From characteristic functions to implied volatility expansions

Antoine Jacquier * Matthew Lorig †

This version: July 9, 2013

Abstract

For any strictly positive martingale $S = \mathrm{e}^X$ for which X has an analytically tractable characteristic function, we provide an expansion for the implied volatility. This expansion is explicit in the sense that it involves no integrals, but only polynomials in $\log(K/S_0)$. We illustrate the versatility of our expansion by computing the approximate implied volatility smile in three well-known martingale models: one finite activity exponential Lévy model (Merton), one infinite activity exponential Lévy model (Variance Gamma), and one stochastic volatility model (Heston). We show how this technique can be extended to compute approximate forward implied volatilities and we implement this extension in the Heston setting. Finally, we illustrate how our expansion can be used to perform a model-free calibration of the empirically observed implied volatility surface.

Keywords: Implied volatility expansions, exponential Lévy, affine class, Heston, additive process,

1 Introduction

While it is rare to find a martingale model for which the transition density is available in closed-form (the Black-Scholes model being a notable exception), there is a veritable zoo of models for which the characteristic function is available explicitly (exponential Lévy models and affine models [9] for instance). The existence of an analytically tractable characteristic function allows for (vanilla) option prices to be computed quickly using (generalised) Fourier transforms [25, 26].

Every model contains unobservable parameters, which are usually calibrated to market data. This calibration procedure is typically performed using implied volatilities rather than option prices, the former being dimensionless. For a given model, one therefore has to compute (by finite difference, Monte Carlo or

*Department of Mathematics, Imperial College London, London, United Kingdom.

†ORFE Department, Princeton University, Princeton, USA. Work partially supported by NSF grant DMS-0739195.

numerical integration) option prices first and then the corresponding implied volatilities by some root-finding algorithm. Both steps require sophisticated numerical tools and occasionally somewhat of an artistic touch. These are computationally expensive and render calibration a long and intensive task.

Over the past decade, many authors have focused on obtaining closed-form approximations for both option prices and implied volatilities, partly in order to speed up this calibration process. Perturbation methods have been used by Lorig and co-authors [28] (see also [10, 11, 13, 19]) to obtain such approximations for diffusion-type models. In extreme regions—where numerical schemes become less efficient—asymptotic expansions of densities and of implied volatilities have been obtained in [3, 7, 15, 20, 33] in the small-maturity case (both for diffusions and jump models) and in [22] for the large-time behaviour of affine stochastic volatility models. Roger Lee [24] pioneered the study of the tails of implied volatility, and more recent (model-dependent and model-free) results have appeared in [2, 7, 12, 17].

The goal of this paper is to derive an approximation for the implied volatility in any model whose characteristic function is available in closed-form. This approximation contains no special function and does not require any numerical integration. It can therefore be used efficiently to accelerate the aforementioned calibration issue. The methodology follows and extends the previous works [27] and is related to some extent to the works by Takahashi and Toda [32]. Indeed, by writing the characteristic function as a perturbation around the Black-Scholes characteristics function, our expansion has the form of a Black-Scholes price perturbed by some additional quantity (which we shall make precise later), which can then be turned into an expansion for the corresponding implied volatility.

The rest of the paper proceeds as follows: in Section 2, we provide a brief review of the characteristic function approach to option pricing and introduce some notations needed later in the paper. Section 3 contains the main results, namely a series expansion for the implied volatility. More precisely, we show (Section 3.1) that, whenever the characteristic function is available in closed-form, the European call price can be written as a regular perturbation around the Black-Scholes price. A similar result then holds for the implied volatility, as detailed in Section 3.2 and 3.3; this can be further applied to the so-called forward implied volatility as outlined in Section 3.4. In Section 4 we numerically test our results and provide practical details about this implementation.

2 Notations and preliminary results

We consider here a given probability space $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$; all the processes studied will be $\mathcal{F}$-adapted. In particular $S = \mathrm{e}^X$ will denote the stock price process, namely a $\mathcal{F}$-adapted martingale under the risk-neutral probability measure $\mathbb{P}$. The dynamics of X may depend on some auxiliary process $Y \in \mathbb{R}^m$ ($m \geq 1$), say some stochastic volatility. The starting point $(X_0, Y_0) = (x, y)$ is assumed to be non-random. For simplicity and notational convenience, we will assume that $m = 1$ and that the risk-free interest rate is zero.

2.1 Pricing via Fourier transforms

Let h be the payoff function of a European call option on S with strike e^k: $h(z) \equiv (\mathrm{e}^z - \mathrm{e}^k)^+$, and denote $\widehat{h}$ its (generalised) Fourier transform

$$\widehat{h}(\lambda) := \int_{\mathbb{R}} \mathrm{e}^{-\mathtt{i}\lambda z} h(z) \mathrm{d}z = \frac{-\mathrm{e}^{k-\mathtt{i}k\lambda}}{\mathtt{i}\lambda + \lambda^2}, \qquad \text{for } \Im(\lambda) < -1.$$

The results obtained below for option prices remain valid for put options with payoff $h(z) \equiv (\mathrm{e}^k - \mathrm{e}^z)^+$, but we shall chiefly consider European call option prices unless otherwise stated. For any $t \geq 0$, define the moment explosions $p^*(t) := \sup\{p \geq 0 : \mathbb{E}_x (S_t^p) < \infty\}$ and $q^*(t) := \sup\{q \geq 0 : \mathbb{E}_x \left(S_t^{-q}\right) < \infty\}$. Since S is a martingale, we have $p^*(t) \geq 1$ and $q^*(t) \geq 0$. We shall further make the stronger assumption:

Assumption 1. For any $t \geq 0$, $p^*(t) > 1$ and $q^*(t) > 0$.

This assumption holds for most models in practice, and allows us to write the value of a call option as

$$u(t, x) := \mathbb{E}_x \, h(X_t) = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{h}(\lambda) \mathbb{E}_x \left(\mathrm{e}^{\mathtt{i}\lambda X_t}\right) \mathrm{d}\lambda_r, \qquad \text{with } \Im(\lambda) \in (-p^*(t), -1), \text{ for all } t \geq 0, \tag{1}$$

where we write $\lambda = \lambda_r + \mathtt{i}\lambda_i$ ($\lambda_r, \lambda_i \in \mathbb{R}$) for a complex number. Of course the function u also depends on y, the starting point of Y, but we shall omit y in the notations for clarity. In this paper, we consider models for which the characteristic function $\mathbb{C} \ni \lambda \mapsto \mathbb{E}_x \left(\mathrm{e}^{\mathtt{i}\lambda X_t}\right)$ admits the representation

$$\log \mathbb{E}_x \left(\mathrm{e}^{\mathtt{i}\lambda X_t}\right) = \mathtt{i}\lambda x + \phi(t, \lambda), \tag{2}$$

for some analytic function $\phi(\cdot, \cdot) : \mathbb{R}_+ \times \mathbb{C} \to \mathbb{C}$, satisfying $\phi(t, -\mathtt{i}) = 0$ for all $t \geq 0$ (martingale property). From (1), this implies that the price of a call option may be written as (see also [25] or [26])

$$u(t, x) = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{h}(\lambda) \mathrm{e}^{\mathtt{i}\lambda x + \phi(t,\lambda)} \mathrm{d}\lambda_r.$$

Several well-known models fit within this class

$$\begin{aligned}
\text{Lévy models :} \qquad & \phi(t, \lambda) = t \left(\mathtt{i}\mu\lambda - \tfrac{1}{2} a^2 \lambda^2 + \int_{\mathbb{R}} \nu(dz)(\mathrm{e}^{\mathtt{i}\lambda z} - 1 - \mathtt{i}\lambda z) \right), \\
\text{Additive models :} \qquad & \phi(t, \lambda) = \mathtt{i}\mu(t)\lambda - \tfrac{1}{2} a^2(t) \lambda^2 + \int_{\mathbb{R}} \nu(t, dz)(\mathrm{e}^{\mathtt{i}\lambda z} - 1 - \mathtt{i}\lambda z), \\
\text{Affine models :} \qquad & \phi(t, \lambda) = C(t, \lambda) + y D(t, \lambda),
\end{aligned}$$

where (μ, a^2, ν) is a Lévy triplet, $(\mu(t), a^2(t), \nu(t))$ are the spot characteristics of an additive process, the function C is fully characterised by $\frac{\mathrm{d}}{\mathrm{d}t}C = D$ and the function D satisfies a Riccati equation. For precise details on Lévy and affine processes, we refer the interested reader to the monograph by Sato [31] and the groundbreaking paper by Duffie, Filipović and Schachermayer [8].

2.2 Black-Scholes and implied volatility

Option prices are commonly quoted in units of implied volatility (rather than in units of currency) first because the latter is dimensionless, and second, because the shape and behaviour of the implied volatility provide more information than option prices. However, implied volatility is scarcely available in closed-form and has to be computed numerically via inversion of the Black-Scholes formula. We derive here a closed-form expansion for the implied volatility for models whose characteristic function is of the form (2). We begin our analysis by defining the Black-Scholes price and the implied volatility.

Definition 2. The *Black-Scholes price* $u^{\mathrm{BS}} : \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}^+$ is given by

$$u^{\mathrm{BS}}(t, x, \sigma_0) := \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{e}^{t\phi_0(\lambda;\sigma_0)} \widehat{h}(\lambda) \mathrm{e}^{\mathtt{i}\lambda x} \mathrm{d}\lambda_r, \qquad \text{where} \qquad \phi_0(\lambda; \sigma_0) := -\frac{1}{2}\sigma_0^2 \left(\lambda^2 + \mathtt{i}\lambda\right).$$

Note that $\phi_0(\cdot; \sigma_0)$ is simply the Lévy exponent of a Brownian motion with volatility σ_0 and drift $-\frac{1}{2}\sigma_0^2$. Thus, this is simply the Fourier representation of the usual Black-Scholes price.

Definition 3. For any maturity t, starting point x and (log) strike k, the *implied volatility* is defined as the unique strictly positive real solution σ to the equation $u^{\mathrm{BS}}(t, x, \sigma) = u$, where u is the (observed or computed) call option price with the same maturity and log strike.

Remark 4. For any $t > 0$, the existence and uniqueness of the implied volatility can be deduced using the general arbitrage bounds for call prices and the monotonicity of u^{BS} (see [11, Section 2.1, Remark (i)]).

For any $t \geq 0$, $x \in \mathbb{R}$, the function $u^{\mathrm{BS}}(t, x, \cdot)$ is analytic on $\mathbb{R}_+^*$, and hence for any $\sigma_0 > 0$ and $\delta \in \mathbb{R}$ such that $\sigma_0 + \delta > 0$, the function $u^{\mathrm{BS}}(t, x, \cdot)$ at the point $\sigma_0 + \delta$ is given by its Taylor series:

$$u^{\mathrm{BS}}(t, x, \sigma_0 + \delta) = \sum_{n=0}^{\infty} \frac{\delta^n}{n!} \partial_\sigma^n u^{\mathrm{BS}}(t, x, \sigma_0), \tag{3}$$

where $\partial_\sigma^n u^{\mathrm{BS}}(t, x, \sigma_0) = \frac{1}{2\pi} \int_{\mathbb{R}} \left(\partial_\sigma^n \mathrm{e}^{t\phi_0(\lambda;\sigma)}\right)\big|_{\sigma=\sigma_0} \widehat{h}(\lambda) \mathrm{e}^{\mathtt{i}\lambda x} \mathrm{d}\lambda_r$. The interchange of the derivative and integral operators is justified by Fubini's theorem. If one observes the option price u, then the following proposition provides a way to compute the corresponding implied volatility.

Proposition 5. *For any $t > 0$, $x \in \mathbb{R}$, let $u : \mathbb{R}_+^* \to \mathbb{R}_+^*$ be defined (as a function of σ) by $u^{\mathrm{BS}}(t, x, \sigma) = u$, and let σ_0 be some strictly positive real number. Then the following expansion holds:*

$$\sigma = \sigma_0 + \sum_{n=1}^{\infty} \frac{b_n}{n!} (u - u^{\mathrm{BS}}(t, x, \sigma_0))^n, \tag{4}$$

where $b_n := \lim_{\sigma \to \sigma_0} \partial_\sigma^{n-1} \left(\frac{\sigma - \sigma_0}{u^{\mathrm{BS}}(t,x,\sigma) - u^{\mathrm{BS}}(t,x,\sigma_0)} \right)^n$.

Proof. Since the function $u^{\mathrm{BS}}(t, x, \cdot)$ is strictly increasing on $\mathbb{R}_+^*$, analytic in a neighbourhood of σ_0 and $\partial_\sigma u^{\mathrm{BS}}(\cdot, \cdot, \sigma_0) \neq 0$, the proposition follows from Lagrange Inversion Theorem [1, Equation 3.6.6]. $\square$

Proposition 5 shows that, for every fixed $t > 0$, $x \in \mathbb{R}$, $\sigma_0 > 0$, there exists some radius of convergence $R > 0$ (depending on (t, x, k)) such that $|u - u^{\mathrm{BS}}(t, x, \sigma_0)| < R$ implies that σ, defined implicitly through the equation $u^{\mathrm{BS}}(t, x, \sigma) = u$, is fully characterised by (4). This result however seems to be only of theoretical interest. Once the option value u is known, computing the implied volatility inverting the Black-Scholes formula is a simple numerical exercise. Moreover, computing the implied volatility using (4) is not numerically efficient since the option price u requires the computation of a (possibly highly oscillatory) Fourier integral. One may wish to use (4) to deduce some properties of the implied volatility, but then the proposition would benefit from precise error bounds when truncating the infinite sum. The rest of the paper focuses on developing a similar expansion, without the need for the (potentially computer-intensive) implementation of the value function u.

3 Implied volatility expansions

3.1 Call prices as perturbations around Black-Scholes

For any $\varepsilon \in (0, 1]$ and $\sigma_0 > 0$ define the function $\phi^\varepsilon(\cdot, \cdot; \sigma_0) : \mathbb{R}_+ \times \mathbb{C} \to \mathbb{C}$ by

$$\phi^\varepsilon(t, \lambda; \sigma_0) := t\, \phi_0(\lambda; \sigma_0) + \varepsilon\, \phi_1(t, \lambda; \sigma_0),$$

where $\phi_1(t, \lambda; \sigma_0) := \phi(t, \lambda) - t\, \phi_0(\lambda; \sigma_0)$, and ϕ_0 is the Black-Scholes characteristic function from Definition 2. Recall from Bochner theorem [29, Theorem 4.2.2] that a complex-valued function f is a characteristic function if and only if it is non-negative definite and $f(0) = 1$. Therefore ϕ^ε is a well-defined characteristic function for any $t \geq 0$, and we can associate to it a (unique up to indistinguishability) stochastic process $(X_t^{\varepsilon,\sigma_0})_{t \geq 0}$, starting at $X_0^{\varepsilon,\sigma_0} = x$, which is a true martingale. The price u^ε of a call option written on X^{ε,σ_0} thus reads

$$u^\varepsilon(t, x, \sigma_0) := \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{d}\lambda_r\, \mathrm{e}^{\phi^\varepsilon(t,\lambda;\sigma_0)} \widehat{h}(\lambda) \mathrm{e}^{\mathrm{i}\lambda x}. \tag{5}$$

Let σ^ε denote the implied volatility corresponding to the option price $u^\varepsilon(t,x,\sigma_0)$. Since $\phi^\varepsilon|_{\varepsilon=1} = \phi$ and $u^\varepsilon|_{\varepsilon=1} = u$, the implied volatility corresponding to the option price u is given by $\sigma = \sigma^\varepsilon|_{\varepsilon=1}$. We now seek an expression for σ^ε. The first step is to show that u^ε can be written as a power series in ε, whose first term corresponds to the Black-Scholes call price with volatility σ_0. To this end, we first expand $\mathrm{e}^{\phi^\varepsilon(t,\lambda;\sigma_0)}$ as

$$\exp\left(\phi^\varepsilon(t,\lambda;\sigma_0)\right) = \mathrm{e}^{t\phi_0(\lambda;\sigma_0)} \sum_{n=0}^{\infty} \frac{1}{n!} \varepsilon^n \phi_1^n(t,\lambda;\sigma_0),$$

and deduce a series representation for u^ε in (5):

$$u^\varepsilon(t,x,\sigma_0) = \sum_{n=0}^{\infty} \varepsilon^n u_n(t,x,\sigma_0), \qquad \text{with} \qquad u_n(t,x,\sigma_0) := \frac{1}{n!}\frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{d}\lambda_r \mathrm{e}^{t\phi_0(\lambda;\sigma_0)} \phi_1^n(t,\lambda;\sigma_0) \widehat{h}(\lambda) \mathrm{e}^{\mathrm{i}\lambda x}, \quad (6)$$

for any $n \geq 0$, where the application of Fubini's theorem is justified since $\int_{\mathbb{R}} \left| \mathrm{e}^{t\phi^\varepsilon(\lambda)} \widehat{h}(\lambda) \mathrm{e}^{\mathrm{i}\lambda x} \right| \mathrm{d}\lambda_r$ is finite. Note in particular that $u_0(t,x,\sigma_0) \equiv u^{\mathrm{BS}}(t,x,\sigma_0)$.

3.2 Series expansion for implied volatility

From (6), it is clear that u^ε is an analytic function of ε (we have explicitly provided its power series representation). Since the composition of two analytic functions is also analytic [5, Section 24, p. 74], the expansion (3) implies that $\sigma^\varepsilon = [u^{\mathrm{BS}}]^{-1}(u^\varepsilon)$ is an analytic function and therefore has a power series expansion in ε, which we write $\sigma^\varepsilon := \sigma_0 + \delta^\varepsilon$, where $\delta^\varepsilon = \sum_{k\geq 1} \varepsilon^k \sigma_k$. The following theorem provides an expansion formula for the coefficients σ_k.

Theorem 6. *Fix $\sigma_0 > 0$, $k \geq 1$, and let R denote the radius of convergence of the expansion* (4). *If $\left|u(t,x) - u^{\mathrm{BS}}(t,x,\sigma_0)\right| < R$ for all $(t,x) \in \mathbb{R}_+ \times \mathbb{R}$, then the following expansion holds:*

$$\sigma_k = \frac{1}{\partial_\sigma u^{\mathrm{BS}}(t,x,\sigma_0)} \left(u_k - \sum_{n=2}^{\infty} \frac{1}{n!} \left(\sum_{j_1+\cdots+j_n=k} \prod_{i=1}^{n} \sigma_{j_i} \right) \partial_\sigma^n u^{\mathrm{BS}}(t,x,\sigma_0) \right). \quad (7)$$

Note that the right-hand side of (7) involves only σ_j for $j \leq k-1$, and hence the sequence $(\sigma_k)_{k\geq 1}$ can be determined recursively.

Proof. Let us fix some $t \geq 0$ and $x \in \mathbb{R}$. Taylor expanding $u^{\mathrm{BS}}(t,x,\sigma^\varepsilon)$ around the point σ_0 we obtain

$$
\begin{aligned}
u^{\mathrm{BS}}(t,x,\sigma^\varepsilon) = u^{\mathrm{BS}}(t,x,\sigma_0+\delta^\varepsilon) &= \sum_{n=0}^{\infty} \frac{1}{n!}(\delta^\varepsilon \partial_\sigma)^n u^{\mathrm{BS}}(t,x,\sigma_0) \\
&= u^{\mathrm{BS}}(t,x,\sigma_0) + \sum_{n=1}^{\infty} \frac{1}{n!} \left(\sum_{k=1}^{\infty} \varepsilon^k \sigma_k \right)^n \partial_\sigma^n u^{\mathrm{BS}}(t,x,\sigma_0) \\
&= u^{\mathrm{BS}}(t,x,\sigma_0) + \sum_{n=1}^{\infty} \frac{1}{n!} \left[\sum_{k=1}^{\infty} \left(\sum_{j_1+\cdots+j_n=k} \prod_{i=1}^{n} \sigma_{j_i} \right) \varepsilon^k \right] \partial_\sigma^n u^{\mathrm{BS}}(t,x,\sigma_0) \\
&= u^{\mathrm{BS}}(t,x,\sigma_0) + \sum_{k=1}^{\infty} \varepsilon^k \left[\sum_{n=1}^{\infty} \frac{1}{n!} \left(\sum_{j_1+\cdots+j_n=k} \prod_{i=1}^{n} \sigma_{j_i} \right) \partial_\sigma^n \right] u^{\mathrm{BS}}(t,x,\sigma_0) \\
&= u^{\mathrm{BS}}(t,x,\sigma_0) + \sum_{k=1}^{\infty} \varepsilon^k \left[\sigma_k \partial_\sigma + \sum_{n=2}^{\infty} \frac{1}{n!} \left(\sum_{j_1+\cdots+j_n=k} \prod_{i=1}^{n} \sigma_{j_i} \right) \partial_\sigma^n \right] u^{\mathrm{BS}}(t,x,\sigma_0).
\end{aligned}
$$

In order to recover the implied volatility from Definition 3, we need to equate the Black-Scholes call price above and the option value u^ε in (6), and collect terms of identical powers of ε:

$$
\begin{aligned}
\mathcal{O}(1): &\qquad u_0(t,x,\sigma_0) = u^{\mathrm{BS}}(t,x,\sigma_0), \\
\mathcal{O}(\varepsilon^k): &\qquad u_k(t,x,\sigma_0) = \sigma_k \partial_\sigma u^{\mathrm{BS}}(t,x,\sigma_0) + \sum_{n=2}^{\infty} \frac{1}{n!} \left(\sum_{j_1+\cdots+j_n=k} \prod_{i=1}^{n} \sigma_{j_i} \right) \partial_\sigma^n u^{\mathrm{BS}}(t,x,\sigma_0), \qquad k \geq 1.
\end{aligned}
$$

Solving the above equations for the sequence $(\sigma_k)_{k\geq 0}$, we find $\sigma_0 = \sigma_0$ at the zeroth order and for any $k \geq 1$, the $\mathcal{O}(\varepsilon^k)$ order is given by (7). $\square$

Remark 7. Explicitly, up to $\mathcal{O}(\varepsilon^5)$ we have

$$
\begin{aligned}
\mathcal{O}(\varepsilon): &\quad \sigma_1 = \frac{1}{\partial_\sigma u_0} u_1, \\
\mathcal{O}(\varepsilon^2): &\quad \sigma_2 = \frac{1}{\partial_\sigma u_0} \Big(u_2 - \tfrac{1}{2}\sigma_1^2 \partial_\sigma^2 u_0 \Big), \\
\mathcal{O}(\varepsilon^3): &\quad \sigma_3 = \frac{1}{\partial_\sigma u_0} \Big(u_3 - \left(\sigma_2\sigma_1 \partial_\sigma^2 + \tfrac{1}{3!}\sigma_1^3 \partial_\sigma^3 \right) u_0 \Big), \\
\mathcal{O}(\varepsilon^4): &\quad \sigma_4 = \frac{1}{\partial_\sigma u_0} \Big(u_4 - \left(\tfrac{1}{2}\left(2\sigma_3\sigma_1 + \sigma_2^2\right) \partial_\sigma^2 + \tfrac{1}{2}\sigma_2\sigma_1^2 \partial_\sigma^3 + \tfrac{1}{4!}\sigma_1^4 \partial_\sigma^4 \right) u_0 \Big), \\
\mathcal{O}(\varepsilon^5): &\quad \sigma_5 = \frac{1}{\partial_\sigma u_0} \Big(u_5 - \Big(\left(\sigma_1\sigma_4 + \sigma_2\sigma_3\right) \partial_\sigma^2 + \tfrac{1}{2}\left(\sigma_1^2\sigma_3 + \sigma_1\sigma_2^2\right) \partial_\sigma^3 + \tfrac{1}{3}\sigma_1^3\sigma_2 \partial_\sigma^4 + \tfrac{1}{5!}\sigma_1^5 \partial_\sigma^5 \Big) u_0 \Big),
\end{aligned}
$$

where all the functions $u_0, \ldots, u_5$ are evaluated at (t,x,σ_0).

Remark 8. Having served its purpose, we now dial ε up to one. The implied volatility is then given by $\sigma = \sum_{k\geq 0} \sigma_k$, where σ_0 is a fixed positive constant and where the sequence $(\sigma_k)_{k\geq 1}$ is given by (7).

3.3 Simplification of the expressions for σ_k

The expression for the coefficients σ_k ($k \geq 1$) in (7) is not straightforward to apply; indeed, one needs to compute first the Fourier integrals u_j ($j \leq k$) via (6), then all the terms of the form $\partial_\sigma^j u_0$ ($j \leq k$). We provide now a more explicit approximation—without integrals or special functions—for σ_k. The key to this simplification is that all the terms $\partial_\sigma^i u_0$ and u_i ($i \in \mathbb{N}$) in (7) can actually be expressed in terms of derivatives of u_0 with respect to x, the starting point of the log stock price process. Indeed, the classical Black-Scholes relation between the Delta, the Gamma and the Vega for call options, $\partial_\sigma u_0(t,x,\sigma)|_{\sigma=\sigma_0} = t\sigma_0(\partial_x^2 - \partial_x)u_0(t,x,\sigma_0)$, implies that the derivative $\partial_\sigma^k u_0$ can be expressed as a sum of terms of the form $a_{k_i}\partial_x^{k_i}(\partial_x^2 - \partial_x)u_0$. We shall also use the equality $p(\lambda)\mathrm{e}^{\mathtt{i}\lambda x} = p(-\mathtt{i}\partial_x)\mathrm{e}^{\mathtt{i}\lambda x}$, which holds for any polynomial p (and actually for any analytic function—simply take p to be its power series). We first start with the following theorem, which provides an approximation for the coefficients u_n in (6) as a differential operator acting on u_0.

Theorem 9. *Fix some $t \geq 0$ and $\sigma_0 > 0$. If the power series $\phi_1(t,\lambda;\sigma_0) = \sum_{k\geq 1} a_k(t;\sigma_0)(\mathtt{i}\lambda)^k$ holds in a complex neighbourhood of the origin, then for any integer $m \geq 2$, u_n defined in (6) can be written as $u_n(t,x,\sigma_0) = u_n^{(m)}(t,x,\sigma_0) + \varepsilon_n^{(m)}(t,x,\sigma_0)$, where*

$$u_n^{(m)}(t,x,\sigma_0) := \frac{1}{n!}\sum_{\substack{2\leq k_2,\ldots,k_m\leq k \\ k_2+\ldots+k_m=n}} \frac{n!}{k_2!\ldots k_m!}\prod_{2\leq j\leq m} a_j^{k_j}(t;\sigma_0)\left(\partial_x^j - \partial_x\right)^{k_j} u_0(t,x,\sigma_0), \tag{8}$$

and where $\varepsilon_n^{(m)}$ only contains derivatives (with respect to x) of u_0 of order higher than n.

Remark 10. Note that the power series for ϕ_1 in the theorem starts at $k = 1$, which follows from the fact that the process $\exp(X^{\varepsilon,\sigma_0})$ is conservative. This expansion holds as soon as all the moments of X_t exist and $\lim_{k\nearrow+\infty}|\lambda|^k\mathbb{E}\left(|X_t|^k\right)/k! = 0$ for $|\lambda|$ small enough, which is valid under Assumption 1.

Proof. Assume that the power series for $\phi_1(t,\cdot;\sigma_0)$ holds around the origin, where the coefficients read

$$a_k(t;\sigma_0) = \frac{(-\mathtt{i})^k}{k!}\,\partial_\lambda^k\phi_1(t,\lambda;\sigma_0)\big|_{\lambda=0}. \tag{9}$$

The martingale condition implies $\phi_1(t,-\mathtt{i};\sigma_0) = \sum_{k\geq 1} a_k(t;\sigma_0) = 0$, and hence

$$\begin{aligned}\phi_1(t,\lambda;\sigma_0) &= \sum_{k=1}^{\infty} a_k(t;\sigma_0)(\mathtt{i}\lambda)^k = \mathtt{i}\lambda a_1(t;\sigma_0) + \sum_{k=2}^{\infty} a_k(t;\sigma_0)\left[(\mathtt{i}\lambda)^k - \mathtt{i}\lambda\right] + \mathtt{i}\lambda\sum_{k=2}^{\infty} a_k(t;\sigma_0) \\ &= \sum_{k=2}^{\infty} a_k(t;\sigma_0)\left[(\mathtt{i}\lambda)^k - \mathtt{i}\lambda\right]. \end{aligned}\tag{10}$$

Let now $\phi_1^{(m)}(t,\cdot;\sigma_0) : \mathbb{C} \to \mathbb{C}$ be the truncation of the series (10) at the m-th order, i.e.

$$\phi_1^{(m)}(t,\lambda;\sigma_0) := \sum_{k=2}^{m} a_k(t;\sigma_0)\left[(\mathtt{i}\lambda)^k - \mathtt{i}\lambda\right], \tag{11}$$

and define the operator δ acting on $\phi_1^{(m)}$ by

$$\delta\phi_1^{(m)}(t,\lambda;\sigma_0) := \phi_1(t,\lambda;\sigma_0) - \phi_1^{(m)}(t,\lambda;\sigma_0) = \frac{\lambda^{m+1}}{2\pi\mathtt{i}}\int_{\partial\Gamma}\frac{\phi_1(t,z;\sigma_0)\mathrm{d}z}{z^{m+1}(z-\lambda)} + \mathtt{i}\lambda\sum_{k=1}^{m}a_k,$$

where Γ is a closed set within the radius of convergence of ϕ_1. It is clear that the integral above is nothing else than the remainder of the series expansion around the point $\lambda = 0$. Hence for any $n \geq 1$, u_n defined in (6), can be written as

$$\begin{aligned}
u_n(t,x,\sigma_0) &= \frac{1}{n!}\frac{1}{2\pi}\int_{\mathbb{R}}\mathrm{d}\lambda_r\,\mathrm{e}^{t\phi_0(t,\lambda,\sigma_0)}\widehat{h}(\lambda)\left(\phi_1(t,\lambda;\sigma_0)\right)^n\,\mathrm{e}^{\mathtt{i}\lambda x}\\
&= \frac{1}{n!}\frac{1}{2\pi}\int_{\mathbb{R}}\mathrm{d}\lambda_r\,\mathrm{e}^{t\phi_0(\lambda;\sigma_0)}\widehat{h}(\lambda)\left(\phi_1^{(m)}(t,\lambda;\sigma_0)+\delta\phi_1^{(m)}(t,\lambda;\sigma_0)\right)^n\,\mathrm{e}^{\mathtt{i}\lambda x}\\
&= \frac{1}{n!}\left(\phi_1^{(m)}(t,-\mathtt{i}\partial_x;\sigma_0)\right)^n\frac{1}{2\pi}\int_{\mathbb{R}}\mathrm{d}\lambda_r\,\mathrm{e}^{t\phi_0(\lambda;\sigma_0)}\widehat{h}(\lambda)\mathrm{e}^{\mathtt{i}\lambda x} + \varepsilon_n^{(m)}(t,x,\sigma_0)\\
&= \frac{1}{n!}\left(\phi_1^{(m)}(t,-\mathtt{i}\partial_x;\sigma_0)\right)^n u_0(t,x,\sigma_0) + \varepsilon_n^{(m)}(\sigma_0) = u_n^{(m)}(t,x,\sigma_0) + \varepsilon_n^{(m)}(t,x,\sigma_0),
\end{aligned} \tag{12}$$

where

$$\begin{aligned}
u_n^{(m)}(t,x,\sigma_0) &:= \frac{1}{n!}\left(\phi_1^{(m)}(t,-\mathtt{i}\partial_x;\sigma_0)\right)^n u_0(t,x,\sigma_0),\\
\varepsilon_n^{(m)}(t,x,\sigma_0) &:= \frac{1}{n!}\sum_{k=1}^{n}\binom{n}{k}\left(\phi_1^{(m)}(t,-\mathtt{i}\partial_x;\sigma_0)\right)^{n-k}\left(\delta\phi_1^{(m)}(t,-\mathtt{i}\partial_x;\sigma_0)\right)^k u_0(t,x,\sigma_0).
\end{aligned}$$

From the decomposition (11), we can write $\phi_1^{(m)}(t,-\mathtt{i}\partial_x;\sigma_0) = \sum_{k=2}^{m}a_k(t;\sigma_0)(\partial_x^k - \partial_x)$, where the coefficients a_k are defined in (9). We can now compute

$$\begin{aligned}
u_n^{(m)}(t,x,\sigma_0) &= \frac{1}{n!}\left(\phi_1^{(m)}(t,-\mathtt{i}\partial_x;\sigma_0)\right)^n u_0(t,x,\sigma_0) = \frac{1}{n!}\left(\sum_{k=2}^{m}a_k(t;\sigma_0)(\partial_x^k-\partial_x)\right)^n u_0(t,x,\sigma_0)\\
&= \frac{1}{n!}\sum_{\substack{2\leq k_2,\ldots,k_m\leq k\\ k_2+\ldots+k_m=n}}\frac{n!}{k_2!\ldots k_m!}\prod_{2\leq j\leq m}a_j^{k_j}(t;\sigma_0)\left(\partial_x^j-\partial_x\right)^{k_j}u_0(t,x,\sigma_0).
\end{aligned}$$

Defining $\gamma_n^{(m)}$ as in the theorem, the result follows. Regarding $\varepsilon_n^{(m)}$, since for any $z \in \Gamma$, there exists $M > 0$ such that $|\phi_1(t,z;\sigma_0)/(z-\lambda)| \leq M$, we have $\left|\delta\phi_1^{(m)}(t,\lambda;\sigma_0)\right| \leq M\left(|\lambda|/R\right)^{m+1} + |\lambda|\left|\sum_{k=1}^{m}a_k\right|$, where R denotes the radius of convergence of ϕ_1. The sequence $(a_k)_{k\geq1}$ is (eventually) decreasing and the sum tends to zero as m tends to infinity, so that the sum can be made arbitrarily small. We then obtain

$$\begin{aligned}
\left|\varepsilon_n^{(m)}(t,x,\sigma_0)\right| &\leq \frac{1}{n!}\sum_{k=1}^{n}\binom{n}{k}\left|\phi_1^{(m)}(t,-\mathtt{i}\partial_x;\sigma_0)\right|^{n-k}\left(M\left(|\lambda|/R\right)^{m+1} + |\lambda|\left|\sum_{k=1}^{m}a_k\right|\right)^k u_0(t,x,\sigma_0)\\
&\leq \frac{1}{n!}\sum_{k=1}^{n}\binom{n}{k}\left|\sum_{j=2}^{m}a_j(t;\sigma_0)\left((\mathtt{i}\lambda)^j - \mathtt{i}\lambda\right)\right|^{n-k}\left(M\left(|\lambda|/R\right)^{m+1} + |\lambda|\left|\sum_{k=1}^{m}a_k\right|\right)^k u_0(t,x,\sigma_0).
\end{aligned}$$

One can then readily check that the sum behaves as $\mathcal{O}(\lambda^n)$ as λ tends to zero. Therefore, $\varepsilon_n^{(m)}$ contains derivatives (with respect to x) of u_0 of at least order n. $\square$

Remark 11. The key to the above proof is to note that $(\phi_1^{(m)}(t,\lambda;\sigma_0))^n$ is the symbol of the differential operator $(\phi_1^{(m)}(t,-\mathtt{i}\partial_x;\sigma_0))^n$ in the sense that $(\phi_1^{(m)}(t,-\mathtt{i}\partial_x;\sigma_0))^n \mathrm{e}^{\mathtt{i}\lambda x} = (\phi_1^{(m)}(t,\lambda;\sigma_0))^n \mathrm{e}^{\mathtt{i}\lambda x}$. The symbol of a differential operator appears naturally in connection with Fourier transforms as follows. If ϕ is the symbol of a differential operator $\mathcal{A}$ and f a function in the Schwartz class of rapidly decaying functions, then $\mathcal{A}f(x) = (2\pi)^{-1}\int_{\mathbb{R}} \mathrm{d}\lambda \phi(\lambda)\mathrm{e}^{\mathtt{i}\lambda x}\widehat{f}(\lambda)$.

The following proposition translates the results above for option prices into expansions for the implied volatility. In particular, it reviews the expansion (7) in light of Theorem 9. For $k \geq 1$, let us introduce the notation $\sigma_k^{(m)}$ as the m-th order approximation of σ_k, i.e. where each u_k in (7) is approximated by $u_k^{(m)}$ defined in Theorem 9. This in turn yields the (n,m)-th order approximation of the implied volatility as

$$\sigma^{(n,m)} := \sum_{j=0}^{n} \sigma_j^{(m)}. \tag{13}$$

Proposition 12. *For any $k \geq 1$ and $m \geq 2$, we have $\sigma_k = \sigma_k^{(m)} + \widehat{\varepsilon}_k^{(m)}$, where $\sigma_k^{(m)} := \Phi_k^{(m)}|_{\sigma=\sigma_0} + \Theta_k^{(m)}|_{\sigma=\sigma_0}$,*

$$\Phi_k^{(m)} := \sum_{\substack{2\leq k_2,\dots,k_m\leq k\\ k_2+\dots+k_m=k}} \prod_{2\leq j\leq m} \frac{a_j^{k_j}(t;\sigma_0)}{k_2!\dots k_m!} \frac{\left(\partial_x^j - \partial_x\right)^{k_j} u^{\mathrm{BS}}}{\partial_\sigma u^{\mathrm{BS}}}, \qquad \Theta_k^{(m)} := -\sum_{n\geq 2} \frac{1}{n!} \left(\sum_{j_1+\cdots+j_n=k} \prod_{i=1}^{n} \sigma_{j_i}^{(m)} \right) \frac{\partial_\sigma^n u^{\mathrm{BS}}}{\partial_\sigma u^{\mathrm{BS}}},$$

and $\widehat{\varepsilon}_k^{(m)} := \varepsilon_k^{(m)}(t,x,\sigma_0)/\partial_\sigma u^{\mathrm{BS}}(t,x,\sigma_0)$. Moreover,

$$\left.\frac{\left(\partial_x^j - \partial_x\right)^k u^{\mathrm{BS}}}{\partial_\sigma u^{\mathrm{BS}}}\right|_{\sigma=\sigma_0} = \frac{1}{t\sigma_0} \sum_{l=0}^{k-1} \sum_{i=2}^{j} \binom{k-1}{l} (-1)^{lj+i} \left(\sigma_0\sqrt{2t}\right)^{3-l(j-1)-k-i} H_{l(j-1)+k-1+i-2}(y), \tag{14}$$

$$\left.\frac{\partial_\sigma^n u^{\mathrm{BS}}}{\partial_\sigma u^{\mathrm{BS}}}\right|_{\sigma=\sigma_0} = \sum_{q=0}^{\lfloor n/2\rfloor} \sum_{p=0}^{n-q-1} \binom{n-q-1}{p} c_{n,n-2q}\sigma_0^{n-2q-1} t^{n-q-1} \left(\sigma_0\sqrt{2t}\right)^{1-p-n+q} H_{p+n-q-1}(y), \tag{15}$$

where $H_n(y) \equiv (-1)^n \mathrm{e}^{y^2}\partial_y^n \mathrm{e}^{-y^2}$ is the n-th Hermite polynomial, $y := \frac{1}{\sigma_0\sqrt{2t}}\left(x - k - \frac{1}{2}\sigma_0^2 t\right)$, and the coefficients $(c_{n,n-2k})$ are defined recursively by $c_{n,n} = 1$ and $c_{n,n-2q} = (n-2q+1)c_{n-1,n-2q+1} + c_{n-1,n-2q-1}$ for any integer $q \in \{1,2,\cdots,\lfloor n/2\rfloor\}$.

The proof of this proposition mainly relies on the following lemma.

Lemma 13. *Let $m \geq 0, n \geq 2$ be integers, $(t,k,\sigma) \in \mathbb{R}_+ \times \mathbb{R} \times (0,\infty)$ Then*

$$\frac{\partial_x^m(\partial_x^n - \partial_x)u^{\mathrm{BS}}(t,x,\sigma)}{(\partial_x^2 - \partial_x)u^{\mathrm{BS}}(t,x,\sigma)} = \sum_{i=2}^{n} \left(-\frac{1}{\sigma\sqrt{2t}}\right)^{m+i-2} H_{m+i-2}\left(\frac{x-k-\frac{1}{2}\sigma^2 t}{\sigma\sqrt{2t}}\right), \qquad \textit{holds for any } x \in \mathbb{R}.$$

Proof. From the Black-Scholes call price formula $u^{\mathrm{BS}}(t,x,\sigma) = \mathrm{e}^x\mathcal{N}(d_+(x)) - \mathrm{e}^k\mathcal{N}(d_-(x))$, with $d_\pm(x) := \frac{1}{\sigma\sqrt{t}}\left(x - k \pm \frac{1}{2}\sigma^2 t\right)$, where $\mathcal{N}$ is the CDF of a standard normal random variable, we immediately obtain

$(\partial_x^2-\partial_x)u^{\mathrm{BS}}(t,x,\sigma)=\frac{1}{\sigma\sqrt{t}}\exp\left(x-\frac{1}{2}d_+^2(x)\right)$. Now, for any integers $m\geq 0$ and $n\geq 2$, we have

$$\begin{aligned}\partial_x^m(\partial_x^n-\partial_x)u^{\mathrm{BS}}(t,x,\sigma)&=\partial_x^m\sum_{i=2}^{n}(\partial_x^i-\partial_x^{i-1})u^{\mathrm{BS}}(t,x,\sigma)=\sum_{i=2}^{n}\partial_x^{m+i-2}(\partial_x^2-\partial_x)u^{\mathrm{BS}}(t,x,\sigma)\\&=\frac{1}{\sigma\sqrt{t}}\sum_{i=2}^{n}\partial_x^{m+i-2}\exp\left(x-\frac{d_+^2(x)}{2}\right).\end{aligned}$$

Finally the lemma follows from the identity $x-\frac{1}{2}d_+^2(x)=-y^2+k$ and from

$$\begin{aligned}\frac{\partial_x^m(\partial_x^n-\partial_x)u^{\mathrm{BS}}(t,x,\sigma)}{(\partial_x^2-\partial_x)u^{\mathrm{BS}}(t,x,\sigma)}&=\frac{\sum_{i=2}^{n}\partial_x^{m+i-2}\exp\left(x-\frac{d_+^2(x)}{2}\right)}{\exp\left(x-\frac{d_+^2(x)}{2}\right)}\\&=\sum_{i=2}^{n}\mathrm{e}^{y^2}\left(\frac{1}{\sigma\sqrt{2t}}\right)^{m+i-2}\partial_y^{m+i-2}\left(\mathrm{e}^{-y^2}\right)=\sum_{i=2}^{n}\left(-\frac{1}{\sigma\sqrt{2t}}\right)^{m+i-2}H_{m+i-2}(y),\end{aligned}$$

where the Hermite polynomials H_i are recalled in Proposition 12. ☐

Proof of Proposition 12. The expression for σ_k as a sum of $\widetilde{\varepsilon}_k^{(m)}$, $\Phi_k^{(m)}$ and $\Theta_k^{(m)}$ is easily deduced by inserting (8) into (7). Thus, we must prove (14) and (15). Below, for clarity, we shall omit the arguments (t,x,σ) for the function u^{BS}. For any integers $j\geq 2$ and $k\geq 2$ we can write

$$\begin{aligned}(\partial_x^j-\partial_x)^k=(\partial_x^j-\partial_x)^{k-1}(\partial_x^j-\partial_x)&=\sum_{l=0}^{k-1}\binom{k-1}{l}(\partial_x^j)^l(-\partial_x)^{k-1-l}(\partial_x^j-\partial_x)\\&=\sum_{l=0}^{k-1}\binom{k-1}{l}(-1)^{k-1-l}\partial_x^{n(j-1)+k-1}(\partial_x^j-\partial_x).\end{aligned}$$

Using Lemma 13 and the identity $\partial_\sigma u^{\mathrm{BS}}=t\sigma(\partial_x^2-\partial_x)u^{\mathrm{BS}}$, we obtain

$$\begin{aligned}\frac{(\partial_x^j-\partial_x)^k u^{\mathrm{BS}}}{\partial_\sigma u^{\mathrm{BS}}}&=\frac{1}{t\sigma}\sum_{l=0}^{k-1}\binom{k-1}{l}(-1)^{k-1-l}\frac{\partial_x^{l(j-1)+k-1}(\partial_x^j-\partial_x)u^{\mathrm{BS}}}{(\partial_x^2-\partial_x)u^{\mathrm{BS}}}\\&=\frac{1}{t\sigma}\sum_{l=0}^{k-1}\sum_{i=2}^{j}\binom{k-1}{l}(-1)^{k-1-l}\left(-\frac{1}{\sigma\sqrt{2t}}\right)^{l(j-1)+k-1+i-2}H_{l(j-1)+k-1+i-2}(y).\end{aligned}\tag{16}$$

Now, define the operator $\mathcal{L}:=t(\partial_x^2-\partial_x)$ so that $\partial_\sigma u^{\mathrm{BS}}=\sigma\mathcal{L}u^{\mathrm{BS}}$. We now claim that the identity

$$\partial_\sigma^n u^{\mathrm{BS}}=\sum_{q=0}^{\lfloor n/2\rfloor}c_{n,n-2q}\sigma^{n-2q}\mathcal{L}^{n-q}u^{\mathrm{BS}},$$

holds for any $n\in\mathbb{N}$, where $c_{n,n}=1$ and $c_{n,n-2q}=(n-2q+1)c_{n-1,n-2q+1}+c_{n-1,n-2q-1}$ for any integer $q\in\{1,2,\cdots,\lfloor n/2\rfloor\}$. Its proof is a simple yet tedious recursion and we omit it for brevity. We obtain

$$\begin{aligned}\frac{\partial_\sigma^n u}{\partial_\sigma u^{\mathrm{BS}}}&=\sum_{q=0}^{\lfloor n/2\rfloor}c_{n,n-2q}\sigma^{n-2q}\frac{\mathcal{L}^{n-q}u}{\partial_\sigma u^{\mathrm{BS}}}=\sum_{q=0}^{\lfloor n/2\rfloor}c_{n,n-2q}\sigma^{n-2q}t^{n-q}\frac{(\partial_x^2-\partial_x)^{n-q}u}{\partial_\sigma u^{\mathrm{BS}}}\\&=\sum_{q=0}^{\lfloor n/2\rfloor}\sum_{p=0}^{n-q-1}c_{n,n-2q}\sigma^{n-2q-1}t^{n-q-1}\binom{n-q-1}{p}(-1)^{n-q-1-p}\left(-\frac{1}{\sigma\sqrt{2t}}\right)^{p+n-q-1}H_{p+n-q-1}(y),\end{aligned}$$

where we have used (16) with $j=2$ and $k=n-q$, and Proposition 12 follows. ☐

Remark 14. This proposition implies that computing each element of the sequence $(\sigma_k^{(m)})_{k\geq 0}$ is a simple algebraic exercise, which can easily be accomplished using a computer algebra system such as Mathematica. For example, with $m=4$, the first two terms read as follows,

$$\sigma_1^{(4)} = \frac{1}{4\sigma_0^5 t^3}\left(4(k-x)^2 a_4 + 4\left[-a_4 + (k-x)\left(a_3+2a_4\right)\right]\sigma_0^2 t + \left(4a_2+6a_3+7a_4\right)\sigma_0^4 t^2\right), \tag{17}$$

$$\begin{aligned}\sigma_2^{(4)} = &-\frac{1}{32\sigma_0^{11}t^6}\Big\{208(k-x)^4 a_4^2 + 32(k-x)^2 a_4\Big[9(k-x)(a_3+2a_4) - 22a_4\Big]\sigma_0^2 t\\
&+ (\sigma_0^2 t)^2\Big[12(k-x)^2 a_3^2 + 6(13(k-x)-10)(k-x)a_3 a_4\\
&\quad + a_4\Big(30a_4 + (k-x)\left((k-x)\left(20a_2+81a_4\right) - 120a_4\right)\Big)\Big]\\
&+ (\sigma_0^2 t)^3\Big[6(3(k-x)-1)a_3^2 + (55(k-x)-42)a_3 a_4\\
&\quad + 2(19(k-x)-22)a_4^2 + 12a_2\Big((k-x)\left(a_3+2a_4\right) - a_4\Big)\Big]\\
&+ (\sigma_0^2 t)^4\Big[\Big(4a_2+4a_3+3(a_4+a_5)\Big)\Big(4a_2+8a_3+11(a_4+a_5)\Big)\Big]\Big\},\end{aligned}$$

where we have used the short-hand notation $a_i = a_i(t;\sigma_0)$. Note that each $\sigma_k^{(m)}$ is expressed as a polynomial in log-moneyness: $(k-x)$. As such, approximate implied volatilities can be computed extremely quickly (in particular faster than option prices, which require a Fourier integration).

3.4 Extension to forward implied volatility

Let $(\mathrm{e}^{X_t})_{t\geq 0}$ be a true martingale. A forward-start call option with forward-start date $t>0$, maturity $t+\tau$ ($\tau>0$), and strike e^k ($k\in\mathbb{R}$) is worth $\mathbb{E}_x\left(\mathrm{e}^{X_\tau^{(t)}} - \mathrm{e}^k\right)^+$ at inception of the contract, where the forward process $X^{(t)}$ is defined pathwise as $X_\tau^{(t)} := X_{t+\tau} - X_t$ for all $\tau\geq 0$. Of course, knowledge of the characteristic function of $X_\tau^{(t)}$ is sufficient to compute the expectation. By independence of increments in the Black-Scholes model, such a forward call option has the same value as a standard vanilla call option, with initial underlying value $x=0$, strike k and maturity τ. We can hence define the *forward implied volatility* analogously to the standard (spot) implied volatility, namely as the unique strictly positive solution σ to the equation $u^{(t)}(\tau) = u^{\mathrm{BS}}(\tau,0,\sigma)$, where $u^{(t)}$ is the option price to be matched. Consider now an affine stochastic volatility model (X,Y) characterised by $\log\mathbb{E}_{x,y}\left(\mathrm{e}^{\mathtt{i}\lambda X_t + \mathtt{i}\mu Y_t}\right) = \mathtt{i}\lambda x + C(t,\lambda,\mu) + y\,D(t,\lambda,\mu)$, for all $(\lambda,\mu)\in\mathbb{C}^2$ such that the expectation exists. We can then compute $\phi^{(t)}$ explicitly as (see also [23])

$$\begin{aligned}\phi^{(t)}(\tau,\lambda;y) &:= \log\mathbb{E}_{x,y}\left(\mathrm{e}^{\mathtt{i}\lambda X_\tau^{(t)}}\right) = \log\mathbb{E}_{x,y}\left(\mathrm{e}^{\mathtt{i}\lambda(X_{t+\tau}-X_t)}\right)\\
&= \log\mathbb{E}_{x,y}\left(\mathrm{e}^{-\mathtt{i}\lambda X_t}\mathbb{E}\left[\mathrm{e}^{\mathtt{i}\lambda X_{t+\tau}}|X_t,Y_t\right]\right) = \log\mathbb{E}_{x,y}\left(\mathrm{e}^{-\mathtt{i}\lambda X_t + \mathtt{i}\lambda X_t + C(\tau,\lambda,0) + Y_t\,D(\tau,\lambda,0)}\right)\\
&= C(\tau,\lambda,0) + C(t,0,-\mathtt{i}D(\tau,\lambda,0)) + yD(t,0,-\mathtt{i}D(\tau,\lambda,0)).\end{aligned}$$

The price of a forward start call option can then be expressed as an inverse Fourier transform, just as in the European call option case: $u^{(t)}(\tau, y) = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{h}(\lambda)e^{\phi^{(t)}(\tau,\lambda;y)}d\lambda_r$, and we are back to the general case with $x = 0$. As such, we are able to compute forward implied volatilities using the same techniques described above for (spot) implied volatility.

4 Numerical implementation: discussions and examples

We now focus on the practical implementation of the results above, namely Proposition 12. Section 4.1 discusses the (arbitrary) choice of the level σ_0 and Section 4.2 proposes a smoothing procedure to further enhance the applicability of our methodology. In Sections 4.3 and 4.4, we implement our implied volatility expansion in two exponential Lévy models (Merton and Variance Gamma) and one stochastic volatility model (Heston). For the Heston model, we also examine how our expansion performs in approximating the forward implied volatility smile.

4.1 Optimal choice of σ_0

Note that the analysis above was for a fixed (t, k) and the choice of σ_0 was arbitrary. When computing implied volatilities over a range of strikes and maturities, one can choose σ_0 to be a function of time to maturity and log strike: $\sigma_0 = \sigma_0(t, k)$. Define $a_i(t; \sigma_0) := \alpha_i(t) + \alpha_i^{\mathrm{BS}}(t; \sigma_0)$, where

$$\alpha_i^{\mathrm{BS}}(t; \sigma_0) := -\frac{(-\mathtt{i})^i t}{i!}\,\partial_\lambda^i \phi_0(\lambda; \sigma_0)\big|_{\lambda=0} \qquad \text{and} \qquad \alpha_i(t) := \frac{(-\mathtt{i})^i}{i!}\,\partial_\lambda^i \phi(t, \lambda)\big|_{\lambda=0}.$$

Note that α_k does not depend on σ_0. Straightforward computations yield

$$\alpha_1^{\mathrm{BS}}(t; \sigma_0) = \frac{\sigma_0^2 t}{2}, \qquad \alpha_2^{\mathrm{BS}}(t; \sigma_0) = -\frac{\sigma_0^2 t}{2} \qquad \text{and} \qquad \alpha_i^{\mathrm{BS}}(t; \sigma_0) = 0, \qquad \text{for any } i \geq 3.$$

This, together with (17), implies

$$\begin{aligned} \sigma_1^{(4)} &= \frac{1}{4t^3\sigma_0^5}\Big(4(k-x)^2 a_4 + 4[-a_4 + (k-x)(a_3 + 2a_4)]\sigma_0^2 t + (4a_2 + 6a_3 + 7a_4)\sigma_0^4 t^2\Big) \\ &= -\frac{\sigma_0}{2} + \frac{1}{4\sigma_0^5 t^3}\Big(4(k-x)^2\alpha_4 + 4[-\alpha_4 + (k-x)(\alpha_3 + 2\alpha_4)]\sigma_0^2 t + (4\alpha_2 + 6\alpha_3 + 7\alpha_4)\sigma_0^4 t^2\Big), \end{aligned}$$

so that $\sigma_1^{(4)} + \frac{\sigma_0}{2}$ is asymptotically equivalent to $(k-x)^2\alpha_4/(\sigma_0^5 t^3)$ as σ_0 tends to zero. Since α_i does not depend either on σ_0 or on (k, x), the only way to prevent this explosive behaviour is to consider σ_0 as a function of (k, x) such that $(k-x)^2/\sigma_0^5$ converges to a constant (possibly zero).

4.2 Smoothing with the SVI parameterisation

Option data is often noisy and limited by the number of strikes at which options are liquidly traded. In [14], Jim Gatheral introduces the following *Stochastic Volatility Inspired* (SVI) parameterisation:

$$\sigma_t^{\mathrm{SVI}}(k) = \left\{ \frac{a}{t} + \frac{b}{t}\left(\rho(k - x - m) + \sqrt{(k - x - m)^2 + \xi^2}\right)\right\}^{1/2}, \tag{18}$$

for any maturity $t > 0$, where $a, b \geq 0$, $\xi > 0$, $m \in \mathbb{R}$, $\rho \in [-1, 1]$. By fitting the SVI parameterisation to noisy option data, one is able to create a smooth implied volatility smile, which then can be used to interpolate implied volatility between strikes and extrapolate implied volatility to strikes which are not traded. The density $p^\sigma(t, x)$ corresponding to a given implied volatility parameterisation $\sigma(t, k)$ can be computed via the Breeden-Litzenberger formula [4]: $p^\sigma(t, k) = \partial_K^2 u^{\mathrm{BS}}(t, x, \sigma(t, \log K); \log K)\big|_{K=\mathrm{e}^k}$. An implied volatility smile $k \mapsto \sigma(t, k)$ is said to be free of *butterfly arbitrage* if the corresponding density is non-negative: $p(t, \cdot) \geq 0$. Let $p_t^{\mathrm{SVI}}(k)$ be the implied volatility smile corresponding to a given SVI parameterisation (18). In general, SVI parameterisation (18) is not guaranteed to be free of butterfly arbitrage. However, for a given set of SVI parameters (a, b, ρ, m, σ), one can easily verify that the corresponding density is non-negative, and therefore free of butterfly arbitrage. This and recent arbitrage-free SVI parameterisations have recently been studied in [16] and [18], and we refer the interested reader to these papers for more details.

As we shall see in the examples considered in Section 4, for finite (n, m), the approximate implied volatility $\sigma^{(n,m)}$ derived in Section 2.2 has a tendency to oscillate around the true implied volatility (see Figures 1, 2 and 4). Taking $\sigma^{(n,m)}$ to be the true implied volatility could lead to arbitrage opportunities. In order to prevent this, we propose to smooth the implied volatility approximation $\sigma^{(n,m)}$ by fitting the SVI parameterisation to it. That is, given a model for the underlying X and a time to maturity t, we first compute the approximate implied volatility $\sigma^{(n,m)}$ as a function of log-strike k, and then fit an arbitrage-free SVI parameterisation σ_t^{SVI} to $\sigma^{(n,m)}$ over some range of strikes, usually chosen to be a symmetric interval around $k = x$.

4.3 Exponential Lévy models

Suppose that X is a Lévy process with Lévy triplet (μ, a^2, ν). Then its characteristic function reads

$$\phi(t, \lambda) = t\left(\mathtt{i}\mu\lambda + \frac{1}{2}a^2(\mathtt{i}\lambda)^2 + \int_{\mathbb{R}} \nu(\mathrm{d}z)(\mathrm{e}^{\mathtt{i}\lambda z} - 1 - \mathbb{1}_{\{|z|<1\}}\mathtt{i}\lambda z)\right),$$

where the drift μ is constrained by the martingale condition $\phi(t, -\mathtt{i}) = 0$: $\mu = -\frac{1}{2}a^2 - \int_{\mathbb{R}} \nu(\mathrm{d}z)(\mathrm{e}^z - 1 - \mathbb{1}_{\{|z|<1\}}z)$. From the expansion $(\mathrm{e}^{\mathtt{i}\lambda z} - 1) = \sum_{n\geq 1} \frac{1}{n!}(\mathtt{i}\lambda z)^n$ we can write

$$\phi_1(t, \lambda; \sigma_0) \equiv \phi(t, \lambda) - t\,\phi_0(\lambda; \sigma_0) \equiv t\left(\mu + I_1 + \tfrac{1}{2}\sigma_0^2\right)\mathtt{i}\lambda + \frac{1}{2}t(a^2 - \sigma_0^2)(\mathtt{i}\lambda)^2 + t\sum_{n=2}^{\infty} \frac{1}{n!} I_n(\mathtt{i}\lambda)^n,$$

where $I_1 := \int_{|z|\geq 1} \nu(\mathrm{d}z)z$ and $I_n := \int_{\mathbb{R}} \nu(\mathrm{d}z)z^n$, for any $n \geq 2$. The existence of I_n is equivalent to the finiteness of the nth moment of X by [31, Theorem 25.3], which is clearly satisfied under Assumption 1. Hence, the coefficients $a_n(t;\sigma_0)$ in (9) are given by

$$a_2(t;\sigma_0) = \frac{t}{2}(a^2 - \sigma_0^2 + I_2), \qquad a_n(t) = \frac{t}{n!}I_n, \quad n \geq 3. \tag{19}$$

We examine two exponential Lévy models in detail—the Merton model [30] and the Variance Gamma model [6]—whose Lévy measures are given by:

$$\text{Merton}: \qquad \nu(\mathrm{d}z) = \frac{\alpha}{\sqrt{2\pi s^2}} \exp\left(\frac{-(z-m)^2}{2s^2}\right)\mathrm{d}z,$$

$$\text{Variance Gamma}: \qquad \nu(\mathrm{d}z) = \alpha\left(\frac{\mathrm{e}^{Gz}}{-z}\mathbb{1}_{\{z<0\}} + \frac{\mathrm{e}^{-Mz}}{z}\mathbb{1}_{\{z>0\}}\right)\mathrm{d}z,$$

where $\alpha, s, G, M > 0$, $m \in \mathbb{R}$. The Merton model is a finite-activity Lévy process ($\nu(\mathbb{R}) < \infty$), whereas the Variance Gamma model has infinite activity ($\nu(\mathbb{R}) = \infty$). For infinite activity Lévy processes, one typically takes the diffusion component to be zero, namely $a = 0$. We now examine the accuracy of the implied volatility expansion above in these models: the Merton model in Figure 1 and the Variance Gamma in Figure 2. For each of these two sets of plots, we fix some parameters, and draw the implied volatility approximations $\sigma^{(n,m)}$ with $m = 7$ for the Merton model and $m = 8$ for the Variance-Gamma one, and $n \in \{1,2,3\}$. We also plot the SVI smoothing of $\sigma^{(3,m)}$ as well as the true implied volatility. The true option price is computed by a quadrature of the inverse Fourier transform representation (1), and the true implied volatility is computed by numerical inversion of the Black-Scholes formula (we use a simple Newton-Raphson algorithm). We also plot the total errors between each approximation (and $\sigma^{(3,\cdot)}$ with SVI smoothing) and the true implied volatility. As discussed above, the implied volatility approximation oscillates around the true implied volatility σ. However, the relative error corresponding to $\sigma^{(3,m)}$ is less than one percent for nearly all log-moneyness to maturity ratios (LMMRs) satisfying $|k-x|/t < 1.4$ for both models, which is well within the implied volatility bid-ask spread of S&P 500 options. Furthermore, the relative error of $\sigma^{(3,m)}$ with SVI smoothing is about one half percent for all $|k-x|/t < 1.0$. As a no-arbitrage consistency check, we also plot the density corresponding to the SVI fit. The parameters for each model are

$$\text{Merton model:} \qquad \sigma_0 = 0.55, a = 0.25, m = -0.15, s = 0.3, \alpha = 1.5, t = 1, x = 0,$$

$$\text{Variance Gamma model:} \qquad \sigma_0 = 0.55, a = 0, M = 7, G = 6, \alpha = 4.5, t = 1, x = 0.$$

In Figure 3 we examine the effect of σ_0 on the approximation proved in Proposition 12 by plotting $\sigma^{(3,8)}$ (in the Variance Gamma model) for different values of σ_0. We observe here that choosing σ_0 below the true implied volatility results in strong oscillations of the implied volatility expansion. Furthermore, the larger one chooses σ_0, the more these oscillations are dampened. This behaviour is not unexpected. Indeed, from Proposition 12 we see that each $\sigma_k^{(m)}$ contains a term $u_k^{(m)}(t,x,\sigma_0)/\partial_\sigma u^{\mathrm{BS}}(t,x,\sigma_0)$ and from (12) we see that $u_k^{(m)}(t,x,\sigma_0)$ can be expressed an integral with respect to λ_r which contains a Gaussian factor $\mathrm{e}^{t\phi_0(\lambda;\sigma_0)} \sim \mathrm{e}^{-t\frac{1}{2}\sigma_0^2\lambda_r^2}$ and a factor $\widehat{h}(\lambda)\left(\phi_1^{(m)}(t,\lambda;\sigma_0)\right)^l \mathrm{e}^{\mathrm{i}\lambda x}$, which grows in absolute value as $\lambda_r \to \pm\infty$. The growth of the non-Gaussian factor depends strongly on l. Choosing a larger σ_0 results in the Gaussian factor decaying more quickly, thus dampening the large $|\lambda_r|$ behaviour of the non-Gaussian factor. As a result, for large values of σ_0, oscillations in the implied volatility approximation are reduced.

4.4 The Heston model

In the Heston model [21], the risk-neutral dynamics of (X,Y) are given by

$$\mathrm{d}X_t = -\frac{1}{2}Y_t\mathrm{d}t + \sqrt{Y_t}\mathrm{d}W_t, \qquad \mathrm{d}Y_t = \kappa(\theta - Y_t)\mathrm{d}t + \delta\sqrt{Y_t}\mathrm{d}B_t, \qquad \mathrm{d}\langle W,B\rangle_t = \rho\,\mathrm{d}t,$$

with $(X_0,Y_0) = (x,y) \in \mathbb{R}\times(0,\infty)$, $\kappa,\theta,\delta > 0$, and where W and B are two standard Brownian motions with correlation $\rho \in [-1,1]$. Its characteristic function reads $\phi(t,\lambda,y) = C(t,\lambda) + yD(t,\lambda)$, where

$$\begin{aligned}
C(t,\lambda) &:= \frac{\kappa\theta}{\delta^2}\left((\kappa - \mathtt{i}\rho\delta\lambda - d(\lambda))t - 2\log\left[\frac{1-\gamma(\lambda)\mathrm{e}^{-d(\lambda)t}}{1-\gamma(\lambda)}\right]\right), \quad & d(\lambda) &:= \left(\delta^2(\lambda^2 + \mathtt{i}\lambda) + (\kappa - \rho\mathtt{i}\lambda\delta)^2\right)^{1/2},\\
D(t,\lambda) &:= \frac{\kappa - \mathtt{i}\rho\delta\lambda - d(\lambda)}{\delta^2}\frac{1-\mathrm{e}^{-d(\lambda)t}}{1-\gamma(\lambda)\mathrm{e}^{-d(\lambda)t}}, & \gamma(\lambda) &:= \frac{\kappa - \mathtt{i}\rho\delta\lambda - d(\lambda)}{\kappa - \mathtt{i}\rho\delta\lambda + d(\lambda)}.
\end{aligned}$$

Unlike the exponential Lévy setting, there is no simple general formula for the coefficients $a_n(t,\sigma_0)$ $(n \geq 2)$. However, from (9), one can compute

$$\begin{aligned}
a_2(t;\sigma_0) =& \frac{\mathrm{e}^{2\kappa t}}{16\kappa^3}\left(4\mathrm{e}^{\kappa t}\left[2(\theta-y)\kappa^2 + 2(y + y\kappa t - \theta(2+\kappa t))\kappa\rho\delta + (\theta + (\theta-y)\kappa t)\delta^2\right] - (2y-\theta)\delta^2\right)\\
&+ \frac{1}{16\kappa^3}\left(8\kappa^2(y + (\kappa t - 1)\theta) - 8(y + \theta(\kappa t - 2))\kappa\rho\delta - ((5 - 2\kappa t)\theta - 2y)\delta^2\right) - \frac{\sigma_0^2 t}{2}.
\end{aligned}$$

Higher order terms $(3 \leq n \leq 6)$ are easily computed using any mathematical software, and are omitted here for clarity. In Figure 4, we plot the function $k \mapsto \sigma_n^{(m)}(k)$ with $m = 6$ and $n \in \{1,2,3\}$, a calibrated SVI to $\sigma_3^{(6)}$ and the true implied volatility (computed exactly as for the Lévy models above). We also plot the relative errors between each approximation (and the SVI smoothing of $\sigma^{(3,6)}$) and the true implied volatility. Again the approximation $\sigma^{(n,m)}$ oscillates around the true implied volatility, but the relative error of $\sigma^{(3,6)}$ is less than two percent for nearly all LMMRs satisfying $|k-x|/t < 2.0$, and that of $\sigma^{(3,6)}$ with SVI smoothing is roughly one percent for all $|k-x|/t < 2.0$. As before, we also plot the density corresponding to the calibrated SVI parameterisation as a no-arbitrage consistency check. We use the following set of parameters: $\sigma_0 = 0.95$, $\kappa = 1$, $\theta = 0.3$, $\delta = 0.7$, $\rho = -0.3$, $t = 1$, $x = 0$, $y = 0.5$.

For the forward implied volatility in the Heston model, we can compute

$$\phi^{(t)}(\tau, \lambda; y) = C(\lambda, \tau) + \frac{D(\lambda, \tau) y \mathrm{e}^{-\kappa t}}{1 - 2\beta_t D(\lambda, \tau)} - \frac{2\kappa\theta}{\delta^2} \log(1 - 2\beta_t D(\lambda, \tau)),$$

with $\beta_t := \frac{\delta^2}{4\kappa}(1 - \mathrm{e}^{-\kappa t})$. Explicit expressions for $a_n^{(t)}(\tau; \sigma_0)$ $(2 \le n \le 6)$ can be computed easily, and we omit their lengthy representations here, and the set of plots in Figure 5 are the analogues of those for the Heston model, but applied to the forward case. We use here the following values for the parameters: $y = 0.5$, $\sigma_0 = 0.8$, $\kappa = 1$, $\theta = 0.3$, $\delta = 0.3$, $\rho = -0.3$, $t = 1$, $\tau = 1$.

4.5 Model-free calibration

As noted previously, the model-specific dependence of the approximate implied volatility expansion $\sigma^{(n,m)}$ is entirely captured by the $a_i(t, \sigma_0)$ $(2 \le i \le m)$. This simple structure allows for a model-free calibration of the implied volatility surface. Assume one observes implied volatilities for maturities $t(_i)_{i=1,\dots,n_T}$ and $(k_j)_{j=1,\dots,n_K}$, where n_T and n_K are two integers. We shall assume for simplicity that the number of available strikes is the same for each maturity. We suggest the following procedure:

(i) Let $\sigma_{i,j} := \sigma(t_i, k_j)$ be the quoted implied volatility for an option with maturity t_i and log strike k_j.

(ii) Let $\sigma_{i,j}^{(n,m)} := \sigma^{(n,m)}(t_i, k_j; \sigma_0)$ be the approximate implied volatility for an option with maturity t_i and log strike k_j computed using the approximation (13).

(iii) At each maturity t_i, leave σ_0 and $a_q(t_i; \sigma_0)$ $(2 \le q \le m)$ as free parameters. Fit $\sigma^{(n,m)}(t_i, \cdot)$ to the market's t_i-maturity implied volatility smile $\sigma(t_i, \cdot)$ by minimising $\sum_j \left|\sigma^{(n,m)}(t_i, k_j) - \sigma_{ij}^{(n,m)}\right|^2$.

(iv) As an initial guess, use the largest quoted implied volatility at each maturity for σ_0, and $a_q(t_i, \sigma_0) = 0$.

Remark 15. With $n = 3$ and $m = 8$, step (iii) is instantaneous using Mathematica's `FitTo` or Matlab's `lsqnonlin` for instance.

We test this procedure on SPX index options from January 4, 2010 with $n = 3$ and $m = 8$. The results for three separate maturities ($t = 0.033$, $t = 0.70$, $t = 1.45$ years) are given on Figure 6. The calibrated parameters are (a_i is a shorthand notation for $a_i(t; \sigma_0)$):

	σ_0	a_2	a_3	a_4	a_5	a_6	a_7	a_8
t =0.033	0.382	-1.64E-3	-1.00E-6	1.64E-6	1.89E-8	8.62E-10	5.22E-12	1.40E-13
t =0.70	0.659	-1.31E-1	-5.00E-3	2.40E-3	4.24E-4	6.96E-5	2.88E-6	4.83E-7
t =1.45	0.436	-8.35E-2	-1.22E-2	1.68E-3	-6.26E-5	3.42E-5	-1.77E-6	3.96E-7

Remark 16. If the stock price is an exponential Lévy model, then (19) implies that $\frac{1}{2}\sigma_0^2 + \frac{1}{t_i}a_2(t_i, \sigma_0) = \frac{1}{2}(a^2 + I_2)$ and $\frac{1}{t_i}a_q(t_i) = \frac{1}{q!}I_q$ $(3 \leq q \leq m)$ should be constant. If this is not so, then exponential Lévy models are probably not the best dynamics to describe the underlying.

Acknowledgments

M. Lorig acknowledges financial support from Imperial College London, and both authors are grateful to Claude Martini and Zeliade Systems for their useful comments and for indicating to us the availability of CBOE data.

References

[1] Abramowitz, M. and I. Stegun (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, Volume 55. Dover publications.

[2] Benaim, S. and P. Friz (2009). Regular variation and smile asymptotics. *Math. Finance 19(1)*, 1–12.

[3] Berestycki, H., J. Busca, and I. Florent (2004). Computing the implied volatility in stochastic volatility models. *Communications on Pure and Applied Mathematics 57 (10)*, 1352–1373.

[4] Breeden, D. T. and R. H. Litzenberger (1978). Prices of state-contingent claims implicit in option prices. *The Journal of Business 51*(4), 621–651.

[5] Brown, J. and R. Churchill (1996). *Complex variables and applications*, Volume 7. McGraw-Hill New York, NY.

[6] Carr, P., D. Madan, and E. Chang (1998). The variance gamma process and option pricing. *European Finance Review 2*(1), 79–105.

[7] Deuschel, J., P. Friz, A. Jacquier, and S. Violante (2013). Marginal density expansions for diffusions and stochastic volatility, part ii: Applications. *To appear in Comm. Pure Applied Math.*.

[8] Duffie, D., D. Filipović, and W. Schachermayer (2003). Affine processes and applications in finance. *Annals of Applied Probability 13 (3)*, 984–1053.

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

[10] Fouque, J.-P., G. Papanicolaou, R. Sircar, and K. Sølna (2003). Singular perturbations in option pricing. *SIAM J. Applied Mathematics 63*(5), 1648–1665.

[11] Fouque, J.-P., G. Papanicolaou, R. Sircar, and K. Solna (2011). *Multiscale Stochastic Volatility for Equity, Interest-Rate and Credit Derivatives.* Cambridge University Press.

[12] Friz, P., S. Gerhold, A. Gulisashvili, and S. Sturm (2011). Refined implied volatility expansions in the heston model. *Quantitative Finance 11 (8)*, 1151–1164.

[13] Fukasawa, M. (2011). Asymptotic analysis for stochastic volatility: Edgeworth expansion. *Electronic Journal of Probability 16*, 764–791.

[14] Gatheral, J. (2004). A parsimonious arbitrage-free implied volatility parameterization with application to the valuation of volatility derivatives. *Presentation at Global Derivatives & Risk Management, Madrid*.

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

[16] Gatheral, J. and A. Jacquier (2013). Arbitrage-free svi volatility surfaces. *To appear in Quant. Finance*.

[17] Gulisahsvili, A. and E. Stein (2010). Asymptotic behavior of the stock price distribution density and implied volatility in stochastic volatility models. *Applied Mathematics & Optimization 61 (3)*, 287–315.

[18] Guo, G., A. Jacquier, C. Martini, and L. Neufcourt (2012). Generalised arbitrage-free svi volatility surfaces. *arXiv:1210.7111*.

[19] Hagan, P., D. Kumar, A. Lesniewski, and D. Woodward (2002). Managing smile risk. *Wilmott Magazine*, 84–108.

[20] Henry-Labordère, P. (2008). *Analysis, geometry and modeling in finance.* Chapman and Hill / CRC.

[21] Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Stud. 6*(2), 327–343.

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

[23] Jacquier, A. and P. Roome (2012). Asymptotics of forward implied volatility. *arxiv 1212.0779*.

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

[25] Lewis, A. (2001). A simple option formula for general jump-diffusion and other exponential Lévy processes.

[26] Lipton, A. (2002). The vol smile problem. *Risk* (February), 61–65.

[27] Lorig, M. (2012). The exact smile of some local volatility models. *arXiv:1207.0750*.

[28] Lorig, M., S. Pagliarani, and A. Pascucci (2013). Implied vol for any local-stochastic vol model. *arXiv:1306.5447v1*.

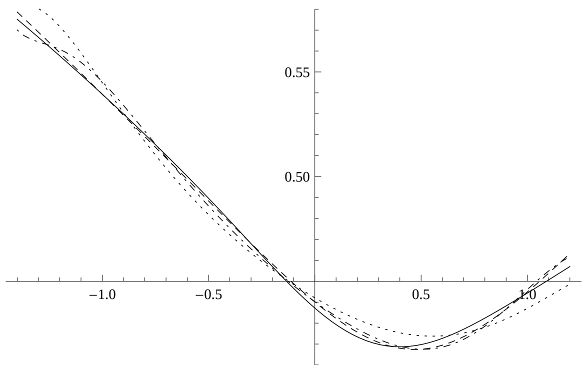
[29] Lukacs, E. (1970). *Characteristic functions (2nd edition)*. Hafner Pub. Co. New York, NY.

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

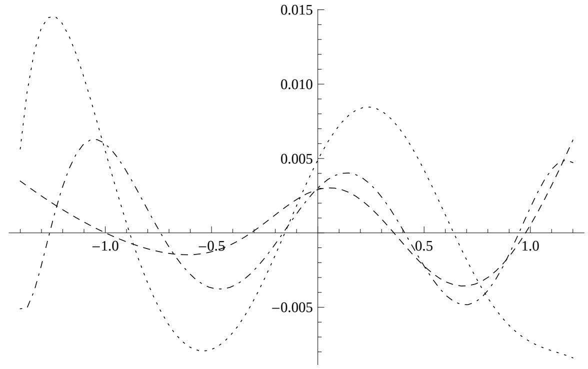
[31] Sato, K. (1999). *Lévy processes and infinitely divisible distributions.* Cambridge University Press.

[32] Takahashi, A. and M. Toda (2013). Note on an extension of an asymptotic expansion scheme. *To appear in International Journal of Theoretical and Applied Finance*.

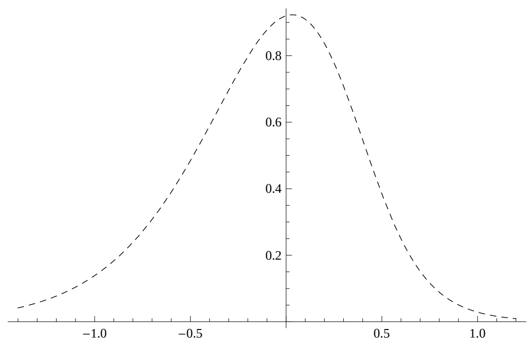
[33] Tankov, P. (2010). Pricing and hedging in exponential lévy models: review of recent results. In *Paris-Princeton Lecture Notes in Mathematical Finance.*



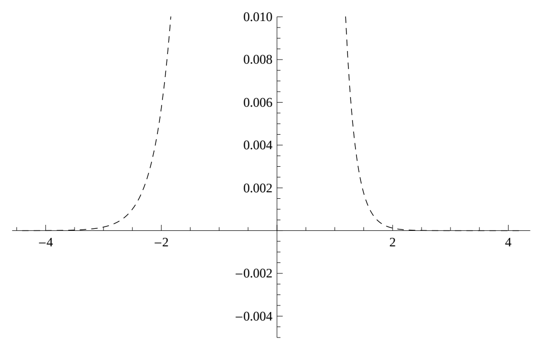
(a) Implied volatilities



(b) Implied volatility errors

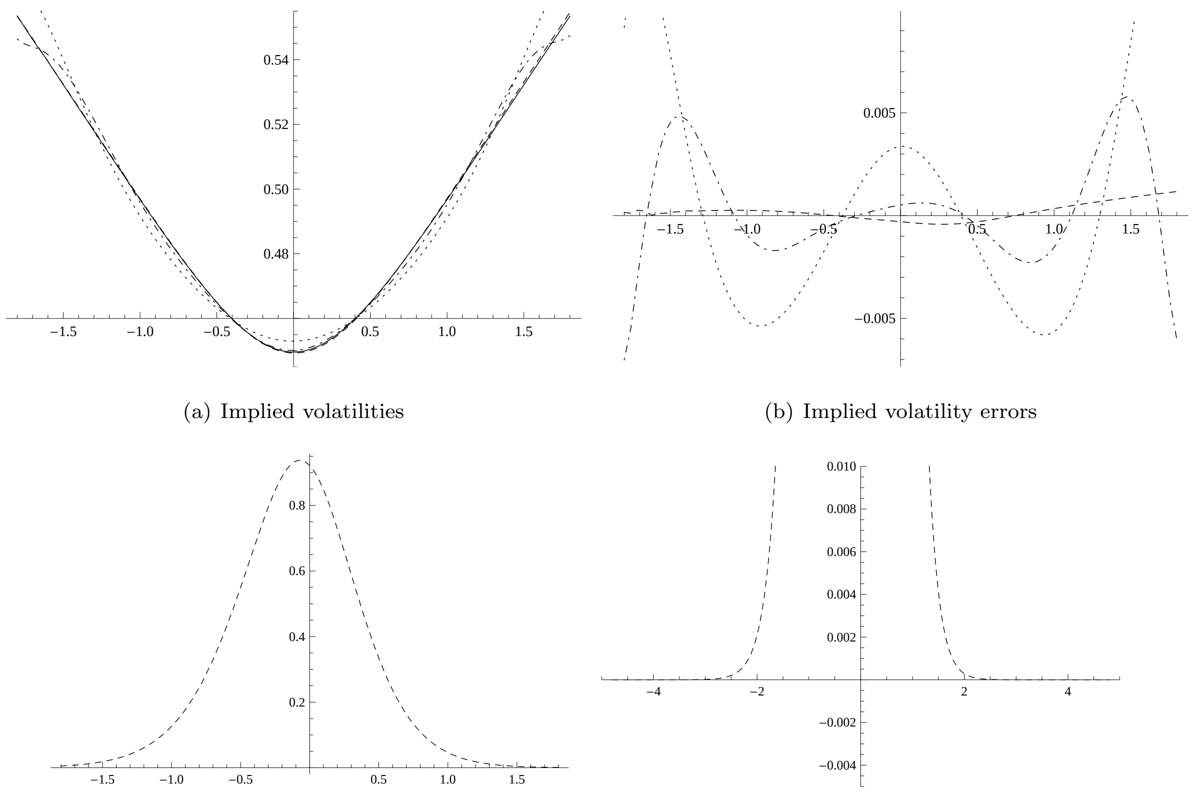


(c) Density



(d) Tails of the density

Figure 1: Numerics for the Merton model detailed in Section 4.3. In the top graphs, the solid line corresponds to the true implied volatility, whereas the others are the different approximations: $\sigma^{(2,7)}$ dots, $\sigma^{(3,7)}$ dots-dashed, and the SVI smoothing of $\sigma^{(3,7)}$ dashed. The two plots below are the density (and its tails) corresponding to the SVI smoothing.



(a) Implied volatilities

(b) Implied volatility errors

(c) Density

(d) Tails of the density

Figure 2: Numerics for the Variance Gamma model detailed in Section 4.3. In the top graphs, the solid line corresponds to the true implied volatility, whereas the others are the approximations $\sigma^{(2,8)}$ dots, $\sigma^{(3,8)}$ dots-dashed, and the SVI smoothing of $\sigma^{(3,8)}$ dashed. The two plots below are the density (and its tails) corresponding to the SVI smoothing.

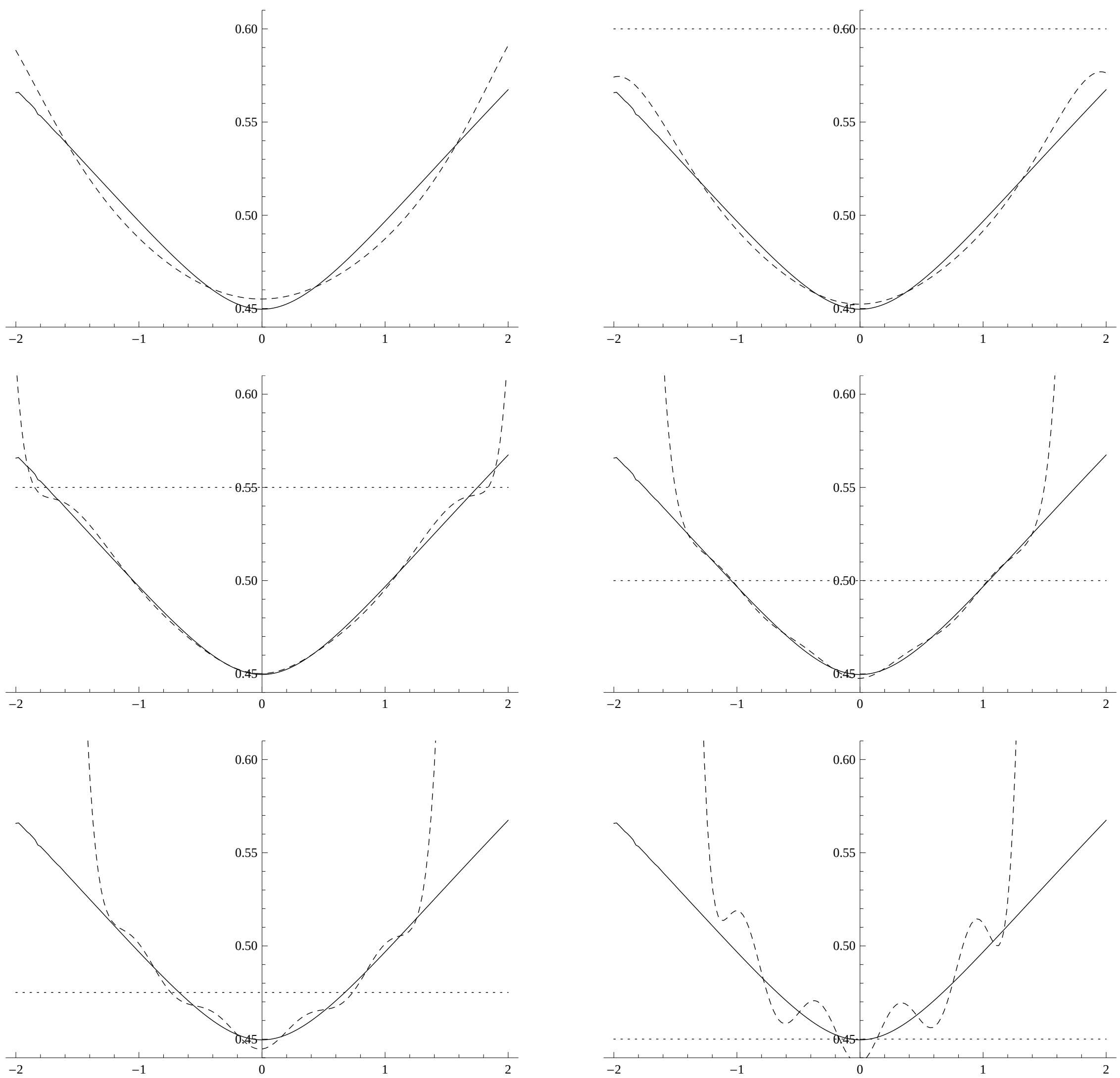
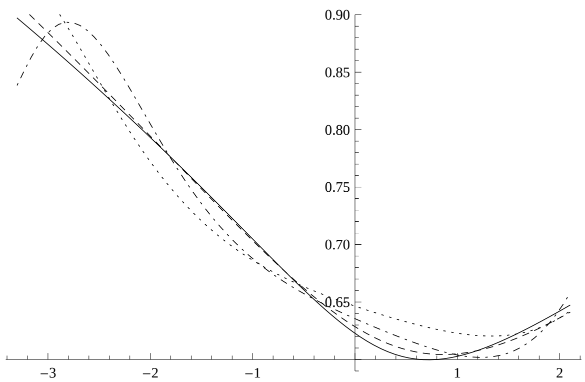
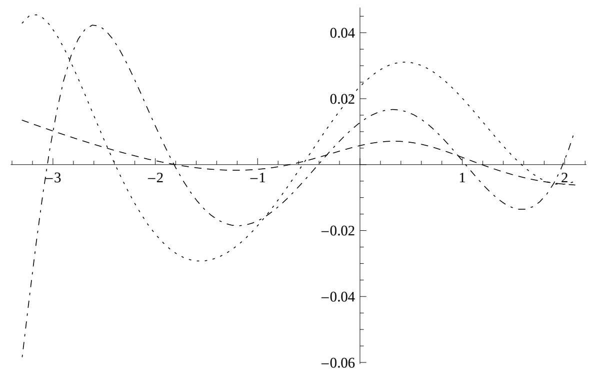


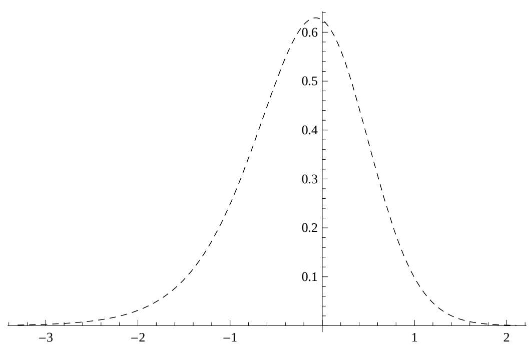
Figure 3: Effect of σ_0 (dotted) on the implied volatility approximation $\sigma^{(3,8)}$ (dashed) in the Variance Gamma model. The true implied volatility corresponds to the solid line. The parameters used for the plots are the same as those used in figure 2: $a = 0.0$, $M = 7.00$, $G = 6.00$, $\alpha = 4.50$, $t = 1.00$, $x = 0.00$.



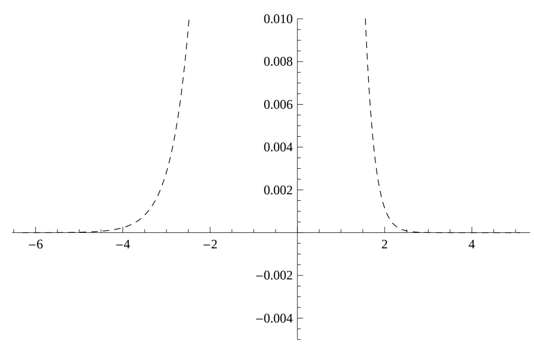
(a) Implied volatilities



(b) Implied volatility errors

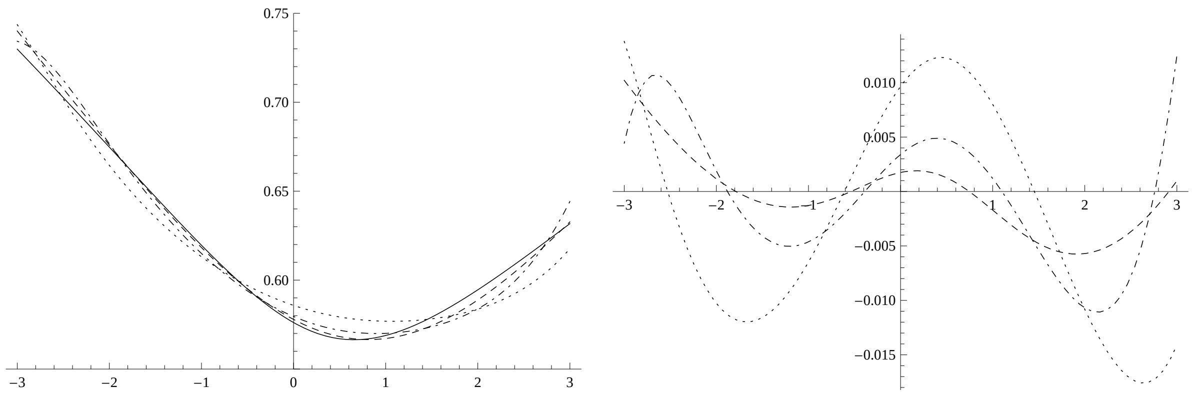


(c) Density

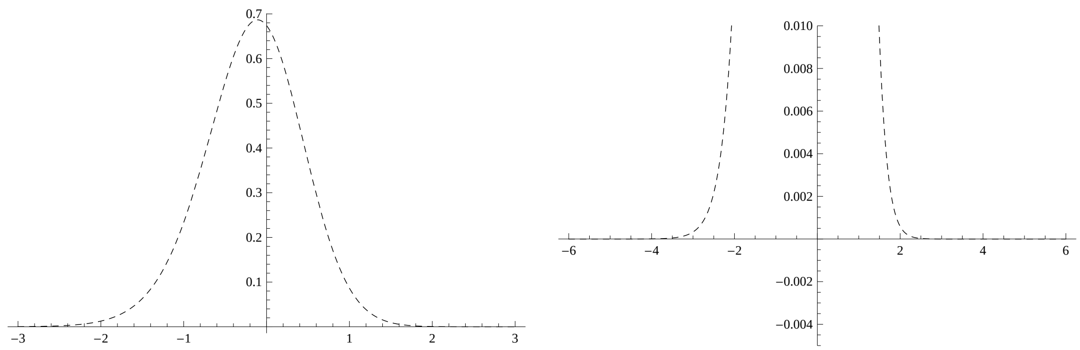


(d) Tails of the density

Figure 4: Numerics for the Variance Gamma model detailed in Section 4.3. In the top graphs, the solid line corresponds to the true implied volatility, whereas the others are the approximations $\sigma^{(2,6)}$ dots, $\sigma^{(3,6)}$ dots-dashed, and the SVI smoothing of $\sigma^{(3,6)}$ dashed. The two plots below are the density (and its tails) corresponding to the SVI smoothing.

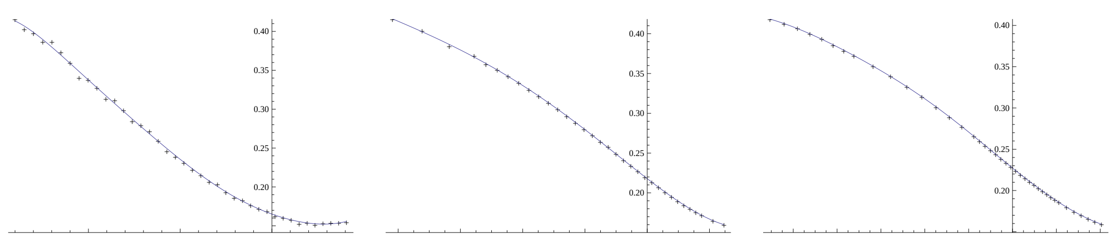


(a) Implied volatilities (b) Implied volatility errors



(c) Density (d) Tails of the density

Figure 5: Numerics for the Variance Gamma model detailed in Section 4.3. In the top graphs, the solid line corresponds to the true implied volatility, whereas the others are the approximations $\sigma^{(2,6)}$ dots, $\sigma^{(3,6)}$ dots-dashed, and the SVI smoothing of $\sigma^{(3,6)}$ dashed. The two plots below are the density (and its tails) corresponding to the SVI smoothing.



(a) $t = 0.033$ years (b) $t = 0.70$ years (c) $t = 1.45$ years

Figure 6: Model-free fit to SPX options from Jan 4, 2010 as explained in Section 4.5. The horizontal axis represents the log-moneyness $(k - x)$.