

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

Elisa Alòs*
Dpt. d'Economia i Empresa
Universitat Pompeu Fabra
c/Ramon Trias Fargas, 25-27
08005 Barcelona, Spain

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

Josep Vives‡
Dpt. de Matemàtiques, Universitat Autònoma de Barcelona
08193 Bellaterra (Barcelona), Spain
Dpt. Probabilitat, Lògica i Estadística, Universitat de Barcelona
Gran Via, 585, 08007 Barcelona, Spain

June 30, 2006

Abstract

In this paper we use the 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 neither a diffusion, nor a Markov process, as the examples in section 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 code: G12, G13

Mathematical Subject Classification: 91B28, 91B70, 60H07.

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

*Supported by grants CICYT BFM2003-04294 and MCYT SEC2003-04476

†Partially supported by the CONACyT grant 45684-F

‡Supported by grant MECD-BFM2003-00261

these extensions, one of the most popular is to allow the volatility to be a stochastic process (see for example Hull and White (1987), Scott (1987), Stein and Stein (1991), Heston (1993) or Ball and Roma (1994), 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 Renault and Touzi (1996)). But the observed implied volatility exhibits dependence not only on the strike price, but also on 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 Lewis (2000), Lee (2004) or Medvedev and Scaillet (2003)). 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 it is shown by Medvedev and Scaillet (2003)), they allow flexible modelling, and generate skews and smiles similar to those observed in market data (see Bates (1996), Barndorff-Nielsen and Shephard (2001a, 2001b) or Carr and Wu (2003)). Recently, Fouque, Papanicolau, Sircar and Solna (2004) 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 is large enough near the maturity time.

The main goal of this paper is to provide a method based on the techniques of the 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. 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 either fractional processes with parameter in $(0, 1)$, Markov processes, or processes with time-varying coefficients, among others.

The paper is organized as follows. In Section 2 we introduce the framework and the notation that we utilize in this paper. In Section 3 we state our basic tool. Namely, an anticipating Itô's formula for the Skorohod integral. As a consequence, in Section 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 Section 5. Section 6 is devoted to the main result of this article. That means, we figure out the short-time limit behavior. Finally, in Section 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 will consider the following model for the log-price of a stock under a risk-neutral probability measure Q :

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

Here, x is the current log-price, r is the instantaneous interest rate, W and B are independent standard Brownian motions, $\rho \in (-1, 1)$, Z is a compound Poisson process with intensity λ and Lévy measure ν , $k = \frac{1}{\lambda} \int_{\mathbb{R}} (e^y - 1) \nu(dy) < \infty$, and σ is a second-order stochastic process adapted to the filtration generated by W . Notice that this model is a generalization of the classical Bates model introduced in Bates (1996), in the sense that we do not assume the volatility to be a diffusion process.

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 an European call with strike price K by the formula

$$V_t = e^{-r(T-t)} E[(e^{X_T} - K)_+ | \mathcal{F}_t], \quad (2)$$

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 will follow here is the same as in Fouque, Papanicolau, Sircar and Solna (2003), where it is assumed that the market selects a unique equivalent martingale measure under which derivative contracts are priced.

In the sequel, we will make use the following notation:

- $v_t := \left(\frac{Y_t}{T-t} \right)^{\frac{1}{2}}$, 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 will write $p(x)$.
- $BS(t, x, \sigma)$ will denote the price of an 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:

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

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 $x_t^* := \ln K - r(T-t)$.

- $\mathcal{L}_{BS}(\sigma)$ will denote the Black-Scholes differential operator, in the log variable, with volatility σ :

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

It is well known that $\mathcal{L}_{BS}(\sigma)BS(\cdot, \cdot, \sigma) = 0$.

- $G(t, x, \sigma) := (\partial_{xx}^2 - \partial_x)BS(t, x, \sigma)$.

3 Preliminaries on Malliavin Calculus

Let us consider a standard Brownian motion $W = \{W_t, t \in [0, T]\}$ defined in a complete probability space (Ω, \mathcal{F}, P) . Set $H = L^2([0, T])$, and denote by $W(h) = \int_0^T h(s)dW_s$ the Wiener integral of a deterministic function $h \in H$.

In this section we introduce the basic notations and results of the Malliavin calculus following closely to Nualart (1995).

Let \mathcal{S} be the set of smooth and cylindrical random variables of the form $F = f(W(h_1), \dots, W(h_n))$, where $n \geq 1$, $f \in C_b^\infty(\mathbb{R}^n)$ (f and all its derivatives are bounded), and $h_1, \dots, h_n \in H$. Given a random variable F of this form, we define its derivative as the stochastic process $\{D_t^W F, t \in [0, T]\}$ given by

$$D_t^W F = \sum_{i=1}^n (\partial_i f)(W(h_1), \dots, W(h_n)) h_i(t), \quad t \in [0, T].$$

The operator D^W and the iterated operators $D^{W,n}$ are closable and unbounded from $L^2(\Omega)$ into $L^2([0, T]^n \times \Omega)$, for all $n \geq 1$. We denote by $\mathbb{D}_W^{n,2}$ the closure of \mathcal{S} with respect to the norm defined by

$$\|F\|_{n,2}^2 = \|F\|_{L^2(\Omega)}^2 + \sum_{k=1}^n \|D^{W,k} F\|_{L^2([0,T]^k \times \Omega)}^2.$$

We denote by δ^W the adjoint of the derivative operator D^W that 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 $\text{Dom} \delta^W$ and the operator δ^W restricted to $L_a^2([0, T] \times \Omega)$ coincides with the Itô stochastic integral. We will make use of 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 prove the following ad-hoc version of Itô's formula, which follows from Alòs and Nualart (1998) and Alòs (2006). In the sequel, we use the notation $D = D^W$ to simplify the exposition.

Proposition 1 Assume model (1) and $\sigma^2 \in \mathbb{L}^{1,2}$. Let $F : [0, T] \times \mathbb{R}^2 \rightarrow \mathbb{R}$ 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

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

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 its continuous part X^c given by the equation:

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

Then, applying Theorem 1 in Alòs (2006) we have that

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

since $dX_t = dX_t^c$ in $[T_i, T_{i+1})$. 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:

$$\begin{aligned}
F(t, X_t, Y_t) &= F(0, X_0, Y_0) + \int_0^t \partial_s F(s, X_s, Y_s) ds \\
&+ \int_0^t \partial_x F(s, X_s, Y_s) dX_s^c - \int_0^t \partial_y F(s, X_s, Y_s) \sigma_s^2 ds \\
&+ \rho \int_0^t \partial_{xy}^2 F(s, X_s, Y_s) \Lambda_s ds + \frac{1}{2} \int_0^t \partial_{xx}^2 F(s, X_s, Y_s) \sigma_s^2 ds \\
&+ \sum_{0 \leq s \leq t} [F(s, X_{s-} + \Delta X_s, Y_s) - F(s, X_{s-}, Y_s)].
\end{aligned}$$

Hence we deduce the desired result. ■

4 An extension of Hull and White formula

In this section, using the Itô's formula and the arguments developed in Alòs (2006), we prove an extension of the Hull and White formula that gives the price of an European call option as a sum of the price when the model has no jumps and no correlation plus three terms: one describes the impact of the correlation on option prices and two of them, which can be presented jointly, describe the impact of jumps in this prices. Hence this formula will be a useful tool to compare the effect of correlation and jumps (see Section 5).

We will need the following result, inspired in Lemma 5 in Fouque, Papanicolaou, Sircar and Solna (2003).

Lemma 2 *Let $0 \leq t \leq s \leq T$, $\rho \in (-1, 1)$ and $\mathcal{G}_t := \mathcal{F}_t \vee \mathcal{F}_T^W \vee \mathcal{F}_T^Z$. Then for every $n \geq 0$, there exists $C = C(n, \rho)$ such that*

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

Proof: A simple calculation gives us that

$$\partial_x BS(s, x, \sigma) = e^x N(d_+) + e^x \frac{N'(d_+)}{\sigma \sqrt{T-t}} - K e^{-r(T-t)} \frac{N'(d_-)}{\sigma \sqrt{T-t}}.$$

Since $e^x N'(d_+) = K e^{-r(T-t)} N'(d_-)$ it follows that

$$\partial_x BS(s, x, \sigma) = e^x N(d_+)$$

and

$$\partial_{xx}^2 BS(s, x, \sigma) = e^x N(d_+) + K e^{-r(T-t)} \frac{N'(d_-)}{\sigma \sqrt{T-t}}.$$

In consequence

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

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

$$E(\partial_x^n G(s, X_s, v_s) | \mathcal{G}_t) = (-1)^n K e^{-r(T-s)} \partial_\mu^n E(p(X_s - \mu, v_s \sqrt{T-s}) | \mathcal{G}_t). \quad (3)$$

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

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

and variance equal to $(1 - \rho^2) \int_t^s \sigma_\theta^2 d\theta$, it follows that

$$\begin{aligned} & E\left(p\left(X_s - \mu, v_s \sqrt{T-s}\right) \middle| \mathcal{G}_t\right) \\ &= \int_{\mathbb{R}} p\left(y - \mu, v_s \sqrt{T-s}\right) p\left(y - \phi, \sqrt{(1 - \rho^2) \int_t^s \sigma_\theta^2 d\theta}\right) dy \\ &= p\left(\phi - \mu, \sqrt{\int_s^T \sigma_s^2 ds + (1 - \rho^2) \int_t^s \sigma_\theta^2 d\theta}\right) \\ &= p\left(\phi - \mu, \sqrt{(1 - \rho^2) \int_t^T \sigma_\theta^2 d\theta + \rho^2 \int_s^T \sigma_\theta^2 d\theta}\right). \end{aligned}$$

Putting this result in (3), we have

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

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

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

and thus the proof is complete. \blacksquare

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

Theorem 3 Assume the model (1) holds with $\sigma \in \mathbb{L}^{1,2}$. Then it follows that

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

Proof: This proof is similar to the one of the main theorem in Alòs (2006). Notice that $BS(T, X_T, v_T) = V_T$. Then, from (2) we have

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

Now, our idea is to apply Proposition 1 to the process $e^{-rt} BS(t, X_t, v_t)$. As the derivatives of $BS(t, x, \sigma)$ are not bounded we will make use of an approximating argument, changing v_t by

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

and $BS(t, x, \sigma)$ by $BS_n(t, x, \sigma) := BS(t, x, \sigma) \psi_n(x)$, where $\psi_n(X_t) := \phi\left(\frac{1}{n}x\right)$, for some $\phi \in \mathcal{C}_b^2$ such that $\phi(x) = 1$ for all $x < 1$ and $\phi(x) = 0$ for all $x > 2$. Now, applying Proposition 1 between t and T and grouping terms according with the type of derivative we obtain:

$$\begin{aligned} &e^{-rT} BS_n(T, X_T, v_T^\delta) \\ &= e^{-rt} BS_n(t, X_t, v_t^\delta) + \int_t^T e^{-rs} \mathcal{L}_{BS}(\sigma_s) BS_n(s, X_s, v_s^\delta) ds \\ &\quad - \frac{1}{2} \int_t^T e^{-rs} \partial_\sigma BS_n(s, X_s, v_s^\delta) \frac{(\sigma_s^2 - (v_s^\delta)^2)}{v_s^\delta(T-s)} ds \\ &\quad - \lambda k \int_t^T e^{-rs} \partial_x BS_n(s, X_s, v_s^\delta) ds \\ &\quad + \int_t^T e^{-rs} \partial_x BS_n(s, X_s, v_s^\delta) \sigma_s (\rho dW_s + \sqrt{1-\rho^2} dB_s) \\ &\quad + \frac{\rho}{2} \int_t^T e^{-rs} \partial_{\sigma_x}^2 BS_n(s, X_s, v_s^\delta) \frac{1}{v_s^\delta(T-s)} \Lambda_s ds \\ &\quad + \int_t^T \int_{\mathbb{R}} e^{-rs} (BS_n(s, X_{s-} + y, v_s^\delta) - BS_n(s, X_{s-}, v_s^\delta)) \tilde{J}_X(ds, dy) \\ &\quad + \int_t^T \int_{\mathbb{R}} e^{-rs} (BS_n(s, X_{s-} + y, v_s^\delta) - BS_n(s, X_{s-}, v_s^\delta)) \nu(dy) ds. \end{aligned}$$

Notice that $\mathcal{L}_{BS}(\sigma_s) BS_n(s, X_s, v_s^\delta) = (\mathcal{L}_{BS}(\sigma_s) BS(s, X_s, v_s^\delta)) \psi_n(X_s) + A_n(s)$,

where

$$A_n(s) = \frac{1}{2}\sigma_s^2 [2\partial_x BS(s, X_s, v_s^\delta)\psi'_n(X_s) + BS(s, X_s, v_s^\delta) (\psi''_n(X_s) - \psi'_n(X_s))] + rBS(s, X_s, v_s^\delta)\psi'_n(X_s).$$

Also note that the classical relation between the *Gamma*, the *Vega* and the *Delta* gives us that

$$\partial_\sigma BS(s, x, \sigma) \frac{1}{\sigma(T-s)} = G(s, x, \sigma).$$

Then we can write

$$\begin{aligned} & e^{-rT} BS_n(T, X_T, v_T^\delta) \\ = & e^{-rt} BS_n(t, X_t, v_t^\delta) + \int_t^T e^{-rs} [(\mathcal{L}_{BS}(\sigma_s)BS)(s, X_s, v_s^\delta)\psi_n(X_s) + A_n(s)] ds \\ & - \frac{1}{2} \int_t^T e^{-rs} G(s, X_s, v_s^\delta)\psi_n(X_s)(\sigma_s^2 - (v_s^\delta)^2) ds \\ & - \lambda k \int_t^T e^{-rs} \partial_x BS_n(s, X_s, v_s^\delta) ds \\ & + \int_t^T e^{-rs} \partial_x BS_n(s, X_s, v_s^\delta) \sigma_s (\rho dW_s + \sqrt{1-\rho^2} dB_s) \\ & + \frac{\rho}{2} \int_t^T e^{-rs} [(\partial_x G)(s, X_s, v_s^\delta)\psi_n(X_s) + G(s, X_s, v_s^\delta)\psi'_n(X_s)] \Lambda_s ds \\ & + \int_t^T \int_{\mathbb{R}} e^{-rs} (BS_n(s, X_{s-} + y, v_s^\delta) - BS_n(s, X_{s-}, v_s^\delta)) \tilde{J}_X(ds, dy) \\ & + \int_t^T \int_{\mathbb{R}} e^{-rs} (BS_n(s, X_{s-} + y, v_s^\delta) - BS_n(s, X_{s-}, v_s^\delta)) \nu(dy) ds \end{aligned}$$

Hence, taking into account that $\mathcal{L}_{BS}(\sigma_s) = \mathcal{L}_{BS}(v_s^\delta) + \frac{1}{2}(\sigma_s^2 - (v_s^\delta)^2)(\partial_{xx}^2 - \partial_x)$ it follows that (using the fact that $\mathcal{L}_{BS}(v_s^\delta)BS(s, X_s, v_s^\delta) = 0$)

$$\begin{aligned} & e^{-rT} BS_n(T, X_T, v_T^\delta) = e^{-rt} BS_n(t, X_t, v_t^\delta) + \int_t^T e^{-rs} A_n(s) ds \\ & - \lambda k \int_t^T e^{-rs} \partial_x BS_n(s, X_s, v_s^\delta) ds \\ & + \int_t^T e^{-rs} \partial_x BS_n(s, X_s, v_s^\delta) \sigma_s (\rho dW_s + \sqrt{1-\rho^2} dB_s) \\ & + \frac{\rho}{2} \int_t^T e^{-rs} [(\partial_x G)(s, X_s, v_s^\delta)\psi_n(X_s) + G(s, X_s, v_s^\delta)\psi'_n(X_s)] \Lambda_s ds \\ & + \int_t^T \int_{\mathbb{R}} e^{-rs} (BS_n(s, X_{s-} + y, v_s^\delta) - BS_n(s, X_{s-}, v_s^\delta)) \tilde{J}_X(ds, dy) \\ & + \int_t^T \int_{\mathbb{R}} e^{-rs} (BS_n(s, X_{s-} + y, v_s^\delta) - BS_n(s, X_{s-}, v_s^\delta)) \nu(dy) ds \end{aligned}$$

Now, taking conditional expectations we obtain that

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

Finally we have that $|BS(t, x, \sigma)| + |\partial_x BS(t, x, \sigma)| \leq 2e^x + K$ and that the processes $\psi_n'(X_s)$ and $\psi_n''(X_s)$ are bounded and tend to zero a.s. Moreover,

Lemma 2 gives us that $E(G(s, X_s, v_s^\delta) | \mathcal{G}_t) \leq C \left(\int_t^T \sigma_s^2 ds \right)^{-\frac{1}{2}}$ and that $E(\partial_x G(s, X_s, v_s^\delta) | \mathcal{G}_t) \leq C \left(\int_t^T \sigma_s^2 ds \right)^{-1}$, for some positive constant C . Also, by Hölder's inequality we have

$$\begin{aligned}
\left| \int_t^T \Lambda_s ds \right| & \leq 2 \int_t^T \sigma_s \int_s^T \sigma_\theta |D_s \sigma_\theta| d\theta ds \\
& \leq 2 \left(\int_t^T \sigma_\theta^2 d\theta \right) \left(\int_t^T \int_t^T (D_r \sigma_\theta)^2 dr d\theta \right)^{\frac{1}{2}}. \quad (4)
\end{aligned}$$

Then, letting first $n \uparrow \infty$ and then $\delta \downarrow 0$, and using the dominated convergence theorem, the proof is complete. ■

5 An expression for the derivative of the implied volatility

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

Proposition 4 *Assume the model (1) holds with $\sigma \in \mathbb{L}^{1,2}$ and for every fixed $t \in [0, T]$, $E \left(\int_t^T \sigma_s^2 ds \middle| \mathcal{F}_t \right)^{-1} < \infty$ a.s. Then it follows that*

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

where

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

Proof: Taking partial derivatives with respect to X_t on the expression $V_t = BS(t, X_t, I_t(X_t))$ we obtain

$$\frac{\partial V_t}{\partial X_t} = \partial_x BS(t, X_t, I_t(X_t)) + \partial_\sigma BS(t, X_t, I_t(X_t)) \frac{\partial I_t}{\partial X_t}(X_t). \quad (5)$$

On the other hand, from Theorem 3 we deduce that

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

which implies that

$$\frac{\partial V_t}{\partial X_t} = E(\partial_x BS(t, X_t, v_t) | \mathcal{F}_t) + E\left(\int_t^T \partial_x F(s, X_s, v_s) ds | \mathcal{F}_t\right). \quad (6)$$

Using now the fact that $E\left(\int_t^T \sigma_s^2 ds | \mathcal{F}_t\right)^{-1} < \infty$ we can check that the conditional expectation $E\left(\int_t^T \partial_x F(s, X_s, v_s) ds | \mathcal{F}_t\right)$ is well defined and finite a.s. Thus, (5) and (6) imply

$$\begin{aligned}
& \frac{\partial I_t}{\partial X_t}(x_t^*) \\
= & \frac{E(\partial_x BS(t, x_t^*, v_t) | \mathcal{F}_t) - \partial_x BS(t, x_t^*, I_t(x_t^*)) + E\left(\int_t^T \partial_x F(s, X_s, v_s) ds | \mathcal{F}_t\right) |_{X_t=x_t^*}}{\partial_\sigma BS(t, x_t^*, I_t(x_t^*))}.
\end{aligned} \quad (7)$$

Notice that

$$E(\partial_x BS(t, x_t^*, v_t) | \mathcal{F}_t) = \partial_x E(BS(t, x, v_t) | \mathcal{F}_t) |_{x=x_t^*} = \partial_x BS(t, x, I_t^0(x)) |_{x=x_t^*}, \quad (8)$$

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

$$\begin{aligned}
& \partial_x (BS(t, x, I_t^0(x)) |_{x=x_t^*}) \\
= & \partial_x BS(t, x^*, I_t^0(x^*)) + \partial_\sigma BS(t, x^*, I_t^0(x^*)) \frac{\partial I_t^0}{\partial x}(x_t^*).
\end{aligned} \quad (9)$$

From Romano and Touzi (1996) we know that $\frac{\partial I_t^0}{\partial x}(x_t^*) = 0$. Then, (7), (8) and (9) imply that

$$\frac{\partial I_t}{\partial X_t}(x_t^*) = \frac{\partial_x BS(t, x_t^*, I_t^0(x_t^*)) - \partial_x BS(t, x_t^*, I_t(x_t^*)) + E\left(\int_t^T \partial_x F(s, X_s, v_s) ds | \mathcal{F}_t\right) |_{X_t=x_t^*}}{\partial_\sigma BS(t, x_t^*, I_t(x_t^*))} \Big|_{X_t=x_t^*}. \quad (10)$$

On the other hand, straightforward calculations lead us to

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

and

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

Then

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

and

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

This, together with (10), implies that the result holds. ■

6 Short-time limit behavior

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

Lemma 5 *Assume the model (1) is satisfied. Then $I_t(x_t^*)\sqrt{T-t}$ tends to 0 a.s. as $T \rightarrow t$.*

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

$$P_t := E[e^{-r(T-t)}(K - e^{X_T})_+ | \mathcal{F}_t] \Bigg|_{X_t=x_t^*} \xrightarrow{T \rightarrow t} (K - e^{x_t^*})_+ = 0, \quad a.s.$$

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

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

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

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

and this allows us to complete the proof. ■

Henceforth we will consider the following hypothesis:

(H1) $\sigma \in \mathbb{L}^{2,4}$

(H2) There exists a constant $a > 0$ such that $\sigma > a > 0$.

(H3) There exists a constant $\delta > -\frac{1}{2}$ such that, for all $0 < t < s < r < T$,

$$E \left((D_s \sigma_r)^2 \middle| \mathcal{F}_t \right) \leq C (r-s)^{2\delta}, \quad (11)$$

$$E \left((D_\theta D_s \sigma_r)^2 \middle| \mathcal{F}_t \right) \leq C (r-s)^{2\delta} (r-\theta)^{-2\delta} \quad (12)$$

Proposition 6 Assume that the model (1) and hypotheses (H1)-(H3) hold. Then:

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

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

Proof: Proposition 4 gives us that

$$\begin{aligned} & \partial_\sigma BS(t, x_t^*, I_t(x_t^*)) \frac{\partial I_t}{\partial X_t}(x_t^*) \\ &= \frac{\rho}{2} E \left(\int_t^T e^{-r(s-t)} (\partial_x - \frac{1}{2}) \partial_x G(s, X_s, v_s) \Lambda_s ds | \mathcal{F}_t \right) \Big|_{X_t=x_t^*} \\ &+ E \left(\int_t^T \int_{\mathbb{R}} e^{-r(s-t)} (\partial_x - \frac{1}{2}) [BS(s, X_s + y, v_s) - BS(s, X_s, v_s)] \nu(dy) ds | \mathcal{F}_t \right) \Big|_{X_t=x_t^*} \\ & \quad - \lambda k E \left(\int_t^T e^{-r(s-t)} (\partial_x - \frac{1}{2}) \partial_x BS(s, X_s, v_s) ds | \mathcal{F}_t \right) \Big|_{X_t=x_t^*} = T_1 + T_2 + T_3. \quad (13) \end{aligned}$$

Now the proof will be decomposed into several steps.

Step 1. Here we see that

$$T_1 = \frac{\rho}{2} E(L(t, x_t^*, v_t) \int_t^T \Lambda_s ds | \mathcal{F}_t) + O(T-t)^{1+2\delta}, \quad (14)$$

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 formula to

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

as in the proof of Theorem 3 and taking conditional expectations with respect \mathcal{F}_t , we obtain that

$$\begin{aligned}
& \frac{\rho}{2} E \left(\int_t^T e^{-r(s-t)} L(s, X_s, v_s) \Lambda_s ds \middle| \mathcal{F}_t \right) = \frac{\rho}{2} E(L(t, X_t, v_t) \left(\int_t^T \Lambda_s ds \right) \middle| \mathcal{F}_t) \\
& + \frac{\rho^2}{4} E \left(\int_t^T e^{-r(s-t)} (\partial_{xxx}^3 - \partial_{xx}^2) L(s, X_s, v_s) \left(\int_s^T \Lambda_r dr \right) \Lambda_s ds \middle| \mathcal{F}_t \right) \\
& + \frac{\rho^2}{2} E \left(\int_t^T e^{-r(s-t)} \partial_x L(s, X_s, v_s) \left(\int_s^T D_s \Lambda_r dr \right) \sigma_s ds \middle| \mathcal{F}_t \right) \\
& + \frac{\rho}{2} E \left(\int_t^T \int_{\mathbb{R}} e^{-r(s-t)} [L(s, X_{s-} + y, v_s) - L(s, X_{s-}, v_s)] \left(\int_s^T \Lambda_r dr \right) \nu(dy) ds \middle| \mathcal{F}_t \right) \\
& - \lambda k \frac{\rho}{2} E \left(\int_t^T e^{-r(s-t)} \partial_x L(s, X_{s-}, v_s) \left(\int_s^T \Lambda_r dr \right) ds \middle| \mathcal{F}_t \right) \\
& = \frac{\rho}{2} E(L(t, X_t, v_t) \left(\int_t^T \Lambda_s ds \right) \middle| \mathcal{F}_t) + S_1 + S_2 + S_3 + S_4.
\end{aligned}$$

Using Lemma 2 we can write

$$\begin{aligned}
S_1 &= \frac{\rho^2}{4} E \left(\int_t^T e^{-r(s-t)} E \left[(\partial_{xxx}^3 - \partial_{xx}^2) L(s, X_s, v_s) \middle| \mathcal{G}_t \right] \left(\int_s^T \Lambda_r dr \right) \Lambda_s ds \middle| \mathcal{F}_t \right) \\
&\leq C \sum_{k=4}^6 E \left[\left(\int_t^T \sigma_s^2 ds \right)^{-\frac{k}{2}} \int_t^T \left| \left(\int_s^T \Lambda_r dr \right) \Lambda_s \right| ds \middle| \mathcal{F}_t \right].
\end{aligned}$$

Taking into account inequality (4) and using Hypotheses (H2) and (H3) we can write

$$\begin{aligned}
S_1 &\leq C \sum_{k=0}^2 E \left[\left(\int_t^T \sigma_\theta^2 d\theta \right)^{-\frac{k}{2}} \left(\int_t^T \int_t^T (D_r \sigma_\theta)^2 dr d\theta \right) \middle| \mathcal{F}_t \right] \\
&\leq C(T-t)^{-1} \left(\int_t^T \int_t^\theta (\theta-r)^{2\delta} dr d\theta \right) \leq C(T-t)^{1+2\delta}.
\end{aligned}$$

Using similar arguments it follows that $S_2 + S_3 + S_4 = O(T-t)^{1+2\delta}$, which proves (14).

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

Step 3. Let us prove that

$$T_3 = -\lambda k E(G(t, x_t^*, v_t)(T-t) \middle| \mathcal{F}_t) + O(T-t). \quad (15)$$

In fact,

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

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

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

Now, using again the same arguments of *Step 1*, (15) follows. Therefore the proof is complete. ■

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

(H4) σ has a.s. right-continuous trajectories.

(H5) For every fixed $t > 0$, $\sup_{s,r,\theta \in [t,T]} E\left(\left(\sigma_s \sigma_r - \sigma_\theta^2\right)^2 \middle| \mathcal{F}_t\right) \rightarrow 0$ as $T \rightarrow t$.

Theorem 7 *Consider the model (1) and suppose that Hypotheses (H1)-(H5) hold:*

1. Assume that δ in (H3) is nonnegative and that there exists a \mathcal{F}_t -measurable random variable $D_t^+ \sigma_t$ such that, for every $t > 0$,

$$\sup_{s,r \in [t,T]} |E\left((D_s \sigma_r - D_t^+ \sigma_t) \middle| \mathcal{F}_t\right)| \rightarrow 0, \quad (16)$$

a.s. as $T \rightarrow t$. Then

$$\lim_{T \rightarrow t} \frac{\partial I_t}{\partial X_t}(x_t^*) = -\frac{1}{\sigma_t} \left(\lambda k + \rho \frac{D_t^+ \sigma_t}{2} \right). \quad (17)$$

2. Assume that δ in (H3) is negative and that there exists a \mathcal{F}_t -measurable random variable $L_t^{\delta,+} \sigma_t$ such that, for every $t > 0$,

$$\frac{1}{(T-t)^{2+\delta}} \int_t^T \int_s^T E(D_s \sigma_r \middle| \mathcal{F}_t) dr ds - L_t^{\delta,+} \sigma_t \rightarrow 0, \quad (18)$$

a.s. as $T \rightarrow t$. Then

$$\lim(T-t)^{-\delta} \frac{\partial I_t}{\partial X_t}(x_t^*) = -\frac{\rho}{\sigma_t} L_t^{\delta,+} \sigma_t \quad (19)$$

Proof: Using Proposition 6 and the facts that

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

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

and

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

we can write

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

By Lemma 5 we know that $I_t(x_t^*)^2(T-t) \rightarrow 0$ as $T \rightarrow t$. Then,

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

Using again Lemma 5, observe that (H4) and the dominated convergence theorem imply that

$$\lim_{T \rightarrow t} S_2 = -\frac{\lambda k}{\sigma_t} \quad (20)$$

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

$$\lim_{T \rightarrow t} \left(S_1 + \frac{\rho}{2\sigma_t} D_t^+ \sigma_t \right) = 0. \quad (21)$$

Indeed, we can write

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

where

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

and

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

Notice that

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

Applying Schwartz inequality it follows that

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

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

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

Using now Schwartz's inequality and the fact that Hypothesis (H3) holds with $\delta \geq 0$ we obtain that

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

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

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

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

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

which tends to zero by 16. Now we have proved (21). Then, (20), (21) and the fact that $\delta \geq 0$ give (17) holds.

Step 2. Finally we show that 19 is true. Let us see that

$$\lim_{T \rightarrow t} \left(S_1 (T-t)^{-\delta} + \frac{\rho}{\sigma_t} L_t^{\delta,+} \sigma_t \right) = 0. \quad (22)$$

Note that

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

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

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

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 (22). Finally, using (22), (20) and the fact that $-\frac{1}{2} < \delta < 0$ the result follows. ■

Remark: Notice that (17) and (19) can be written in terms of $\frac{\partial I_t}{\partial Z}$, being $Z = \log K$, the log-strike, simply by changing the sign of the limits.

7 Examples

7.1 Diffusion stochastic volatilities

Assume that the volatility σ_r can be written as $\sigma_r = f(Y_r)$, where $f \in C_b^1(\mathbb{R})$ and Y_r is the solution of a stochastic differential equation:

$$dY_r = a(r, Y_r) dr + b(r, Y_r) dW_r, \quad (23)$$

for some real functions $a, b \in C_b^1(\mathbb{R})$. Then, classical arguments (see for example Nualart (1995)) give us that $Y \in \mathbb{L}^{1,2}$ and that

$$D_s Y_r = \int_s^r \frac{\partial a}{\partial x}(u, Y_u) D_s Y_u du + b(s, Y_s) + \int_s^r \frac{\partial b}{\partial x}(u, Y_u) D_s Y_u dW_u. \quad (24)$$

Taking now into account that $D_s \sigma_r = f'(Y_r) D_s Y_r$ it can be easily deduced from (24) that (H3) holds with $\delta = 0$ and that

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

as $T \rightarrow t$. Then, Theorem 7 gives us that

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

that agrees with the results in Medvedev and Scaillet (2004).

In particular, if Y_r is an Ornstein-Uhlenbeck process of the form

$$Y_r = m + (Y_t - m) e^{-\alpha(r-t)} + c \int_t^r \sqrt{2\alpha} \exp(-\alpha(r-s)) dW_s, \quad (25)$$

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

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

7.2 Fractional stochastic volatility models

Assume that the volatility σ can be written as $\sigma_r = f(Y_r)$, where $f \in C_b^1(\mathbb{R})$ and Y_r is a process of the form

$$Y_r = m + (Y_t - m) e^{-\alpha(r-t)} + c\sqrt{2\alpha} \int_t^r e^{-\alpha(r-s)} dW_s^H, \quad (26)$$

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 Comte and Renault (1998), assume the volatility model (26), for some $H > 1/2$. Notice that (see for example Alòs, Mazet and Nualart (2001)) $\int_t^r e^{-\alpha(r-s)} dW_s^H$ 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 where it follows easily that $\sup_{s,r \in [t,T]} |E(D_s \sigma_r | \mathcal{F}_t)| \rightarrow 0$ as $T \rightarrow t$. Then, Theorem 7 gives us that $\lim_{T \rightarrow t} \frac{\partial I_t}{\partial X_t}(x_t^*) = -\frac{\lambda k}{\sigma_t}$. 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 (26), taking $0 < H < 1/2$. It can be proved (see for example Alòs, Mazet and Nualart (2001)) that $\int_t^r e^{-\alpha(r-s)} dW_s^H$ can be expressed as

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

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

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

Then, Theorem 7 gives us that

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

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, Papanicolau, Sircar and Solna (2004) 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_r can be written as $\sigma_r = f(Y_r)$, where f is a regular enough function and Y_r is a diffusion process of the form (25), being $\sqrt{\alpha(s)}$ 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_r to be a diffusion process of the form (25), with $\sqrt{\alpha(s)} = (T_{n(s)} - s)^{-\frac{1}{2}+\varepsilon}$, for some $\varepsilon > 0$. It is now easy to see that $Y_r \in \mathbb{L}^{1,2}$ 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 the 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 do not require the volatility to be a diffusion, nor a Markov process. Moreover, with these techniques the short-time behaviour 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 Hull and White formula with applications to option pricing aproximation. To appear in Finance and Stochastics (2006).
- [2] Alòs, E., Mazet, O. and Nualart, D.: Stochastic Calculus with respect to Gaussian processes. The Annals of Probability **29**(2), 766-801 (2001).
- [3] Alòs, E. and Nualart, D. : An extension of Itô's formula for anticipating processes. Journal of Theoretical Probability **11**(2), 493-514 (1998).
- [4] Ball, C. A. and Roma, A.: Stochastic volatility option pricing. Journal of Financial and Quantitative Analysis, **29**, 589-607 (1994).
- [5] Barndorff-Nielsen, O. E. and Shephard, N.: Modelling by Lévy processes for financial econometrics. In Lévy processes-Theory and Applications, Barndorff-Nielsen, O. Mikosch, T. and Resnick, S., eds., Birkhäuser, 283-318 (2001).
- [6] Barndorff-Nielsen, O. E. and Shephard, N.: Econometric analysis of realized volatility and its use in estimating stochastic volatility models. J. R. Statistic. Soc. B, **64**, 253-280 (2002).
- [7] Bates, D.: Jumps and stochastic volatility: the exchange rate processes implicit in Deutschemark options. The Review of Financial Studies, **9** (1), 69-107 (1996).

- [8] Carr, P. and Wu, L.: The finite moment logstable process and option pricing. *Journal of Finance*, **58**, 753-778 (2003).
- [9] Comte, F. and Renault, E.: Long-memory in continuous-time stochastic volatility models. *Mathematical Finance*, **8**, 291-323 (1998).
- [10] Fouque, J. P., Papanicolaou, G., Sircar, K. R. and Solna, K.: Singular Perturbations in Option Pricing. *SIAM J. Applied Math*, **63**(5), 1648-1665 (2003).
- [11] Fouque, J. P., Papanicolaou, G. Sircar, K. R. and Solna, K.: Maturity Cycles in Implied Volatility, *Finance and Stochastics*, **8**(4) 451-477 (2004).
- [12] Heston, S. L.: A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, **6**, 327-343 (1993).
- [13] Hull, J. C. and White, A.: The pricing of options on assets with stochastic volatilities. *Journal of Finance*, **42**, 281-300 (1987).
- [14] Lee, R.: Implied Volatility: Statics, Dynamics and Probabilistic Interpretation, in *Recent Advances in Applied Probability*. Springer (2004).
- [15] Lewis, A.: *Option Valuation Under Stochastic Volatility*. Finance press (2000).
- [16] Medvedev, A. and Scaillet, O.: A simple calibration procedure of stochastic volatility models with jumps by short term asymptotics. Discussion paper HEC, Genève and FAME, Université de Genève (2004).
- [17] Nualart, D.: *The Malliavin calculus and related topics*. Springer (1995).
- [18] Renault, E and Touzi, N.: Option hedging and implied volatilities in a stochastic volatility model. *Mathematical Finance* **6** (3), 279-302 (1996).
- [19] Scott, L. O.: Option pricing when the variance changes randomly: Theory, estimation and application. *Journal of Financial and Quantitative Analysis*, **22**, 419-438 (1987).
- [20] Stein, E. M. and Stein, J. C.: Stock price distributions with stochastic volatility: An analytic approach. *The review of Financial Studies* **4**, 727-752 (1991).