Putting this result in (4.1), we have

$$\begin{aligned} &E\left(\partial_x^n G(s, X_s, v_s) | \mathcal{G}_t\right) \\ &\quad = (-1)^n K e^{-r(T-s)} \partial_\mu^n p\left(\phi - \mu, \sqrt{(1-\rho^2) \int_t^T \sigma_\theta^2 \, d\theta + \rho^2 \int_s^T \sigma_\theta^2 \, d\theta}\right). \end{aligned}$$

A simple calculation and the fact that, for all positive constants c, d, the function $x^c e^{-dx^2}$ is bounded, give us that

$$\left|\partial_\mu^n p\left(\phi-\mu, \sqrt{(1-\rho^2)\int_t^T \sigma_s^2\,ds+\rho^2\int_s^T \sigma_s^2\,ds}\right)\right|$$
$$\leq C\left((1-\rho^2)\int_t^T \sigma_s^2\,ds+\rho^2\int_s^T \sigma_s^2\,ds\right)^{-\frac{1}{2}(n+1)} \leq C\left(\int_t^T \sigma_s^2\,ds\right)^{-\frac{1}{2}(n+1)},$$

and thus the proof is complete. □

Now we are able to prove the main result of this section, the extended Hull and White formula.

Theorem 4.2 *Assume the model* (2.1) *holds with $\sigma \in \mathbb{L}^{1,2}$. Then it follows that*

$$\begin{aligned} V_t &= E\big(\mathrm{BS}(t, X_t, v_t)|\mathcal{F}_t\big)+\frac{\rho}{2}E\left(\int_t^T e^{-r(s-t)}\partial_x G(s, X_s, v_s)\Lambda_s\,ds\,\bigg|\,\mathcal{F}_t\right)\\ &\quad+E\left(\int_t^T\int_{\mathbb{R}} e^{-r(s-t)}\big(\mathrm{BS}(s, X_s+y, v_s)-\mathrm{BS}(s, X_s, v_s)\big)\nu(dy)\,ds\,\bigg|\,\mathcal{F}_t\right)\\ &\quad-\lambda k E\left(\int_t^T e^{-r(s-t)}\partial_x \mathrm{BS}(s, X_s, v_s)\,ds\,\bigg|\,\mathcal{F}_t\right). \end{aligned}$$

Proof This proof is similar to the one of the main theorem in [1], so we only sketch it. Notice that $\mathrm{BS}(T, X_T, v_T)=V_T$. Then, from (2.2), we have

$$e^{-rt}V_t=E\big(e^{-rT}\mathrm{BS}(T, X_T, v_T)|\mathcal{F}_t\big).$$

Now our idea is to apply Proposition 3.1 to the process $e^{-rt}\mathrm{BS}(t, X_t, v_t)$. As the derivatives of $\mathrm{BS}(t, x, \sigma)$ are not bounded, we use an approximating argument, changing v_t to

$$v_t^\delta := \sqrt{\frac{1}{T-t}(Y_t+\delta)},$$

and $\mathrm{BS}(t, x, \sigma)$ to $\mathrm{BS}_n(t, x, \sigma) := \mathrm{BS}(t, x, \sigma)\psi_n(x)$, where $\psi_n(x) := \phi(\frac{1}{n}x)$, for some $\phi \in \mathcal{C}_b^2$ such that $\phi(x)=1$ for all $|x|<1$ and $\phi(x)=0$ for all $|x|>2$. Applying Proposition 3.1 between t and T, proceeding as in Theorem 3 in [1] and observing that

$$\mathcal{L}_{\mathrm{BS}}(\sigma_s)\,\mathrm{BS}_n\big(s, X_s, v_s^\delta\big)=\big(\mathcal{L}_{\mathrm{BS}}(\sigma_s)\,\mathrm{BS}\big(s, X_s, v_s^\delta\big)\big)\psi_n(X_s)+A_n(s),$$

where

$$\begin{aligned} A_n(s) &= \frac{1}{2}\sigma_s^2\big[2\partial_x \mathrm{BS}\big(s, X_s, v_s^\delta\big)\psi_n'(X_s)+\mathrm{BS}\big(s, X_s, v_s^\delta\big)\big(\psi_n''(X_s)-\psi_n'(X_s)\big)\big]\\ &\quad+r\,\mathrm{BS}\big(s, X_s, v_s^\delta\big)\psi_n'(X_s), \end{aligned}$$

we obtain

$$
\begin{aligned}
&E\big(e^{-rT}\,\mathrm{BS}_n\big(T,X_T,v_T^\delta\big)|\mathcal{F}_t\big)\\
&\quad = E\bigg(e^{-rt}\,\mathrm{BS}_n\big(t,X_t,v_t^\delta\big)+\int_t^T e^{-rs}A_n(s)\,ds-\lambda k\int_t^T e^{-rs}\partial_x\,\mathrm{BS}_n\big(s,X_s,v_s^\delta\big)\,ds\\
&\qquad +\frac{\rho}{2}\int_t^T e^{-rs}\big[(\partial_x G)\big(s,X_s,v_s^\delta\big)\psi_n(X_s)+G\big(s,X_s,v_s^\delta\big)\psi_n'(X_s)\big]\Lambda_s\,ds\\
&\qquad +\int_t^T\int_{\mathbb{R}} e^{-rs}\big(\mathrm{BS}_n\big(s,X_{s-}+y,v_s^\delta\big)-\mathrm{BS}_n\big(s,X_{s-},v_s^\delta\big)\big)\nu(dy)\,ds\,\bigg|\,\mathcal{F}_t\bigg).
\end{aligned}
$$

Now, letting first $n\uparrow\infty$ and then $\delta\downarrow 0$, using Lemma 4.1 and dominated convergence arguments, the result follows. □

5 An expression for the derivative of the implied volatility

Let $I_t(X_t)$ denote the implied volatility process, which satisfies $V_t=\mathrm{BS}(t,X_t,I_t(X_t))$, by definition. In this section we prove a formula for its at-the-money derivative that we use in Sect. 6 to study the short-time behavior of the implied volatility.

Proposition 5.1 *Assume the model (2.1) holds with $\sigma\in\mathbb{L}^{1,2}$ and for every fixed $t\in[0,T)$, $E(\int_t^T\sigma_s^2\,ds|\mathcal{F}_t)^{-1}<\infty$ a.s. Then it follows that*

$$
\frac{\partial I_t}{\partial X_t}(x_t^*)=\frac{E(\int_t^T(\partial_x F(s,X_s,v_s)-\frac{1}{2}F(s,X_s,v_s))\,ds|\mathcal{F}_t)}{\partial_\sigma\,\mathrm{BS}(t,x_t^*,I_t(x_t^*))}\bigg|_{X_t=x_t^*},\qquad a.s.,
$$

where

$$
\begin{aligned}
F(s,X_s,v_s):={}&\frac{\rho}{2}e^{-r(s-t)}\partial_x G(s,X_s,v_s)\Lambda_s\\
&+\int_{\mathbb{R}} e^{-r(s-t)}\big[\mathrm{BS}(s,X_s+y,v_s)-\mathrm{BS}(s,X_s,v_s)\big]\nu(dy)\\
&-\lambda k e^{-r(s-t)}\partial_x\mathrm{BS}(s,X_s,v_s).
\end{aligned}
$$

Proof Taking partial derivatives of the expression $V_t=\mathrm{BS}(t,X_t,I_t(X_t))$ with respect to X_t, we obtain

$$
\frac{\partial V_t}{\partial X_t}=\partial_x\,\mathrm{BS}\big(t,X_t,I_t(X_t)\big)+\partial_\sigma\,\mathrm{BS}\big(t,X_t,I_t(X_t)\big)\frac{\partial I_t}{\partial X_t}(X_t).\tag{5.1}
$$

On the other hand, from Theorem 4.2 we deduce that

$$
V_t=E\big(\mathrm{BS}(t,X_t,v_t)|\mathcal{F}_t\big)+E\bigg(\int_t^T F(s,X_s,v_s)\,ds\,\bigg|\,\mathcal{F}_t\bigg),
$$

which implies that

$$\frac{\partial V_t}{\partial X_t} = E\big(\partial_x \operatorname{BS}(t, X_t, v_t)|\mathcal{F}_t\big) + E\left(\int_t^T \partial_x F(s, X_s, v_s)\, ds \,\middle|\, \mathcal{F}_t\right). \tag{5.2}$$

Using now the fact that $E(\int_t^T \sigma_s^2\, ds|\mathcal{F}_t)^{-1} < \infty$ we can check that the conditional expectation $E(\int_t^T \partial_x F(s, X_s, v_s)\, ds|\mathcal{F}_t)$ is well-defined and finite a.s. Thus, (5.1) and (5.2) imply

$$\frac{\partial I_t}{\partial X_t}(x_t^*) = \left.\frac{E(\partial_x \operatorname{BS}(t,x_t^*,v_t)|\mathcal{F}_t) - \partial_x \operatorname{BS}(t,x_t^*,I_t(x_t^*)) + E(\int_t^T \partial_x F(s,X_s,v_s)\, ds|\mathcal{F}_t)}{\partial_\sigma \operatorname{BS}(t,x_t^*,I_t(x_t^*))}\right|_{X_t = x_t^*}. \tag{5.3}$$

Notice that

$$\begin{aligned} E\big(\partial_x \operatorname{BS}(t, x_t^*, v_t)|\mathcal{F}_t\big) &= \partial_x E\big(\operatorname{BS}(t, x, v_t)|\mathcal{F}_t\big)\big|_{x=x_t^*} \\ &= \partial_x \operatorname{BS}\big(t, x, I_t^0(x)\big)\big|_{x=x_t^*}, \end{aligned} \tag{5.4}$$

where $I_t^0(X_t)$ is the implied volatility in the case $\rho = \lambda = 0$.

Also, by the classical Hull and White formula, we have

$$\partial_x \big(\operatorname{BS}\big(t, x, I_t^0(x)\big)\big)\big|_{x=x_t^*} = \partial_x \operatorname{BS}\big(t, x^*, I_t^0(x^*)\big) + \partial_\sigma \operatorname{BS}\big(t, x^*, I_t^0(x^*)\big) \frac{\partial I_t^0}{\partial x}(x_t^*). \tag{5.5}$$

From [21] we know that $\frac{\partial I_t^0}{\partial x}(x_t^*) = 0$. Then, (5.3), (5.4), and (5.5) imply that

$$\frac{\partial I_t}{\partial X_t}(x_t^*) = \left.\frac{\partial_x \operatorname{BS}(t,x_t^*,I_t^0(x_t^*)) - \partial_x \operatorname{BS}(t,x_t^*,I_t(x_t^*)) + E(\int_t^T \partial_x F(s,X_s,v_s)\, ds|\mathcal{F}_t)}{\partial_\sigma \operatorname{BS}(t,x_t^*,I_t(x_t^*))}\right|_{X_t = x_t^*}. \tag{5.6}$$

On the other hand, straightforward calculations lead us to

$$\partial_x \operatorname{BS}(t, x_t^*, \sigma) = e^{x_t^*} N\left(\frac{1}{2}\sigma\sqrt{T-t}\right)$$

and

$$\operatorname{BS}(t, x_t^*, \sigma) = e^{x_t^*}\left(N\left(\frac{1}{2}\sigma\sqrt{T-t}\right) - N\left(-\frac{1}{2}\sigma\sqrt{T-t}\right)\right).$$

Then

$$\partial_x \operatorname{BS}(t, x_t^*, \sigma) = \frac{1}{2}\big(e^{x_t^*} + \operatorname{BS}(t, x_t^*, \sigma)\big)$$

and

$$\begin{aligned}
&\partial_x \operatorname{BS}\big(t, x_t^*, I_t^0(x_t^*)\big) - \partial_x \operatorname{BS}\big(t, x_t^*, I_t(x_t^*)\big)\\
&\quad= \frac{1}{2}\big(\operatorname{BS}\big(t, x_t^*, I_t^0(x_t^*)\big) - \operatorname{BS}\big(t, x_t^*, I_t(x_t^*)\big)\big)\\
&\quad= \frac{1}{2}\big(E\big(\operatorname{BS}(t, x_t^*, v_t) - V_t(x_t^*)|\mathcal{F}_t\big)\big) = -\frac{1}{2}E\left(\int_t^T F(s, X_s, v_s)\,ds \,\middle|\, \mathcal{F}_t\right)\Big|_{X_t=x_t^*}.
\end{aligned}$$

This, together with (5.6), implies the result. □

6 Short-time limit behavior

Here our purpose is to study the limit of $\frac{\partial I_t}{\partial X_t}(x_t^*)$ when $T \downarrow t$. To this end, we need the following lemma:

Lemma 6.1 *Assume the model* (2.1) *is satisfied. Then* $I_t(x_t^*)\sqrt{T-t}$ *tends to* 0 *a.s., as* $T \to t$.

Proof Using the dominated convergence theorem it is easy to see that

$$P_t := E\big(e^{-r(T-t)}\big(K - e^{X_T}\big)_+|\mathcal{F}_t\big)\big|_{X_t=x_t^*} \underset{T\to t}{\longrightarrow} \big(K - e^{x_t^*}\big)_+ = 0, \quad \text{a.s.}$$

Now, by the classical call-put parity relation, we obtain

$$V_t = E\big(e^{-r(T-t)}\big(e^{X_T} - K\big)_+|\mathcal{F}_t\big)\big|_{X_t=x_t^*} \longrightarrow \big(e^{x_t^*} - K\big)_+ = 0.$$

Hence, taking into account that, in the at-the-money case, $V_t = \operatorname{BS}(t, x_t^*, I_t(x_t^*))$, we deduce that

$$\operatorname{BS}\big(t, x_t^*, I_t(x_t^*)\big) = 2Ke^{-r(T-t)}\left[N\left(\frac{I(x_t^*)\sqrt{T-t}}{2}\right) - \frac{1}{2}\right] \longrightarrow 0,$$

and this allows us to complete the proof. □

Henceforth, we consider the following hypotheses:

(H1) $\sigma \in \mathbb{L}^{2,4}$.
(H2) There exists a constant $a > 0$ such that $\sigma > a$.
(H3) There exists a constant $\delta > -\frac{1}{2}$ such that, for all $0 < t < s < r < T$,

$$E\big((D_s\sigma_r)^2|\mathcal{F}_t\big) \le C(r-s)^{2\delta}, \tag{6.1}$$

$$E\big((D_\theta D_s\sigma_r)^2|\mathcal{F}_t\big) \le C(r-s)^{2\delta}(r-\theta)^{-2\delta}. \tag{6.2}$$

Proposition 6.2 *Assume that the model* (2.1) *and hypotheses* (H1)–(H3) *hold. Then*

$$
\begin{aligned}
&\partial_\sigma \mathrm{BS}\big(t, x_t^*, I_t(x_t^*)\big) \frac{\partial I_t}{\partial X_t}(x_t^*) \\
&\quad = \frac{\rho}{2} E\bigg(L(t, x_t^*, v_t) \int_t^T \Lambda_s \, ds \,\bigg|\, \mathcal{F}_t \bigg) \\
&\qquad - \lambda k E\big(G(t, x_t^*, v_t)(T-t) | \mathcal{F}_t \big) + O(T-t)^{(1+2\delta)\wedge 1},
\end{aligned}
$$

as $T \to t$ and where $L(t, x_t^*, v_t) = (\partial_{xx}^2 - \frac{1}{2}\partial_x) G(t, x_t^*, v_t)$.

Proof Proposition 5.1 gives us that

$$
\begin{aligned}
&\partial_\sigma \mathrm{BS}\big(t, x_t^*, I_t(x_t^*)\big) \frac{\partial I_t}{\partial X_t}(x_t^*) \\
&= \frac{\rho}{2} E\bigg(\int_t^T e^{-r(s-t)} \Big(\partial_x - \frac{1}{2} \Big) \partial_x G(s, X_s, v_s) \Lambda_s \, ds \,\bigg|\, \mathcal{F}_t \bigg) \bigg|_{X_t = x_t^*} \\
&\quad + E\bigg(\int_t^T \int_{\mathbb{R}} e^{-r(s-t)} \Big(\partial_x - \frac{1}{2} \Big) \\
&\qquad\quad \times \big[\mathrm{BS}(s, X_s + y, v_s) - \mathrm{BS}(s, X_s, v_s) \big] \nu(dy) \, ds \,\bigg|\, \mathcal{F}_t \bigg) \bigg|_{X_t = x_t^*} \\
&\quad - \lambda k E\bigg(\int_t^T e^{-r(s-t)} \Big(\partial_x - \frac{1}{2} \Big) \partial_x \mathrm{BS}(s, X_s, v_s) \, ds \,\bigg|\, \mathcal{F}_t \bigg) \bigg|_{X_t = x_t^*} \\
&= T_1 + T_2 + T_3.
\end{aligned}
\tag{6.3}
$$

Now the proof will be decomposed into several steps.

Step 1. Here we claim that

$$
T_1 = \frac{\rho}{2} E\bigg(L(t, x_t^*, v_t) \int_t^T \Lambda_s \, ds \,\bigg|\, \mathcal{F}_t \bigg) + O(T-t)^{1+2\delta}, \tag{6.4}
$$

where $L(s, X_s, v_s) = (\partial_{xx}^2 - \frac{1}{2}\partial_x) G(s, X_s, v_s)$. In fact, applying Itô's formula to

$$
\frac{\rho}{2} e^{-r(s-t)} L(s, X_s, v_s) \bigg(\int_s^T \Lambda_r \, dr \bigg)
$$

as in the proof of Theorem 4.2 and taking conditional expectations with respect to $\mathcal{F}_t$, we obtain that

$$
\begin{aligned}
&\frac{\rho}{2} E\bigg(\int_t^T e^{-r(s-t)} L(s, X_s, v_s) \Lambda_s \, ds \,\bigg|\, \mathcal{F}_t \bigg) \\
&\quad = \frac{\rho}{2} E\bigg(L(t, X_t, v_t) \bigg(\int_t^T \Lambda_s \, ds \bigg) \,\bigg|\, \mathcal{F}_t \bigg) \\
&\qquad + \frac{\rho^2}{4} E\bigg(\int_t^T e^{-r(s-t)} (\partial_{xxx}^3 - \partial_{xx}^2) L(s, X_s, v_s) \bigg(\int_s^T \Lambda_r \, dr \bigg) \Lambda_s \, ds \,\bigg|\, \mathcal{F}_t \bigg)
\end{aligned}
$$

$$+\frac{\rho^2}{2}E\bigg(\int_t^T e^{-r(s-t)}\partial_x L(s,X_s,v_s)\bigg(\int_s^T D_s\Lambda_r\,dr\bigg)\sigma_s\,ds\,\bigg|\,\mathcal{F}_t\bigg)$$

$$+\frac{\rho}{2}E\bigg(\int_t^T\int_{\mathbb{R}} e^{-r(s-t)}\big[L(s,X_{s-}+y,v_s)-L(s,X_{s-},v_s)\big]$$

$$\times\bigg(\int_s^T \Lambda_r\,dr\bigg)\nu(dy)\,ds\,\bigg|\,\mathcal{F}_t\bigg)$$

$$-\lambda k\frac{\rho}{2}E\bigg(\int_t^T e^{-r(s-t)}\partial_x L(s,X_{s-},v_s)\bigg(\int_s^T \Lambda_r\,dr\bigg)ds\,\bigg|\,\mathcal{F}_t\bigg)$$

$$=\frac{\rho}{2}E\bigg(L(t,X_t,v_t)\bigg(\int_t^T \Lambda_s\,ds\bigg)\,\bigg|\,\mathcal{F}_t\bigg)+S_1+S_2+S_3+S_4.$$

Using Lemma 4.1 we can write

$$S_1=\frac{\rho^2}{4}E\bigg(\int_t^T e^{-r(s-t)}E\big(\big(\partial^3_{xxx}-\partial^2_{xx}\big)L(s,X_s,v_s)|\mathcal{G}_t\big)\bigg(\int_s^T \Lambda_r\,dr\bigg)\Lambda_s\,ds\,\bigg|\,\mathcal{F}_t\bigg)$$

$$\le C\sum_{k=4}^{6}E\bigg[\bigg(\int_t^T\sigma_s^2\,ds\bigg)^{-\frac{k}{2}}\int_t^T\bigg|\bigg(\int_s^T\Lambda_r\,dr\bigg)\Lambda_s\bigg|ds\,\bigg|\,\mathcal{F}_t\bigg].$$

Hence, using hypotheses (H2) and (H3) we can write

$$S_1\le C\sum_{k=0}^{2}E\bigg(\bigg(\int_t^T\sigma_\theta^2\,d\theta\bigg)^{-\frac{k}{2}}\bigg(\int_t^T\int_t^T(D_r\sigma_\theta)^2\,dr\,d\theta\bigg)\,\bigg|\,\mathcal{F}_t\bigg)$$

$$\le C(T-t)^{-1}\bigg(\int_t^T\int_t^\theta(\theta-r)^{2\delta}\,dr\,d\theta\bigg)\le C(T-t)^{1+2\delta}.$$

Using similar arguments it follows that $S_2+S_3+S_4=O(T-t)^{1+2\delta}$, which proves (6.4).

Step 2. As $|\mathrm{BS}(t,x,\sigma)|+|\partial_x\mathrm{BS}(t,x,\sigma)|\le 2e^x+K$ it follows that $T_2=O(T-t)$.

Step 3. Let us prove that

$$T_3=-\lambda k E\big(G(t,x_t^*,v_t)(T-t)|\mathcal{F}_t\big)+O(T-t).\tag{6.5}$$

In fact,

$$E\bigg(\int_t^T e^{-r(s-t)}\bigg(\partial_x-\frac{1}{2}\bigg)\partial_x\mathrm{BS}(s,X_s,v_s)\,ds\,\bigg|\,\mathcal{F}_t\bigg)\bigg|_{X_t=x_t^*}$$

$$=E\bigg(\int_t^T e^{-r(s-t)}G(s,X_s,v_s)\,ds\,\bigg|\,\mathcal{F}_t\bigg)\bigg|_{X_t=x_t^*}$$

$$+\frac{1}{2}E\bigg(\int_t^T e^{-r(s-t)}\partial_x\mathrm{BS}(s,X_s,v_s)\,ds\,\bigg|\,\mathcal{F}_t\bigg)\bigg|_{X_t=x_t^*}.$$

As $|\partial_x \mathrm{BS}(t,x,\sigma)| \le e^x$ it follows easily that the second term on the right-hand side of this equality is $O(T-t)$. On the other hand, Itô's formula allows us to write

$$
\begin{aligned}
&E\big(e^{-r(s-t)}G(s,X_s,v_s)|\mathcal{F}_t\big)\\
&\quad = E\big(G(t,X_t,v_t)|\mathcal{F}_t\big) + \frac{\rho}{2}E\left(\int_t^s e^{-r(u-t)}\big(\partial^3_{xxx}-\partial^2_{xx}\big)G(u,X_u,v_u)\Lambda_u\,du\,\middle|\,\mathcal{F}_t\right)\\
&\qquad + E\left(\int_t^s\int_R e^{-r(u-t)}\big(G(u,X_u+y,v_u)-G(u,X_u,v_u)\big)\nu(dy)\,du\,\middle|\,\mathcal{F}_t\right)\\
&\qquad - \lambda k E\left(\int_t^s e^{-r(u-t)}\partial_x G(u,X_u,v_u)\,du\,\middle|\,\mathcal{F}_t\right).
\end{aligned}
$$

Now, using again the same arguments as in Step 1, (6.5) follows. Therefore, the proof is complete. □

Now we can state the main result of this paper. We consider the following hypotheses:

(H4) σ has a.s. right-continuous trajectories.
(H5) For every fixed $t>0$, $\sup_{s,r,\theta\in[t,T]}E((\sigma_s\sigma_r-\sigma_\theta^2)^2|\mathcal{F}_t)\to 0$, as $T\to t$.

Theorem 6.3 *Consider the model* (2.1) *and suppose that hypotheses* (H1)–(H5) *hold.*

1. *Assume that δ in* (H3) *is nonnegative and that there exists an $\mathcal{F}_t$-measurable random variable $D_t^+\sigma_t$ such that, for every $t>0$,*
$$\sup_{s,r\in[t,T]}\big|E\big(\big(D_s\sigma_r-D_t^+\sigma_t\big)\big|\mathcal{F}_t\big)\big|\to 0,\quad a.s., \tag{6.6}$$
as $T\to t$. Then
$$\lim_{T\to t}\frac{\partial I_t}{\partial X_t}(x_t^*) = -\frac{1}{\sigma_t}\left(\lambda k+\rho\frac{D_t^+\sigma_t}{2}\right). \tag{6.7}$$
2. *Assume that δ in* (H3) *is negative and that there exists an $\mathcal{F}_t$-measurable random variable $L_t^{\delta,+}\sigma_t$ such that, for every $t>0$,*
$$\frac{1}{(T-t)^{2+\delta}}\int_t^T\int_s^T E(D_s\sigma_r|\mathcal{F}_t)\,dr\,ds - L_t^{\delta,+}\sigma_t\to 0,\quad a.s., \tag{6.8}$$
as $T\to t$. Then
$$\lim_{T\to t}(T-t)^{-\delta}\frac{\partial I_t}{\partial X_t}(x_t^*) = -\frac{\rho}{\sigma_t}L_t^{\delta,+}\sigma_t. \tag{6.9}$$

Proof Using Proposition 6.2 and the facts that

$$\partial_\sigma \mathrm{BS}\big(t,x_t^*,I_t(x_t^*)\big) = \frac{Ke^{-r(T-t)}e^{\frac{-I_t(x_t^*)^2(T-t)}{8}}\sqrt{T-t}}{\sqrt{2\pi}},$$

$$L(t,x_t^*,v_t) = -Ke^{-r(T-t)}\frac{1}{\sqrt{2\pi}}e^{-\frac{v_t^2(T-t)}{8}}v_t^{-3}(T-t)^{-\frac{3}{2}}$$

and

$$G(t,x_t^*,v_t)=\frac{Ke^{-r(T-t)}e^{\frac{-v_t{}^2(T-t)}{8}}}{v_t\sqrt{2\pi(T-t)}},$$

we can write

$$\begin{aligned}\frac{\partial I_t}{\partial X_t}(x_t^*) &= -\frac{\rho}{2}e^{\frac{I_t(x_t^*)^2(T-t)}{8}}(T-t)^{-2}E\left(e^{-\frac{v_t^2(T-t)}{8}}v_t^{-3}\int_t^T\Lambda_s\,ds\,\middle|\,\mathcal{F}_t\right)\\ &\quad-\lambda ke^{\frac{I_t(x_t^*)^2(T-t)}{8}}E\big(e^{\frac{-v_t{}^2(T-t)}{8}}v_t^{-1}\mid\mathcal{F}_t\big)+O(T-t)^{(\frac{1}{2}+2\delta)\wedge\frac{1}{2}}\\ &=: S_1+S_2+O(T-t)^{(\frac{1}{2}+2\delta)\wedge 1}.\end{aligned}$$

By Lemma 6.1 we know that $I_t(x_t^*)^2(T-t)\to 0$, as $T\to t$. Then

$$\lim_{T\to t}S_1=-\frac{\rho}{2}\lim_{T\to t}\left[(T-t)^{-2}E\left(e^{-\frac{v_t^2(T-t)}{8}}v_t^{-3}\int_t^T\Lambda_s\,ds\,\middle|\,\mathcal{F}_t\right)\right].$$

Using again Lemma 6.1, observe that (H4) and the dominated convergence theorem imply that

$$\lim_{T\to t}S_2=-\frac{\lambda k}{\sigma_t}. \tag{6.10}$$

Now the proof will be decomposed into two steps.

Step 1. Here we analyze the case $\delta\geq 0$. In this case we only need to show that

$$\lim_{T\to t}\left(S_1+\frac{\rho}{2\sigma_t}D_t^+\sigma_t\right)=0. \tag{6.11}$$

Indeed, we can write

$$\lim_{T\to t}\left(S_1+\frac{\rho}{2\sigma_t}D_t^+\sigma_t\right)=\lim_{T\to t}E\left(A_TB_T+\frac{\rho}{2\sigma_t}D_t^+\sigma_t\,\middle|\,\mathcal{F}_t\right),$$

where

$$A_T:=\frac{\rho}{v_t}\exp\left(-\frac{(v_t^2)(T-t)}{8}\right)$$

and

$$B_T:=-\frac{1}{v_t^2(T-t)^2}\int_t^T\int_s^T\sigma_r\sigma_sD_s\sigma_r\,dr\,ds.$$

Notice that

$$\begin{aligned}&\lim_{T\to t}E\left(A_TB_T+\frac{\rho}{2\sigma_t}D_t^+\sigma_t\,\middle|\,\mathcal{F}_t\right)\\ &\quad=\lim_{T\to t}E\left(\left(A_T-\frac{\rho}{\sigma_t}\right)B_T\,\middle|\,\mathcal{F}_t\right)+\frac{\rho}{\sigma_t}\lim_{T\to t}E\left(\left(B_T+\frac{D_t^+\sigma_t}{2}\right)\,\middle|\,\mathcal{F}_t\right)\\ &\quad=\lim_{T\to t}U_1+\frac{\rho}{\sigma_t}\lim_{T\to t}U_2.\end{aligned}$$

Applying the Cauchy–Schwarz inequality yields that

$$U_1 \le \left[E\left(\left(A_T-\frac{\rho}{\sigma_t}\right)^2 \,\middle|\, \mathcal{F}_t\right)\right]^{\frac{1}{2}}\left[E\left(B_T^2 \mid \mathcal{F}_t\right)\right]^{\frac{1}{2}}.$$

Using the dominated convergence theorem it is easy to see that $E((A_T-\frac{\rho}{\sigma_t})^2|\mathcal{F}_t)$ tends to zero, as $T\to t$, and a simple calculation gives us that $E(B_T^2|\mathcal{F}_t)$ is bounded; whence, we deduce that $\lim_{T\to t} U_1 = 0$. On the other hand,

$$\begin{aligned}
|U_2| &= \left|\frac{1}{(T-t)^2}E\left(\int_t^T\int_s^T\left(\frac{\sigma_s\sigma_r}{v_t^2}D_s\sigma_r - D_t^+\sigma_t\right)dr\,ds\,\middle|\,\mathcal{F}_t\right)\right| \\
&\le \frac{C}{(T-t)^2}\left|E\left(\int_t^T\int_s^T\left(\frac{\sigma_s\sigma_r}{v_t^2}-1\right)D_s\sigma_r\,dr\,ds\,\middle|\,\mathcal{F}_t\right)\right| \\
&\quad + \frac{C}{(T-t)^2}\left|E\left(\int_t^T\int_s^T\left(D_s\sigma_r - D_t^+\sigma_t\right)dr\,ds\,\middle|\,\mathcal{F}_t\right)\right| \\
&=: |U_{2,1}| + |U_{2,2}|.
\end{aligned}$$

Using now the Cauchy–Schwarz inequality and the fact that hypothesis (H3) holds with $\delta \ge 0$, we obtain that

$$\begin{aligned}
|U_{2,1}| &\le \frac{C}{(T-t)^2}\left(E\left(\int_t^T\int_s^T\left(\frac{\sigma_s\sigma_r}{v_t^2}-1\right)^2 dr\,ds\,\middle|\,\mathcal{F}_t\right)\right)^{\frac{1}{2}} \\
&\quad \times\left(E\left(\int_t^T\int_s^T (D_s\sigma_r)^2\,dr\,ds\,\middle|\,\mathcal{F}_t\right)\right)^{\frac{1}{2}} \\
&\le \frac{C}{(T-t)}\left(E\left(\int_t^T\int_s^T\left(\frac{\sigma_s\sigma_r}{v_t^2}-1\right)^2 dr\,ds\,\middle|\,\mathcal{F}_t\right)\right)^{\frac{1}{2}}.
\end{aligned}$$

Now (H2) and (H4) allow us to write

$$\begin{aligned}
|U_{2,1}| &\le \frac{C}{(T-t)}\left(\int_t^T\int_s^T E\left(\left(\sigma_s\sigma_r - v_t^2\right)^2|\mathcal{F}_t\right)dr\,ds\right)^{\frac{1}{2}} \\
&= \frac{C}{(T-t)}\left(\int_t^T\int_s^T E\left(\left(\sigma_s\sigma_r-\left(\frac{1}{T-t}\int_t^T\sigma_\theta^2\,d\theta\right)\right)^2\,\middle|\,\mathcal{F}_t\right)dr\,ds\right)^{\frac{1}{2}} \\
&\le \frac{C}{(T-t)^{\frac{3}{2}}}\left(\int_t^T\int_s^T\int_t^T E\left(\left(\sigma_s\sigma_r-\sigma_\theta^2\right)^2|\mathcal{F}_t\right)d\theta\,dr\,ds\right)^{\frac{1}{2}},
\end{aligned}$$

which tends to zero by hypothesis (H5). Similarly,

$$|U_{2,2}| \le \frac{C}{(T-t)^2}\left|\int_t^T\int_s^T E\left(\left(D_s\sigma_r - D_t^+\sigma_t\right)|\mathcal{F}_t\right)dr\,ds\right|,$$

which tends to zero by (6.6). Now we have proved (6.11). Then, (6.10), (6.11), and the fact that $\delta \geq 0$ give (6.7).

Step 2. Finally, we show that (6.9) is true. Let us prove that

$$\lim_{T\to t}\left(S_1(T-t)^{-\delta}+\frac{\rho}{\sigma_t}L_t^{\delta,+}\sigma_t\right)=0. \tag{6.12}$$

Note that

$$\lim_{T\to t}\left(S_1(T-t)^{-\delta}+\frac{\rho}{\sigma_t}L_t^{\delta,+}\sigma_t\right)=\lim_{T\to t}E\left(A_T\tilde{B}_T+\frac{\rho}{\sigma_t}L_t^{\delta,+}\sigma_t\,\middle|\,\mathcal{F}_t\right),$$

where A_T is defined as in Step 1 and

$$\tilde{B}_T:=-\frac{1}{v_t^2(T-t)^{2+\delta}}\int_t^T\int_s^T\sigma_r\sigma_s D_s\sigma_r\,dr\,ds.$$

But

$$\begin{aligned}
&\lim_{T\to t}E\left(A_T\tilde{B}_T+\frac{\rho}{\sigma_t}L_t^{\delta,+}\sigma_t\,\middle|\,\mathcal{F}_t\right)\\
&\quad=\lim_{T\to t}E\left(\left(A_T-\frac{\rho}{\sigma_t}\right)\tilde{B}_T\,\middle|\,\mathcal{F}_t\right)+\frac{\rho}{\sigma_t}\lim_{T\to t}E\big(\big(\tilde{B}_T+L_t^{\delta,+}\sigma_t\big)\big|\mathcal{F}_t\big).
\end{aligned}$$

Then, using similar arguments as in the proof of Step 1, we can easily see that this expression is equal to zero. Now we have proved (6.12). Finally, using (6.12), (6.10), and the fact that $-\frac{1}{2}<\delta<0$ the result follows. □

Remark Notice that (6.7) and (6.9) can be written in terms of $\frac{\partial I_t}{\partial Z}$, where $Z=\log K$ is the log-strike, by simply changing the sign of the limits.

7 Examples

7.1 Diffusion stochastic volatilities

Assume that the volatility σ can be written as $\sigma=f(Y)$, where $f\in\mathcal{C}_b^1(\mathbb{R})$ and Y is the solution of a stochastic differential equation

$$dY_r=a(r,Y_r)\,dr+b(r,Y_r)\,dW_r, \tag{7.1}$$

for some real functions $a,b\in\mathcal{C}_b^1(\mathbb{R})$. Then, classical arguments (see, for example, Theorem 2.2.1 in [20]) give us that $Y\in\mathbb{L}^{1,2}$ and that

$$D_sY_r=\int_s^r\frac{\partial a}{\partial x}(u,Y_u)D_sY_u\,du+b(s,Y_s)+\int_s^r\frac{\partial b}{\partial x}(u,Y_u)D_sY_u\,dW_u. \tag{7.2}$$

Taking now into account that $D_s\sigma_r = f'(Y_r)D_sY_r$, it can be easily deduced from (7.2) that (H3) holds with $\delta = 0$ and that

$$\sup_{s,r\in[t,T]} \left|E\left(\left(D_s\sigma_r - f'(Y_t)b(t,Y_t)\right)\middle|\mathcal{F}_t\right)\right| \to 0,$$

as $T \to t$. Then Theorem 6.3 gives us that

$$\lim_{T\to t} \frac{\partial I_t}{\partial X_t}(x_t^*) = -\frac{1}{\sigma_t}\left(\lambda k + \frac{\rho}{2} f'(Y_t)b(t,Y_t)\right),$$

which agrees with the results in [19].

In particular, if Y is an Ornstein–Uhlenbeck process of the form

$$Y_r = m + (Y_t - m)e^{-\alpha(r-t)} + c\int_t^r \sqrt{2\alpha}\exp\left(-\alpha(r-s)\right) dW_s, \tag{7.3}$$

$D_sY_r = c\sqrt{2\alpha}\exp(-\alpha(r-s))$ for all $t \le s < r$ and then it follows that

$$\lim_{T\to t} \frac{\partial I_t}{\partial X_t}(x_t^*) = -\frac{1}{\sigma_t}\left(\lambda k + c\sqrt{2\alpha}\frac{\rho}{2} f'(Y_t)\right).$$

7.2 Fractional stochastic volatility models

Assume that the volatility σ can be written as $\sigma = f(Y)$, where $f \in \mathcal{C}_b^1(\mathbb{R})$ and Y is a process of the form

$$Y_r = m + (Y_t - m)e^{-\alpha(r-t)} + c\sqrt{2\alpha}\int_t^r e^{-\alpha(r-s)}\, dW_s^H, \tag{7.4}$$

where $W_s^H := \int_0^s (s-u)^{H-\frac{1}{2}}\, dW_u$.

7.2.1 Case $H > \frac{1}{2}$

As in [10], assume the volatility model (7.4), for some $H > 1/2$. Notice that (see, for example, [3]) $\int_t^r e^{-\alpha(r-s)}\, dW_s^H$ can be written as

$$\left(H - \frac{1}{2}\right)\int_0^r \left(\int_s^r \mathbb{1}_{[t,r]}(u)e^{-\alpha(r-u)}(u-s)^{H-\frac{3}{2}}\, du\right) dW_s,$$

from which it follows easily that $\sup_{s,r\in[t,T]} |E(D_s\sigma_r|\mathcal{F}_t)| \to 0$, as $T \to t$. Then Theorem 6.3 gives us that $\lim_{T\to t} \frac{\partial I_t}{\partial X_t}(x_t^*) = -\frac{\lambda k}{\sigma_t}$. That is, the at-the-money short-dated skew slope of the implied volatility is not affected by the correlation in this case.

7.2.2 Case $H < \frac{1}{2}$

Assume again the model (7.4), taking $0 < H < 1/2$. It can be proved (see, for example, [3]) that $\int_t^r e^{-\alpha(r-s)}\, dW_s^H$ can be expressed as

$$\left(\frac{1}{2} - H\right) \int_0^r \left(\int_s^r \left[\mathbb{1}_{[t,r]}(u) e^{-\alpha(r-u)} - \mathbb{1}_{[t,r]}(s) e^{-\alpha(r-s)} \right] (u-s)^{H-\frac{3}{2}}\, du \right) dW_s$$
$$+ \int_t^r e^{-\alpha(r-s)} (r-s)^{H-\frac{1}{2}}\, dW_s.$$

Then it follows that (H3) holds for $\delta = H - \frac{1}{2}$ and we can easily check that

$$E\left(\frac{1}{(T-t)^{2+H-\frac{1}{2}}} \int_t^T \int_s^T D_s^W \sigma_r\, dr\, ds - c\sqrt{2\alpha} f'(Y_t) \,\middle|\, \mathcal{F}_t \right) \to 0, \quad \text{as } T \to t.$$

Then Theorem 6.3 gives us that

$$\lim_{T \to t} (T-t)^{\frac{1}{2}-H} \frac{\partial I_t}{\partial X_t}(x_t^*) = -c\sqrt{2\alpha} \frac{\rho}{\sigma_t} f'(Y_t).$$

That is, the introduction of fractional components with Hurst parameter $H < 1/2$ in the definition of the volatility process allows us to reproduce a skew slope of order $O(T-t)^{\delta}$, for every $\delta > -1/2$.

7.3 Time-varying coefficients

Fouque et al. [13] have introduced a new approach to capture the maturity-dependent behavior of the implied volatility, by allowing the volatility coefficients to depend on the time till the next maturity date. Namely, they assume that the volatility σ can be written as $\sigma = f(Y)$, where f is a regular enough function and Y is a diffusion process of the form (7.3), with $\sqrt{\alpha(s)}$ a suitable cutoff of the function $(T_{n(s)} - s)^{-\frac{1}{2}}$, with fixed maturity dates $\{T_k\}$ (the third Friday of each month) and $n(t) = \inf\{n : T_n > s\}$.

Following this idea, we can consider Y to be a diffusion process of the form (7.3), with $\sqrt{\alpha(s)} = (T_{n(s)} - s)^{-\frac{1}{2}+\varepsilon}$, for some $\varepsilon > 0$. It is now easy to see that $Y \in \mathbb{L}^{1,2}$ and that

$$\frac{1}{(T-t)^{2+(\frac{1}{2}-\varepsilon)}} \int_t^T \int_s^T E(D_s\sigma_r | \mathcal{F}_t)\, dr\, ds + \rho c \left(\frac{1}{-1/2+\varepsilon} \right) \left(\frac{1}{1/2+\varepsilon} \right) \frac{f'(Y_t)}{2}$$

tends to zero, as $T - t$ tends to zero. Hence, we deduce that, in this case, the short-date skew slope of the implied volatility is of the order $O(T-t)^{-\frac{1}{2}+\varepsilon}$.

8 Conclusions

We have seen that Malliavin calculus may provide a natural approach to deal with the short-date behavior of the implied volatility for jump-diffusion models with stochastic volatility. This theory does not require the volatility to be a diffusion or a

Markov process. Moreover, with these techniques the short-time behavior of the implied volatility can be analyzed for known and new volatility models—in particular, models that reproduce short-date skews of order $O(T-t)^{\delta}$, for $\delta > -\frac{1}{2}$.

References

1. Alòs, E.: A generalization of the Hull and White formula with applications to option pricing approximation. Finance Stoch. **10**, 353–365 (2006)
2. Alòs, E., Nualart, D.: An extension of Itô's formula for anticipating processes. J. Theor. Probab. **11**, 493–514 (1998)
3. Alòs, E., Mazet, O., Nualart, D.: Stochastic calculus with respect to Gaussian processes. Ann. Probab. **29**, 766–801 (2001)
4. Bakshi, G., Cao, C., Chen, Z.: Empirical performance of alternative option pricing models. J. Finance **52**, 2003–2049 (1997)
5. Ball, C., Roma, A.: Stochastic volatility option pricing. J. Finance Quant. Anal. **29**, 589–607 (1994)
6. Barndorff-Nielsen, O.E., Shephard, N.: Modelling by Lévy processes for financial econometrics. In: Barndorff-Nielsen, O.E., Mikosch, T., Resnick, S.I. (eds.) Lévy Processes: Theory and Applications, pp. 283–318. Birkhäuser, Basel (2001)
7. Barndorff-Nielsen, O.E., Shephard, N.: Econometric analysis of realized volatility and its use in estimating stochastic volatility models. J. Roy. Stat. Soc. Ser. B Stat. Methodol. **64**, 253–280 (2002)
8. Bates, D.S.: Jumps and stochastic volatility: exchange rate processes implicit in Deutsche Mark options. Rev. Finance Stud. **9**, 69–107 (1996)
9. Carr, P., Wu, L.: The finite moment log stable process and option pricing. J. Finance **58**, 753–778 (2003)
10. Comte, F., Renault, E.: Long memory in continuous-time stochastic volatility models. Math. Finance **8**, 291–323 (1998)
11. Duffie, D., Pan, J., Singleton, K.: Transform analysis and asset pricing for affine jump-diffusions. Econometrica **68**, 1343–1376 (2000)
12. Fouque, J.-P., Papanicolaou, G., Sircar, R., Solna, K.: Singular perturbations in option pricing. SIAM J. Appl. Math. **63**, 1648–1665 (2003)
13. Fouque, J.-P., Papanicolaou, G., Sircar, R., Solna, K.: Maturity cycles in implied volatility. Finance Stoch. **8**, 451–477 (2004)
14. Heston, S.L.: A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev. Finance Stud. **6**, 327–343 (1993)
15. Hull, J., White, A.: The pricing of options on assets with stochastic volatilities. J. Finance **42**, 281–300 (1987)
16. Jacod, J., Protter, P.: Risk neutral compatibility with option prices. Preprint (2006). http://legacy.orie.cornell.edu/~protter/finance.html
17. Lee, R.W.: Implied volatility: statics, dynamics, and probabilistic interpretation. In: Baeza-Yates, R., Glaz, J., Gzyl, H., et al. (eds.) Recent Advances in Applied Probability, pp. 241–268. Springer, Berlin (2004)
18. Lewis, A.L.: Option Valuation Under Stochastic Volatility with Mathematica Code. Finance Press, Newport Beach (2000)
19. Medvedev, A., Scaillet, O.: Approximation and calibration of short-term implied volatilities under jump-diffusion stochastic volatility. Rev. Finance Stud. **20**(2), 427–459 (2007)
20. Nualart, D.: The Malliavin Calculus and Related Topics. Springer, Berlin (1995)
21. Renault, E., Touzi, N.: Option hedging and implied volatilities in a stochastic volatility model. Math. Finance **6**, 279–302 (1996)
22. Schweizer, M., Wissel, J.: Term structures of implied volatilities: absence of arbitrage and existence results. Math. Finance (2006, to appear)
23. Scott, L.O.: Option pricing when the variance changes randomly: theory, estimation and an application. J. Finance Quant. Anal. **22**, 419–438 (1987)
24. Stein, E.M., Stein, J.C.: Stock price distributions with stochastic volatility: an analytic approach. Rev. Finance Stud. **4**, 727–752 (1991)