Asymptotic Methods for Computing Implied Volatilities Under Stochastic Volatility

Alexey Medvedev

First version: November 2004
Current version: December 2004

This research has been carried out within the NCCR FINRISK project on “Financial Econometrics for Risk Management”.
ASYMPTOTIC METHODS FOR COMPUTING IMPLIED VOLATILITIES UNDER STOCHASTIC VOLATILITY

Alexey MEDVEDEV

a HEC Genève and FAME, Université de Genève, 102 Bd Carl Vogt, CH - 1211 Genève 4, Suisse.
medvedev@hec.unige.ch

First draft: November 2004
Current draft: December 2004

Abstract

In this paper we propose analytical approximations for computing implied volatilities when time-to-maturity \( \tau \) is small. The analysis is performed in the framework of a two-factor model with local and stochastic volatility. We describe an algorithm for building the power series approximation of implied volatility. In the case of CEV volatility of volatility we first obtain a quasi-analytical solution for the limit of implied volatilities \( \Upsilon \) as \( \tau \to 0 \). Then we show that implied volatilities of short term options can be accurately computed by a proper transformation of \( \Upsilon \). We introduce a class of models for which this method may be accurate also for \( \tau \gg 0 \). In the particular case of SABR model we obtain an extension of the formula derived in Hagan et al. (2002).

Key words: Option pricing, stochastic volatility, local volatility, implied volatility, short term asymptotics.

JEL Classification: G12.

---

1Financial support by the National Centre of Competence in Research "Financial Valuation and Risk Management" (NCCR FINRISK) is gratefully acknowledged.
1 Introduction

It is now generally accepted that Black-Scholes model cannot be used to accurately price European options in practice. To make it comply with empirical facts the model has been generalized in numerous ways. Stochastic and local volatility models are the extensions of the Black-Scholes model most widely used in practice. The first generalization is made by introducing the second stochastic factor driving the volatility of underlying asset. The second generalization involves a dependence of the volatility on the price of the underlying and time. Alternative extensions are made by adding jumps, however, the use of them is still limited. These models require the specification of probability distributions related to rare events, and, more importantly, they do not allow perfect hedging.

Stochastic volatility models do not yield closed form solutions for European options except for few cases. It is therefore not surprising that various asymptotic techniques have been applied to obtain analytical approximations of option prices and implied volatilities. Table 1 below lists the papers dealing with asymptotic approximations, and the methods used to obtain them.

Insert Table 1 somewhere here

In this paper we derive approximations for implied volatilities when time-to-maturity $\tau$ is small. We perform the analysis under a general two-factor specification with local and stochastic volatility. We obtain several results. First, we describe an algorithm for computing the power series representation of implied volatilities up to virtually any order. Here we generalize the results obtained in Medvedev and Scaillet (2004) in the class of pure diffusion models. Second, we derive a quasi-analytical solution for the limit of implied volatilities $\Upsilon$ as $\tau \to 0$ in the case of CEV volatility of volatility. We combine both results, and propose a method for computing implied volatilities for $\tau > 0$ via the transformation of $\Upsilon$.

In a closely related paper Berestycki et al. (2004) deal with the short term asymptotics under a multifactor setting. They propose the direct use of the short term asymptotics to compute implied volatilities. The asymptotic expansion, however, is not available in a closed form. Since we have to compute it numerically the advantage of this approach is not obvious.

Numerical experiments based on realistic model parameters suggest that our method works well for short term option (as short as one month). Im-
portantly the method yields accurate approximations whenever the short term asymptotics converges, hence, there is no need to solve for the latter as proposed in Berestycki et al. (2004).

In the paper we introduce a class of stochastic volatility models for which the method is accurate not only near expiry. The popular SABR model is a particular case. The approximation formula derived by Hagan et. al. (2002) for SABR model appears to be a particular case of our method.

The paper is organized as follows. In the next section we introduce stochastic and local volatility models. In the section that follows we describe the algorithm for computing the power series expansion of implied volatilities. In Section 4 we relate our results to other asymptotic expansions, and in Section 5 we provide some numerical examples highlighting the problems with the use of the series expansion. In Section 6 we assume CEV volatility of volatility, derive the formula for the limit of implied volatility \( \Upsilon \), and describe the method of computing implied volatilities for \( \tau > 0 \). The case of SABR model is considered separately in details.

2 Stochastic, local and implied volatility

Throughout the paper we will be dealing with pricing of stock options. The theory can be equally applied to other important derivatives such as options on forward rates (caps and caplets), and options on foreign exchange.

It is well known that if stock price \( S_t \) follows a log-normal diffusion process:

\[
dS_t = \mu_t S_t dt + \sigma S_t dW_t,
\]

then the European, let say, call option price expiring at date \( T \) is given by the Black-Scholes formula:

\[
C_t(K, \tau) = C_{BS}(X_t, K, \tau, \sigma) \equiv Ke^{-r\tau} \left[ e^{X_t} N(d_1) - N(d_2) \right],
\]

where

\[
d_{1,2} = \frac{X_t}{\sigma \sqrt{\tau}} \pm \frac{1}{2} \sigma \sqrt{\tau}, \quad N(d) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d} e^{-\frac{s^2}{2}} ds
\]

\( r \) and \( \delta \) are the constant interest rate and the constant dividend yield respectively, \( K \) is the option strike price, \( \tau = T - t \) is the option time-to-maturity,
\[ X_t = \log \left( \frac{S_t e^{(r-\delta)\tau}}{K} \right) \] is the option moneyness. An equivalent way of finding the option price is to consider the risk-neutral process for \( S_t \):

\[ dS_t = (r-\delta)S_t dt + \sigma S_t dW_t, \tag{2} \]

and compute the discounted expectations of the option’s final payoff:

\[ C_t(S_t, K) = E_t \left[ e^{-r\tau} \max \{S_T - K, 0\} \right]. \tag{3} \]

Let us consider a generalization of the risk-neutral process (2) by making the volatility stochastic and dependent on \( S_t \) and \( t \):

\[
\begin{align*}
dS_t &= (r-\delta)S_t dt + \sigma_t S_t F(S_t, t) dW_t^{(1)}, \\
d\sigma_t &= a(\sigma_t) dt + b(\sigma_t) \left( \rho dW_t^{(1)} + \sqrt{1-\rho^2} dW_t^{(2)} \right),
\end{align*} \tag{4}
\]

Here \( W_t^{(1)} \) and \( dW_t^{(2)} \) are independent Brownian motions, and \( F \) is the local volatility function. The call option price is again given by (3). Most of models used in practice are particular cases of (4). Two well known classes of models are local volatility models (\( \sigma \) is constant) introduced by Dupire (1994), and stochastic volatility models (\( F \equiv 1 \)) first studied in Hull and White (1987).

In practice options are quoted in implied volatilities rather than prices. The (Black-Scholes) implied volatility of, let say, a call option is defined as the value of the volatility parameter in the Black-Scholes formula that, other things equal, leads to the correct option price. Formally:

\[ C_t(K, \tau) = C^{BS}(X_t, K, \tau, \text{ImVol}), \tag{5} \]

Let us write the dynamics of \( X_t \) which easily follows from (4) by applying Itô’s lemma:

\[
\begin{align*}
dx_t &= -\frac{1}{2} \sigma_t^2 f^2(X_t, t) dt + \sigma_t f(X_t, t) dW_t^{(1)}, \\
d\sigma_t &= a(\sigma_t) dt + b(\sigma_t) \left( \rho dW_t^{(1)} + \sqrt{1-\rho^2} dW_t^{(2)} \right),
\end{align*} \tag{6}
\]

where

\[ f(X_t, t) \equiv F \left( K e^{X_t-(r-\delta)\tau}, t \right), \quad \tau = T - t \tag{7} \]

and from (3) the call option price is given by:

\[ C_t(K, \tau) = KE_t \left[ e^{-r\tau} \max \{e^{X_T} - 1, 0\} \right]. \tag{8} \]

Let us substitute (8), and the Black-Scholes formula (1) in the identity (5) getting rid of \( K \):

\[ E_t \left[ e^{-rT} \max \{e^{X_T} - 1, 0\} \right] = \frac{1}{K} C^{BS}(X_t, K, \tau, \text{ImVol}). \tag{9} \]
From (6) we observe that the distribution of $X_T$ is determined by $X_t$ and $\sigma_t$. From (9) it follows that the implied volatility is a deterministic function of four parameters:

$$ImVol = I(X_t, t, \tau, \sigma_t),$$

An important remark is in order. The form of function $f$ depends on the option under consideration (see (7)). Keeping this in mind we can develop a theory of option pricing for arbitrary $f$, and then use the relationship (7) to determine the correct form of this function. In the rest of the paper we will assume model (6) without further justification. To simplify the exposition we will also skip the dependence of the local volatility function on $t$ since all the results can be easily generalized by adding $t$ in the list of arguments.

The call option price and the implied volatility under model (6) do not admit closed form solutions except for few special cases. In the simplest case of deterministic volatility ($b = 0$, $f \equiv 1$) the implied volatility is equal to the square root of the average variance over the life of the option:

$$I(\tau, \sigma_t) = \left( \frac{1}{\tau} \int_t^{t+\tau} \sigma_s^2 ds \right)^{\frac{1}{2}}. \tag{10}$$

In general the option price and the implied volatility are related to the average variance in a non-trivial way.

In this paper we deal with analytical approximations of $I$ based on its short term asymptotics, that is, when $\tau$ is small. To fix the terminology let us call the short term asymptotics of $I$ the following representation:

$$I(X, \tau, \sigma) = \sum_{n=0}^{\infty} I_n(X, \sigma) \tau^n. \tag{11}$$

In (11) the asymptotic terms $I_n$ can be further substituted by their power series expansions in $X$. The resulting representation:

$$I(X, \tau, \sigma) = \sum_{n,m=0}^{\infty} I_{nm}(\sigma) X^m \tau^n, \tag{12}$$

we name the power series expansion of $I$. The important difference between (11) and (12) is that the latter requires convergence not only in $\tau$ but also in $X$. We will return to this point in next sections.

Making the use of formula (10) we can easily obtain the asymptotic ex-
pansion of $I$ in the deterministic volatility case:

$$I(\tau, \sigma_t) \simeq \left( \frac{1}{\tau} \int_{t}^{t+\tau} (\sigma_t + a(\sigma_t)(s-t))^2 \, ds \right)^{\frac{1}{2}}$$

$$\simeq \left( \frac{1}{\tau} \int_{t}^{t+\tau} (\sigma_t^2 + 2\sigma_t a(\sigma_t)(s-t)) \, ds \right)^{\frac{1}{2}}$$

$$\simeq \sigma_t + \frac{a(\sigma_t)}{2}\tau. \quad (13)$$

Let us observe that when $\tau \to 0$ implied volatility converges to spot volatility $\sigma_t$. In general, this is true only for at-the-money (ATM) implied volatilities (see Ledoit and Santa-Clara (2002) e.g.), that is, when $X = 0$. Let us denote:

$$\Upsilon(X, \sigma) = \lim_{\tau \to 0} I(X, \tau, \sigma) = I_0(X, \sigma),$$

then we have:

$$\Upsilon(0, \sigma) = I_{00} = \sigma f(0). \quad (14)$$

In the special case of $b = 0$ the limit of implied volatility $\Upsilon$ is given by (see Berestycki et al. (2002) e.g.):

$$\Upsilon(X, K, \sigma) = \sigma X \left( \int_{0}^{X} \frac{ds}{f(s)} \right)^{-1}. \quad (15)$$

In the paper we generalize these results by solving for $\Upsilon$ under model (6) assuming CEV type of the volatility of volatility (vovol) $b(\sigma) = \beta \sigma^\nu$.

Before we proceed with a technical section let us introduce a convenient parameterization of implied volatility. Denoting $x = X/\sqrt{\tau}$ we define new function $i$:

$$i(x, \tau, \sigma) \equiv I(x\sqrt{\tau}, \tau, \sigma). \quad (16)$$

In the next section we derive the short term asymptotics of $i$ in the form:

$$i(x) = i_0(x) + i_1(x)\sqrt{\tau} + i_2(x)\tau + i_3(x)\tau\sqrt{\tau}... \quad (16)$$

Here and occasionally in the rest of the paper we will not indicate the dependence on $\sigma$ when it is not important.

Let us note that keeping $x$ fixed and letting $\tau \to 0$ means that $X$ is also forced to converge to zero. Hence, using (14) we conclude that $i_0 = \sigma f(0)$. The other terms in (16) can be obtained by means of a recursive relationship described in the Proposition 1. They have to be polynomials in $x$ so that by passing from $x = X/\sqrt{\tau}$ to $X$ we arrive at the power series expansion of $I$ (12).
3 Short term asymptotics of $i$

In this section we derive an algorithm for computing the short term asymptotic expansion (16). This section is technical and can be skipped by a reader not interested in details of the derivation.

Under model (6) call option price satisfies:

$$-C_\tau + \frac{1}{2}(C_{XX} - C_X)\sigma^2 f^2(X) + \frac{1}{2}C_{\sigma\sigma}b^2 + C_{\sigma\bar{x}}\sigma b \rho f(X) = 0. \quad (17)$$

Let us proceed with finding the partial differential equation (PDE) for $I$. To do this we substitute $C = C^{BS}(X, K, \tau, I(X, \sigma, \tau),)$ into (17). The resulting equation is simplified by dividing by $C^{BS}_\sigma$, and subsequent simplification (see Ledoit and Santa-Clara (2002) for a similar derivation). The PDE for $I$ takes the following complicated looking form:

$$0 = \frac{\rho b \sigma f}{I^{1/\tau}} d_2 I_\sigma - \frac{1}{2} b^2 d_1 d_2 (I_\sigma)^2 - b \rho \sigma f I_{X\sigma} + \frac{I}{2\tau} + I_\tau$$

$$-\frac{\sigma^2 f^2}{2I \tau} + \frac{\sigma^2 f^2}{I^{1/\tau}} d_2 I_X - \frac{\rho b \sigma f}{I} d_1 d_2 I_X I_\sigma - \frac{1}{2} \frac{\sigma^2 f^2}{I} d_1 d_2 (I_X)^2$$

$$-\frac{1}{2} \frac{\sigma^2 f^2 (I_{XX} - I_X)}{I^{1/\tau}} - \frac{b^2}{2} I_{\sigma\sigma} - a I_\sigma, \quad (18)$$

where

$$d_{1,2} = \frac{X}{I^{1/\tau}} \pm \frac{I^{1/\tau}}{2}.$$

The boundary condition for this equation does not easily follow from that of PDE for the option price since implied volatility is indeterminable at $\tau = 0$. To pin down the right solution of (18) we impose a regularity condition. Namely we will be looking for those solutions that admit power series representation (12) in some neighborhood of $X, \tau = 0$. As it will become clear shortly there is only one solution of (18) which satisfies this regularity condition. To motivate the choice of the solution, let us observe that the slow-volatility-variation asymptotics of implied volatility derived by Lee (2001) for the case $f \equiv 1$ represents the part of the series (12) with $m + n \leq 2$.

Although equation (18) looks complicated it has a nice structure that allows solving for the power series expansion of its regular solution virtually up to any order. Let us first derive the short term asymptotics of $i$ given by
(15), which is related to (12) via the substitution $X = x\sqrt{\tau}$:

$$i(x, \sigma, \tau) = \sum_{n,m=0}^{\infty} I_{nm}(\sigma) (x\sqrt{\tau})^n \tau^m$$

$$= I_{00} + I_{10} x\sqrt{\tau} + I_{01} \tau + I_{20} x^2 \tau + I_{11} x\tau \sqrt{\tau} + I_{02} \tau^2 + ...$$

$$= I_{00} + [I_{10} x \sqrt{\tau} + I_{20} x^2 + I_{01}] \tau + ...$$

$$= \sum_{m=0}^{\infty} i_m(x, \sigma) (\sqrt{\tau})^m.$$  

Careful observation reveals that $i_m$ should necessarily be a polynomial of order $m$ in $x$ having the form

$$i_m(x) = i_{mm} x^m + i_{mm-2} x^{m-2} + ... i_{mm-2M} x^{m-2M}, \quad M = [m/2]. \quad (19)$$

Let us now make the use of this nice property to find a recursive relationship for asymptotic terms $i_m$. This relationship will involve PDEs whose solutions are easy to find given their particular polynomial form (19).

We proceed with finding the PDE for $i$. Let us make the following substitutions into (18):

$$I_X = i_x \tau^{-1/2} \quad I_{XX} = i_{xx} \tau^{-1}$$

$$I_{X\sigma} = i_{x\sigma} \tau^{-1/2} \quad I_{\tau} = -\frac{1}{2} x i_x \tau^{-1} + \frac{1}{2} i_x \sqrt{\tau} \tau^{-1/2}$$

$$I_{\sigma} = i_{\sigma}.$$ 

After some tedious but trivial algebra we arrive at the following equation for $i$:

$$0 = \Phi(i)$$

$$= \frac{i}{2} - \frac{\sigma^2 f^2}{2i} - \frac{i_x x}{2} - \frac{\sigma^2 f^2}{2i} i_{xx} + \frac{\sigma^2 f^2}{2i^2} x i_x - \frac{\sigma^2 f^2}{2i^3} x^2 (i_x)^2$$

$$+ \left[ \rho \sigma f \left( \frac{x i_x}{i^2} - i_{xx} - \frac{x^2 i_x i_{\sigma}}{i^3} \right) + \frac{1}{2} i_x \sqrt{\tau} \right] \tau$$

$$+ \left[ -\frac{\rho \sigma f i_{x\sigma}}{2} - \frac{b^2 x^2}{2i^3} (i_x)^2 + \frac{\sigma^2 f^2}{8} (i_x)^2 i_x - b^2 (i_{\sigma\sigma} - \alpha i_x) \right] \tau$$

$$+ \left[ \frac{\rho \sigma f i_{x\sigma i_x i_{\sigma}}}{4} \right] \tau \sqrt{\tau}$$

$$+ \left[ \frac{b^2 i (i_x)^2}{8} \right] \tau^2,$$
where $f$ stands for $f(x\sqrt{\tau})$. Let us denote by $\Phi_m(h)$ $m$-th order term in the short term expansion of function $\Phi(h)$. That is:

$$\Phi(h) = \Phi_0(h) + \Phi_1(h)\sqrt{\tau} + ... \Phi_m(h)(\sqrt{\tau})^m + ...,$$

If $\Phi_m(h)$ is a polynomial in $x$ then we will denote by $\Phi_{mn}(h)$ the coefficient before $n$-th power of $x$, that is:

$$\Phi_m(h)(x) = \Phi_{m0}(h) + \Phi_{m1}(h)x + \Phi_{m2}(h)x^2 + ...$$

The last piece of notation is $i^{(m)}$ which denotes the short term expansion of $i$ up to $m$-th order of $\sqrt{\tau}$:

$$i^{(m)} \equiv i_0 + i_1\sqrt{\tau} + ... + i_m(\sqrt{\tau})^m.$$

We can now state the main result of the section.

**Proposition 1** Under model (6) implied volatility function $i(x, \sigma, \tau)$ has short term asymptotics:

$$i(x, \sigma, \tau) = \sum_{m=0}^{\infty} i_m(x, \sigma)(\sqrt{\tau})^m,$$

with

$$i_m(x) = i_{mm}x^m + i_{mm-2}x^{m-2} + ... + i_{mm-2M}x^{m-2M},$$

where $M = [m/2]$ and $i_{mm-2k}$ satisfy the following recursive relationship:

$$i_{mm-2k} = \frac{\sigma^2 f(0)^2(m - 2k + 2)(m - 2k + 1)i_{mm-2k+2} - 2\Phi_{mm-2k}(i^{(m-1)})}{2m - 2k + 2},$$

$k = 0...M$, $i_{mm+2} \equiv 0$ and $i_{00} = \sigma f(0)$.

**Proof.** see Appendix A.

Let us note that the procedure described in Proposition 1 can be easily implemented in any symbolic software like, for example, Maple that is used by the author. This method is universal and can be applied to any model of type (6).

4 Relation to other expansion of implied volatility

The power series expansion of implied volatilities (12) can be easily obtained from (20) by passing back from $x = X/\sqrt{\tau}$ to $X$. The resulting expression
analytically coincides with some asymptotic expansions obtained in the literature using different approaches. Indeed, let us for illustration consider here the short term asymptotics of $i$ to the order of $\tau$ assuming $f \equiv 1^2$:

$$i = \sigma + i_1 \sqrt{\tau} + i_2 \tau + O(\tau \sqrt{\tau}),$$  \hfill (21)

with

$$i_1 = -\frac{\rho b}{2\sigma},$$

$$i_2 = \left[ -\frac{5}{12} \frac{\rho^2 b^2}{\sigma^3} + \frac{1}{6} \frac{b^2}{\sigma^3} + \frac{1}{6} \frac{\rho^2 b\sigma}{\sigma^2} \right] \frac{x^2}{2}$$

$$+ \left[ a + \frac{\rho b \sigma}{4} + \frac{1}{24} \frac{\rho^2 b^2}{\sigma} + \frac{1}{12} \frac{b^2}{\sigma} - \frac{1}{6} \frac{\rho^2 b b'}{\sigma} \right] \frac{\tau}{2}$$

$$+ \left[ -\frac{5}{12} \frac{\rho^2 b^2}{\sigma^3} + \frac{1}{6} \frac{b^2}{\sigma^3} + \frac{1}{6} \frac{\rho^2 b b'}{\sigma^2} \right] x^2 \text{Residual}. \hfill (22)$$

Here $b'$ denotes the derivative of $b(\sigma)$.

Let us assume $a = 0$, and denote $b(\sigma) = \epsilon \bar{b}(\sigma)$ understanding $\epsilon$ as a small number. Let us observe that $i_1$ is of order $\epsilon$, and $i_2$ involves terms of order $\epsilon^2$. Following the analogy we guess, which in fact is true, that $i_3$ will involve terms of order $\epsilon^3$ etc. This observation suggests that the small volatility expansion\(^3\) of the implied volatility (derived in Lewis (2000)) should yield the same analytical expression as the short term asymptotics presented here.

Substitution of $x = X/\sqrt{\tau}$ into the expressions for $i_1$ and $i_2$ yields the series expansion of $I$:

$$I = I_{00} + I_{10} X + I_{01} \tau + I_{20} X^2 + \text{Residual}$$

$$= \sigma + \left[ -\frac{\rho b}{2\sigma} \right] X$$

$$+ \left[ a + \frac{\rho b \sigma}{4} + \frac{1}{24} \frac{\rho^2 b^2}{\sigma} + \frac{1}{12} \frac{b^2}{\sigma} - \frac{1}{6} \frac{\rho^2 b b'}{\sigma} \right] \frac{\tau}{2}$$

$$+ \left[ -\frac{5}{12} \frac{\rho^2 b^2}{\sigma^3} + \frac{1}{6} \frac{b^2}{\sigma^3} + \frac{1}{6} \frac{\rho^2 b b'}{\sigma^2} \right] X^2 + \text{Residual.} \hfill (22)$$

Let us compare (22) with the analytical results obtained by Lee (2001) to see that the slow-volatility-variation asymptotics derived there yields the same expression. The last observation is not surprising at all. Indeed, if $\tau$ is sufficiently small the volatility is slow varying relative to the option time-to-maturity. Let us further observe that by setting $b = 0$ in (22) we come up with the asymptotic expansion (13).

\(^2\)This formula was also derived in Medvedev and Scaillet (2004)

\(^3\)The asymptotics of $I$ for small $\epsilon$ instead of $\tau$.  

10
Formula (22) can be used to approximate \( \Upsilon(X, \sigma) \). Setting \( \tau = 0 \) in (22) we obtain:

\[
\Upsilon(X, \sigma) = I_{00} + I_{10}X + I_{20}X^2 + O(X^3)
\]

\[
= \sigma + \left[ \frac{\rho b}{2\sigma} \right] X
+ \frac{1}{2} \left[ -\frac{5 \rho^2 b^2}{6 \sigma^3} + \frac{1}{3} \frac{b^2}{\sigma^3} + \frac{1}{3} \frac{\rho^2 b^3}{\sigma^2} \right] X^2 + O(X^3).
\]

(23)

Implied volatility is known to be a symmetric function of \( X \) if \( \rho = 0 \) (Renault and Touzi (1996)). The latter can indeed be observed from (23). On the contrary, if the correlation is different from zero then near ATM implied volatilities form a skew, whose magnitude is given by \(-\frac{\rho b}{2\sigma}\). The skew of actual implied volatilities when plotted against \( K \) is typically negative. This is indeed predicted by (23) if we recall the our definition of moneyness \( X = \log(S/K) \), and note that \( \rho < 0 \) in practice.

Although the quadratic approximation of \( \Upsilon \) (23) seems to be promising it is not sufficiently accurate. The implied volatility smile appears to have much more complex structure. To obtain a more accurate representation of implied volatilities we need to find higher order expansions of \( I \) using the algorithm proposed in Proposition 1. Let us note that the derivation of the second order asymptotics of \( i \) (21) is already involving if done in the paper. Fortunately, we are able to program the algorithm in a symbolic software and derive asymptotics up to virtually any order of \( \sqrt{\tau} \).

In Figure 1 we plot the graph of the time required for building the short term asymptotics of \( i \) as a function of the order of expansion. For example, MAPLE requires only one and a half minute on a computer with Intel processor Pentium IV, 2.66 MG and non-binding RAM to build the asymptotics of \( i \) up to \((\sqrt{\tau})^9\). In particular, this allows to obtain the power series expansion of \( \Upsilon \) up to \( X^9 \). Importantly, this expansion should be constructed only once. Once found the expression involves only elementary functions, hence, the computation of implied volatilities is extremely fast. On the contrary, Monte-Carlo simulations should be run each time an option price is needed.

At this point one may have an illusion that we can apply the power expansion of implied volatilities (12) to price options with any strike and time-to-maturity. Unfortunately, the life appears to be not as simple as that. In the next section we will observe that series (12) converges only in the neighborhood of \( X = \tau = 0 \), and this neighborhood is not always sufficiently wide for the method to be universal. The main problem seems to be the divergence of the series for far-from-the-money (FFM) options (\( |X| >> 0 \)) even if \( \tau = 0 \). In particular, this does not permit computing the risk-neutral
price density from implied volatilities.

5 Numerical examples

In this section we study the performance of the series expansion of implied volatilities (12) based on realistic model parameters. In the numerical experiments we will use three benchmark models. Three examples share some common features: we set \( f \equiv 1 \) (no local volatility), assume CEV function for the vovol, and choose the drift of the variance \( \sigma^2 \) to be a linear function (volatility mean-revertion). In terms of our specification (6) we will have:

\[
a(\sigma) = \frac{\kappa(\omega - \sigma^2) - b^2}{2\sigma}, \\
b(\sigma) = \beta\sigma^\nu, \quad \beta \geq 0.
\]

We borrow three sets of model parameters from Ait-Sahalia and Kimmel (2004), where different stochastic volatility model specifications, including general CEV type one, have been fitted to the market data on S&P500.

The first example is the so called Heston model with \( \nu = 0 \). Under this specification option prices are known in a closed form, and for this reason the Heston model is extremely popular in the empirical literature. The second example is the GARCH model with \( \nu = 1 \) which does not admit a closed form solution. The third example results from the fit of the general CEV model to the market data, which yielded \( \nu = 0.7 \).

The actual and approximate values of implied volatilities of one-month options are shown in Figure 2. Here we used Monte-Carlo simulation to find option prices under GARCH and CEV specifications employing two variance reduction techniques: antitetic variables and conditional simulation (see Jones (2003) for details). Option prices under the Heston specification were computed using known closed form solution via the Fourier transform. The series expansion of \( I \) corresponds to the short term asymptotics of \( i \) to the order of \( (\sqrt{\tau})^9 \).

Let us observe that the power series approximation is extremely accurate in some neighborhood around \( X = -\log(K/S) = 0 \) or for sufficiently close ATM options. The asymptotic approximation is extremely poor outside of this neighborhood suggesting that the power series does not converge there. The accuracy improves noticeably as the value of the spot volatility \( \sigma \) increases, and the interval of convergence clearly widens.

These observations have important implications going beyond the present paper. As it has been noted in the previous section the slow-volatility-variation asymptotics of Lee (2001), and the small-vovol asymptotics of Lewis...
(2000) analytically coincide with the short term asymptotics presented here\(^4\). Since the series convergence is not always guaranteed one can now doubt the applicability of these results. Good news is that in our case the interval of convergence can be easily found by the comparison of the asymptotic approximations at different orders. The difference should obviously be negligible at higher orders if the series converges.

Numerical experiments show that the divergence of the power series expansion for FFM options might impact its practical value. It is therefore desirable to obtain the asymptotics of \( I \) (11), as pursued by Berestycki (2004), rather than its power series expansion (12). However, this appears to be an extremely difficult task in general case. So in the rest of the paper we will consider the particular case of CEV vovol.

6 Short term asymptotics of \( I \): the case of CEV vovol

In this section we consider the particular case of CEV vovol, which is the most common specification used in practice. We first solve for the limit of implied volatilities \( \Upsilon \). Then we propose an improved method for computing implied volatilities.

6.1 Derivation of \( \Upsilon \)

Let us investigate the limit of implied volatilities as \( \tau \) goes to zero for given moneyness \( X \). From (18) it is straightforward to obtain the PDE for the limiting smile \( \Upsilon \) by collecting terms with the lowest power of \( \tau \), and setting their sum equal to zero:

\[
\Upsilon^2 \sigma^2 - \Upsilon^4 - 2X \sigma f \Upsilon (\rho b \Upsilon + \sigma Y_X) + X^2 (b^2 \Upsilon^2 + \sigma^2 f^2 \Upsilon_X^2 + 2 \rho b \sigma f \Upsilon_X \Upsilon_Y) = 0,
\]

This equation under different notation can be also found in Schönbucher (1999) and Berestycki et al. (2004). In both papers the PDE is written down but not solved explicitly.

Let us now assume a particular shape of vovol:

\[
b(\sigma) = \beta \sigma^\nu, \quad \beta \geq 0.
\]

\(^4\)Strictly speaking the small vovol asymptotics yields the same expression if the volatility drift \( a \) is zero.
As it was mentioned in the previous section we are looking for the regular solution of (24). That is the one that admits a power series expansion in $X$. The next Proposition characterizes $\Upsilon$.

**Proposition 2** Assume model (6) and CEV vovol (25) with $\gamma < 2$ then $\Upsilon$ can be written as:

$$\Upsilon(X, \sigma) = \sigma X \zeta \left( \int_0^X \frac{ds}{f(s)} \cdot \int_{\alpha_0}^{\alpha_1} \frac{1}{\sqrt{1 + G^2(s)}} ds \right)^{-1} \text{ for } |\rho| < 1,$$

(26)

where $G$ satisfies:

$$G' = \frac{1 + G^2}{(2 - \nu) s G + 1}, \quad G(\alpha_0) = \alpha_0 (2 - \nu),$$

(27)

with

$$\zeta = \frac{\beta}{\sigma^{2-\nu}} \int_0^X \frac{ds}{f(s)},$$

$$\alpha_0 = -\frac{\rho}{(2 - \nu) \sqrt{1 - \rho^2}},$$

$$\alpha_1 = \frac{\zeta}{\sqrt{1 - \rho^2}} + \alpha_0;$$

and

$$\Upsilon(X, \sigma) = \sigma (1 - \nu) X \zeta \left\{ \int_0^X \frac{ds}{f(s)} \cdot \left[ 1 - (1 - \zeta (2 - \nu))^{\frac{1}{\nu - 2}} \right] \right\}^{-1} \text{ for } |\rho| = 1,$$

with

$$\zeta = \frac{\rho \beta}{\sigma^{2-\nu}} \int_0^X \frac{ds}{f(s)};$$

**Proof.** see Appendix B.

Proposition 2 reduces the complicated PDE (24) to a simple ordinary differential equation. The latter can be solved by propagating the function from its initial value using the explicit expression for the derivative (27).

In the rest of the section we assume $f \equiv 1$ and $|\rho| < 1$. In the particular case of $\nu = 1$ the analytical solution to (27) is given by $G(s) = s$, and (26) yields:

$$\Upsilon(X, \sigma) = \frac{\sigma \zeta}{\ln \left( \frac{-\rho + \sqrt{1 - 2\rho + \zeta^2}}{1 - \rho} \right)}.$$
This solution was also found in Hagan et al. (2002) with the help of a perturbation approach, and Berestycki et al. (2004) by solving a differential equation. The expression for $\Upsilon$ contains a radical which does not admit a powers series expansion in $\zeta$ everywhere.

In the general case we have:

$$\Upsilon(X, \sigma) = \sigma \zeta \left( \int_{\alpha_0}^{\alpha_1(\zeta)} \frac{1}{\sqrt{1 + G^2(s)}} ds \right)^{-1}.$$ 

This expression suggests that the implied volatility limit $\Upsilon$ depends on $X$ through $\zeta = \frac{\beta X}{\sigma^2\tau}$. As a consequence, the interval of convergence of the power expansion of $\Upsilon$ should be wider for larger $\sigma$, which we indeed observed in the numerical experiments.

### 6.2 Implied volatility as a transformation of $\Upsilon$

In the previous section we obtained a quasi-analytical solution for the limit of implied volatility $\Upsilon$. The derivation of the other asymptotic terms in (11) seems to an extremely difficult task. In this section we take another approach by proposing an alternative to the asymptotic expansion of implied volatilities.

Let us recall that the power series expansion (12) works relatively well for ATM implied volatilities as suggested by numerical experiments. If $\tau$ is small then the shape of implied volatility smile should not be "much" different from that of the limit $\Upsilon$. Hence we can approximate implied volatilities for $\tau \approx 0$ by the curve $\Upsilon(X)$ shifted, let say proportionally, to exactly match the ATM implied volatility. In fact, we can do better then this.

Let us proceed in a formal way. We start by noting that in the neighborhood of $X = \tau = 0$ the implied volatility can be written as:

$$I(X, \tau) = \Upsilon \left( X \lambda(X, \tau) \right) \Pi(\tau),$$  

where $\lambda$ and $\Pi$ are some functions. Indeed, from (28) we have:

$$\Pi(\tau) = \frac{I(0, \tau)}{\Upsilon(0)},$$

and assuming:

$$\lambda(X, \tau) = \Lambda(\tau) + \Lambda_1(\tau) X + ...$$

\footnote{Here we skip the dependence on $\sigma$ to simplify the exposition.}
we obtain:

\[
\Lambda(\tau) = \frac{I_X(0, \tau)}{\Upsilon_X(0) \Pi^{-1}(\tau)},
\]

(31)

\[
\Lambda_1(\tau) = \frac{I_{XX}(0, \tau) - \Upsilon_{XX}(0) \Lambda(\tau)}{\Upsilon_X(0) \Pi^{-1}(\tau)},
\]

\[
\ldots
\]

**Assumption:** In (30) we can drop terms with \(X\):

\[
\lambda(X, \tau) \simeq \lambda(0, \tau) = \Lambda(\tau)
\]

(32)

It follows that \(^6\):

\[
I(X, \tau) \simeq \Upsilon(X\Lambda(\tau)) \Pi(\tau).
\]

(33)

To construct approximation (33) we should determine \(\Pi\) and \(\Lambda\), which are functions of \(\tau\) only. Hence, we have eliminated the problem of power series divergence of \(I(X, \tau)\) for \(|X| \gg 0\). From the (29) and (31) short term asymptotics of \(\Pi\) and \(\Lambda\) easily follows. Indeed, the asymptotic expansions of \(I(0, \tau)\) and \(I_X(0, \tau)\) can be obtained using the algorithm proposed in the previous sections. We first find the expansion of \(i(x, \tau)\) - the implied volatility parameterized by \(x = X/\sqrt{\tau}\). By setting \(x = 0\) we determine \(I(0, \tau) = i(0, \tau)\). By taking the derivatives of \(i\) with respect to \(x\) and setting \(x = 0\) we find \(I_X(0, \tau) = i_x(0, \tau)/\sqrt{\tau}\). Note that to obtain the asymptotic expansion of the skew \(I_X(0, \tau)\) up to \(\tau^m\) we need to find the asymptotics of \(i\) up to \(\sqrt{\tau}^{2m+1}\).

An important remark is in order. We could have eliminated the problem of the series divergence in a simpler way by dropping terms with large powers of \(X\) in the expansion of \(I(X, \tau)\) (12). For example, we could have chosen a quadratic approximation similar to (23). Unfortunately, as it has been noted previously, this straightforward approach does not yield an accurate formula for FFM implied volatilities. The representation (33) is also intuitively preferable since it preserves the shape of implied volatility smile, and turns to exact formula at \(\tau = 0\).

Let us note that the equalities (29) and (31) ensure that ATM implied volatility and the skew are exact:

\[
ATM = I(0, \tau) = \Upsilon(0) \Pi(\tau)
\]

\[
Skew = I_X(0, \tau) = \Upsilon_X(0) \Lambda(\tau) \Pi(\tau).
\]

As a consequence the approximation (33) can be interpreted as a transformation of the initial curve by a proportional shift, and a change of scale in such

\(^6\)To be strict we should consider separately the case \(\rho = 0\) when the skew is zero, and both the nominator and the denominator in the right hand side of (31) are equal to zero. The formula, however, appears to be valid also for \(\rho = 0\).
way that ATM implied volatility and the skew coincide with the true ones. Assumption (32) basically means that by fitting two local characteristics of \( I \) we hope to obtain an accurate approximation globally.

To compute the approximation of implied volatility (33) we need to evaluate \( \Pi \) and \( \Lambda \) using the asymptotic equalities given by (31). Let us denote by \( A(n,m) \) the approximation (33) with \( \Pi \) and \( \Lambda \) substituted by their asymptotic expansions up to \( \tau^n \) and \( \tau^m \) correspondingly.

6.3 Performance of \( A(n,m) \)

Let us use three numerical examples as in Figure 2 to study the accuracy of the approximation introduced in the previous section. In Figure 3 we plot the actual implied volatilities and their approximations based on (33) \( A(2,2) \) and \( A(4,4) \) for 1-month options and different levels of spot volatility \( \sigma \).

The approximation (33) yields accurate values of implied volatilities in all three examples when \( n = m = 4 \) with considerable improvement over the series approximation depicted in Figure 2. Let us observe that with the exception of the low volatility case it is sufficient to use the second order expansions of \( \Pi \) and \( \Lambda \) \( A(2,2) \) for the approximation to be reasonably accurate.

Figure 4 shows the accuracy of (33) for 2-month options under the same three sets of model parameters. The immediate observation is that at low levels of the volatility the approximation fails to yield accurate values of implied volatilities. Although the ATM implied volatility seems to be correct the skew is not. To clarify the picture we experimented with higher order expansions and found that they do not improve the accuracy. It is, therefore, possible that the power series expansion of the skew does not converge in this case.

Let us note that if the short term asymptotics of the skew does not converge then we cannot also apply the short term expansion of \( I \) (11) as proposed by Berestycki et al. (2004). Indeed, to compute the option delta we need the derivative of the implied volatility with respect to \( X \), given only by its short term asymptotics. If the latter does not converge then the approximation is clearly not applicable. Taking into account that the short term asymptotics of implied volatility (11) can only be found numerically, it seems not worthwhile to pursue the idea.

From numerical experiments we observe that the level of volatility \( \sigma \) clearly affects the validity of the approximation. In particular, the larger is \( \sigma \) the wider is the range of \( \tau \) where our approach yields accurate results. There exist, however, a subclass of models (6) for which this relationship does not hold. In the next section we introduce these models, and show that
approximation (33) appears to be valid also for long term options.

6.4 A class of stochastic volatility models

Let us obtain some intuition about why the accuracy of the asymptotic approach depends on the level of $\sigma$. Using series expansion (22) we can write down the leading order expansion of ATM implied volatilities:

$$I(0, \tau, \sigma) = \sigma + \left( a + \frac{\rho b}{4} + \frac{\rho^2 b^2}{6} \sigma - \frac{\rho^2 b b'}{6} \right) \tau + O(\tau^2).$$  \hspace{1cm} (34)

Recalling the shape of vovol $b$ (25) it is straightforward to verify that if $\nu < 1$ then the leading term in expansion (34) does not have a finite limit as $\sigma \to 0$. On the contrary, the leading order expansion is well defined at $\sigma = 0$ if $\nu \geq 1$. Let us now consider the leading order expansion of the skew ($f \equiv 1$):

$$I_X(0, \tau, \sigma) = -\frac{\rho b}{2\sigma} + \rho \left[ \frac{3\rho^2 b^3}{16\sigma^3} + \frac{\rho^2 b^2}{24\sigma^2} + \frac{b^2 b'}{8\sigma^2} - \frac{ab'}{6\sigma} + \frac{b b'}{6\sigma} \right] \tau + O(\tau^2).$$  \hspace{1cm} (35)

It can be easily verified that the leading term in expansion (35) is well defined at $\sigma = 0$ if $\nu \geq 1$ and $a(\sigma) = \alpha \sigma^\eta$ with $\eta \geq 1$. In fact, these are sufficient conditions for the short term asymptotics of implied volatilities to exist at $\sigma = 0$. The next Proposition states the general result.

**Proposition 3** Assume model (6) with CEV shape of vovol (25). The terms $I_{nm}(X, \tau, \sigma)$ in the power series expansion (12) satisfy:

$$\lim_{\sigma \to 0} |I_{nm}(X, \tau, \sigma)| < \infty,$$

if the volatility process has the form:

$$d\sigma_t = dt \sum_s \alpha_s \sigma_t^{\eta_s} + d\tilde{W}^{(2)}_t \sum_s \beta_s \sigma_t^{\nu_s}, \quad \eta_s, \nu_s \geq 1.$$  \hspace{1cm} (36)

where

$$d\tilde{W}^{(2)}_t = \rho dW^{(1)}_t + \sqrt{1 - \rho^2} dW^{(2)}_t$$
Proof. For not very large $n,m$ this can be directly verified by computing $I_{nm}$. For other terms we guess that this regularity should hold.

The class of models (36) is sufficiently wide to incorporate mean-reversion in stochastic volatility. Indeed, let us consider the following specification:

$$\sigma = \frac{\phi}{\xi \tau},$$

and $\xi_t$ follows mean-reverting (or mean-fleeing) process:

$$d\xi_t = \kappa(\bar{\xi} - \xi_t)dt + \beta \xi_t d\widetilde{W}(2)^t.$$

Applying Itô’s lemma we find:

$$ds_t = \left(\frac{\beta^2 \gamma(1 + \gamma)}{2} \sigma_t^{2+\gamma-2\delta} + \kappa \gamma \sigma_t - \kappa \overline{\xi} \sigma_t^{1+\gamma} \right) dt + \beta \sigma_t^{1+\gamma-\delta} d\widetilde{W}(2)^t.$$

If $\gamma > 0$ and $1 \geq \delta$ then this process has the form (36). Let us also recall that Proposition 2 imposes a restriction on the vovol, which is satisfied if $\frac{1+\gamma-\delta}{\gamma} < 2$ or $\gamma + \delta > 1$. In the particular case of $\gamma = \delta = 1$ we have:

$$ds_t = \left((\beta^2 + k) \sigma_t - \kappa \overline{\xi} \sigma_t^2 \right) dt + \beta \sigma_t d\widetilde{W}(2)^t.$$

Setting $\bar{\xi} = 0$ we arrive at the model introduced by Hull and White (1987):

$$ds_t = (\beta^2 + k) \sigma_t dt + \beta \sigma_t d\widetilde{W}(2)^t.$$

Its version with $\kappa = -\beta^2$ and local volatility is studied in Hagan et al. (2002) under the name of SABR. In the next section we consider the accuracy of our method in the framework of SABR model, which is inspired by its popularity in practice.

6.5 SABR model

Let us consider SABR model, which is a particular case of (6):

$$dX_t = -\frac{1}{2} \sigma_t^2 (X_t)^{2\theta} + \sigma_t (X_t)^{\theta} dW_t^{(1)},$$

$$ds_t = (\beta^2 + k) \sigma_t dt + \beta \sigma_t dW_t^{(2)}.$$

In this section we will study its log-normal version with $\theta = 0$. Using the result of the previous section we construct the following approximation of implied volatilities:

$$I(X, \sigma, \tau) \simeq \sigma \zeta_\tau / \ln \left( \frac{\zeta_\tau - \rho + \sqrt{1 - 2 \rho \zeta_\tau + \zeta_\tau^2}}{1 - \rho} \right) \Pi(\tau),$$

(37)
where
\[ \zeta = \frac{\beta}{\sigma} \left[ \Lambda(\tau) X \right]. \] (38)

In particular, we have:
\[
A(1, 0) = \frac{\sigma \zeta}{\ln (\zeta - \rho + \sqrt{1 - 2\rho\zeta + \zeta^2})} \\
\times \left[ 1 + \left( \frac{\beta \rho \sigma}{4} - \frac{\beta^2(3\rho^2 - 2)}{24} \right) \tau \right],
\]
with \[ \zeta = \frac{\beta}{\sigma} X. \]

Let us observe that \( A(1, 0) \) coincides with the approximation obtained in Hagan et al. (2002) using a perturbation approach.

In Table 2 we compare the performance of different approximations using the model parameters from Hagan et al. (2002). The HKL approximation refers to the one derived in Hagan et al. (2002), \( A(k, k) \) stands for the approximation (37) with \( \Lambda \) and \( \Pi \) substituted by their short term asymptotics up to \( \tau^k \). As it is evident from the Table these asymptotics converge relatively quickly, and it is sufficient to choose \( k = 2 \).

Although the difference between approximations shown in Table 2 is negligible for \( K \approx S \), our approach seems to perform better than HKL for \( K \ll S \). Let us note that when the pricing of European options is concerned all the approximations seem to be equally accurate. Indeed, in practice FFM options are rarely priced in the market. The situation is different when we are concerned with pricing of more complex contingent claims. In this case the conditional probability density of the underlying price \( S_T(T > t) \) should be accurately evaluated also around its boundary \( S_T \approx 0 \).

Let us recall that the conditional price density:
\[
p(s|S_t, \sigma_t, t, T) \equiv \frac{\partial}{\partial \delta} \Pr \{ S_T \in (s, s + \delta)|S_t, \sigma_t, t \}
\]
equals to the second \( K \)-derivative of the European call option price evaluated at \( K = s \). Recalling the definition of implied volatility, we obtain:
\[
p(s|S_t, \sigma_t, t, T) = \frac{\partial^2}{\partial K^2} C_t(K, \tau)|_{K=s}
\]
\[= \left\{ C_{BS}^2(I_K)^2 + 2C_{BS}^2 I_K + C_{BS}^2 I_{KK} + C_{BS}^2 \right\} |_{K=s},(39)\]
where $C^{BS}$ and $I$ are shortcuts for $C^{BS}(X_t, K, \tau, I)$ and $I(\log(S_t/K), K, \tau, \sigma)$ respectively\textsuperscript{7}. From (39) we observe that the accuracy of approximation of the density at $s \approx 0$ clearly depends on the accuracy of the approximation of implied volatility and its $K-$derivatives at $K \approx 0 < S$.

In Figure 5 we plot the approximations of the price density obtained with the help of HKL and $A(2, 2)$, and its sample estimate based on Monte Carlo simulations. The model parameters are such that HKL produces negative estimates of the density around the boundary. Our approximation is clearly more accurate being positive everywhere. We have also experimented with higher order approximations, and discovered that they do not improve the accuracy. Consequently, under these model parameters the asymptotics of $\Lambda$ and $\Pi$ possibly do not converge.

The numerical experiments of this section show that approximation $A(2, 2)$ is preferable for SABR model. Indeed, in all cases $A(2, 2)$ seems to perform better than HKL. If the short term asymptotics of $\Lambda$ and $\Pi$ converge then the use of their higher order expansions provide only minor improvements. On the other hand, if these asymptotics do not converge then higher order approximations may appear to be even less accurate.

### 7 Concluding remarks

In this paper we assume a two-factor stochastic volatility model. However, some results may be generalized to a multifactor setting. In particular an algorithm for computing the power series expansion (12) can be derived for any continuous finite factor model. When jumps are introduced the power series expansion of $I$ does not converge for $(X, \tau) \neq (0, 0)$. Nevertheless, if jumps are of finite variation then the short term asymptotics of $i$ is still well defined but difficult to solve for (see Medvedev and Scaillet (2004)).

The method of computing implied volatilities for $\tau \approx 0$ via the transformation of the implied volatility smile at $\tau = 0$ is applied to models with CEV volatility of volatility. In the paper we describe the type of stochastic volatility models for which this approach might be accurate even for $\tau >> 0$. This class includes well-known Hull and White (1987) and SABR models, and is sufficiently wide to incorporate mean-reverting volatility. One area for further research is to study the performance of these models on the market data, and apply the method proposed in the paper to compute implied volatilities.

\textsuperscript{7}When computing these derivatives we should take into account that $f$ also depends on $K$ as suggested by (7). To emphasize this we write $K$ in the list of arguments for $I$. 

21
References


Appendix A. Proof of Proposition 1

The fact that ATM implied volatilities converge to the spot volatility allows us to write \( i_0 = \sigma F(K) \). To find the elements in the expansion (20) we formally should substitute it in (18), find the short term expansion of \( \Phi(i) \) and equate each element to zero. Let us consider the structure of the \( m \)-th element of this expansion denoted by \( \Phi_m(i) \). This element can be found by substitution the part of \( i \) up to \( m \)-th order denoted by \( i^{(m)} \). Let us note that we have to substitute \( i^{(m)} \) only in the following part of \( \Phi \):

\[
\Psi(h) = \frac{h}{2} - \frac{\sigma^2 f'^2}{2h} - \frac{h x x - \sigma^2 f'^2}{2h^2} + \frac{\sigma^2 f'^2}{h^2} x h + \frac{1}{2} h \sqrt{\tau},
\]

That is:

\[
\Phi_m(i) = \Psi_m(i^{(m)}) + (\Phi - \Psi)_m(i^{(m-1)})
\]

\[
= \Psi_m(i^{(m)}) - \Psi_m(i^{(m-1)}) + \Phi_m(i^{(m-1)}).
\]

The difference \( \Psi_m(i^{(m)}) - \Psi_m(i^{(m-1)}) \) can be found from (40). For example:

\[
\left\{ \frac{\sigma^2 f'^2}{2i^{(m)}} \right\}_m = \left\{ \frac{\sigma^2 f'^2}{2 \left( i^{(m-1)} + i_m(\sqrt{\tau})^m \right)} \right\}_m
\]

\[
= \left\{ \frac{\sigma^2 f'^2}{2 \left[ i^{(m-1)} \right]^2} (i^{(m-1)} - i_m(\sqrt{\tau})^m) \right\}_m
\]

\[
= \left\{ \frac{\sigma^2 f'^2}{2i^{(m-1)}} \right\}_m - \left\{ \frac{\sigma^2 f'^2}{2 \left[ i^{(m-1)} \right]^2} \right\}_0 i_m
\]

\[
= \left\{ \frac{\sigma^2 f'^2}{2i^{(m-1)}} \right\}_m - \frac{i_m}{2}.
\]

Proceeding in the similar fashion we will find:

\[
\Psi_m(i^{(m)}) - \Psi_m(i^{(m-1)}) = \left(1 + \frac{m}{2}\right) i_m + \frac{1}{2} \frac{d i_m}{d x} + \frac{\sigma^2 f'^2(0)}{2} \frac{d^2 i_m}{d x^2}.
\]

Equalizing \( \Phi_m(i) \) to zero using (41) we arrive at:

\[
- \left(1 + \frac{m}{2}\right) i_m - \frac{1}{2} \frac{d i_m}{d x} + \frac{\sigma^2 f'^2(0)}{2} \frac{d^2 i_m}{d x^2} = \Phi_m(i^{(m-1)}).
\]

Let us recall that \( i_m \) is polynomial in \( x \):

\[
i_m(x) = i_m x^m + i_{m-2} x^{m-2} + \ldots + i_{m-2M} x^{m-2M}.
\]
as well as $\Phi_{m}(i^{(m-1)})$ and their dimension should necessarily match:

$$
\Phi_{m}(i^{(m-1)}) (x) = \Phi_{mm}(i^{(m-1)}) x^m + \Phi_{mm-1}(i^{(m-1)}) x \\
+ \ldots + \Phi_{mm-2M}(i^{(m-1)}) x^{m-2M}.
$$

(44)

Now the recursive relationship in Proposition 1 easily follows by solving (42) taking into account (43) and (44).
Appendix B. Proof of Proposition 2

Let us substitute $g = 1/\Upsilon$ and find the PDE in terms of $g$:

$$\sigma^2 g^2 - 1 + 2X \sigma f g (\rho bg_\sigma + \sigma g_X) + X^2 (b^2 g_\sigma^2 + \sigma^2 f^2 g_X^2 + 2\rho b \sigma f g_X g_\sigma).$$

Let us further rearrange the expression:

$$(\sigma f g + X \sigma f g_X)^2 + 2\rho b X g_\sigma (\sigma f g + X \sigma g_X) + b^2 X^2 g_\sigma^2 = 1.$$ \hspace{1cm} (45)

Now an obvious substitution follows $p = Xg$, which allows us to obtain a compact form:

$$\sigma^2 f^2 p_X^2 + 2\rho b \sigma f p_X p_\sigma + b^2 p_\sigma^2 = 1.$$ \hspace{1cm} (46)

Note that from the regularity assumption we have:

$$p(X, \sigma) = \frac{X}{Y(X, \sigma)} = \frac{X}{\sigma f(0) + X(... + ...},$$

meaning that the correct solution of (46) should satisfy:

$$p(0, \sigma) = 0, \quad p_X(0, \sigma) = \frac{1}{\sigma f(0)}.$$ \hspace{1cm} (47)

The PDE (46) can be further simplified by getting rid of the middle term and function $f$. This can be accomplished by introducing a new variable $y = \int_0^X \frac{ds}{f(s)} - r(\sigma)$. Assuming $p(x, \sigma) = \tilde{p}(y, \sigma)$ from the equation (46) we have:

$$\sigma^2 \tilde{p}_y^2 + 2\rho b \sigma \tilde{p}_y (-\tilde{p}_y r' + \tilde{p}_\sigma) + b^2 (-\tilde{p}_y r' + \tilde{p}_\sigma)^2 = 1,$$

or

$$(\sigma^2 - 2\rho b \sigma r' + b^2 r'^2) \tilde{p}_y^2 + (2\rho b \sigma - 2b^2 r') \tilde{p}_y \tilde{p}_\sigma + b^2 \tilde{p}_\sigma^2 = 1.$$ \hspace{1cm} (48)

and initial conditions:

$$\tilde{p}(0, \sigma) = 0, \quad \tilde{p}_y(0, \sigma) = \frac{1}{\sigma}.$$ \hspace{1cm} (49)

Case 1 ($|\rho| = 1$). Let us set $r = 0$, and first consider the case $\rho = 1$. Then from (47) we find:

$$(\sigma \tilde{p}_y + b \tilde{p}_\sigma)^2 = 1.$$ \hspace{1cm} (50)

The form of the solution (46) suggests that we can write:

$$\sigma \tilde{p}_y + b \tilde{p}_\sigma = 1.$$ \hspace{1cm} (51)
Now let us guess that \( \hat{p} \) can be written as:

\[
\hat{p}(y, \sigma) = \hat{p}_1(y, \sigma) + \hat{p}_2(\sigma).
\]  

(48)

with:

\[
b \frac{d\hat{p}_2}{d\sigma} = 1,
\]

and

\[
\frac{\partial \hat{p}_1}{\partial y} + b \frac{\partial \hat{p}_1}{\partial \sigma} = 0,
\]  

(49)

The solution to the first equation:

\[
\hat{p}_2(\sigma) = \int b^{-1}(s)ds = \frac{\sigma^{1-\nu}}{\beta(1-\nu)} + C_0.
\]

(50)

The solution to the second equation we look in the form:

\[
\hat{p}_1(y, \sigma) = f(h_1(y) + h_2(\sigma)),
\]

(51)

From (49) we have:

\[
\frac{dh_1}{dy} + b \frac{dh_2}{d\sigma} = 0.
\]

It immediately follows that:

\[
h_1(y) = C_1y + C_2,
\]

(52)

and

\[
h_2(\sigma) = -C_1 \int sb^{-1}(s)ds = -C_1 \frac{\sigma^{2-\nu}}{(2-\nu)\beta} + C_3.
\]

(53)

Let us now put together (48), (50), (51), (52) and (53), and drop redundant constants:

\[
\hat{p}(y, \sigma) = \frac{\sigma^{1-\nu}}{\beta(1-\nu)} + f\left(-\frac{\sigma^{2-\nu}}{(2-\nu)\beta} + y\right).
\]

Recalling the initial condition \( \hat{p}(0, \sigma) = 0 \) we have:

\[
0 = \frac{\sigma^{1-\nu}}{\beta(1-\nu)} + f\left(-\frac{\sigma^{2-\nu}}{(2-\nu)\beta}\right),
\]

or

\[
f(s) = -\frac{(-s(2-\nu))^{\frac{1-\nu}{2-\nu}}}{\beta(2-\nu)}.
\]
Now the correct solution is:

$$\hat{p}(y, \sigma) = \frac{\sigma^{1-\nu}}{\beta(1-\nu)} \left[ 1 - \left( 1 - \frac{\beta(2-\nu)y}{\sigma^{2-\nu}} \right)^{\frac{1}{2-\nu}} \right].$$

The expression for $\Upsilon$ can be compactly written:

$$\Upsilon(X, \sigma) = \sigma(1-\nu)X \zeta \left\{ \int_0^X \frac{ds}{f(s)} \cdot \left[ 1 - (1 - \zeta(2-\nu))^{\frac{1}{2-\nu}} \right] \right\}^{-1},$$

by denoting:

$$\zeta = \frac{\beta}{\sigma^{2-\nu}} \int_0^X ds \cdot f(s).$$

The case $\rho = -1$ can be clearly reduced to $\rho = 1$ by substitution of $\beta$ with $-\beta$.

**Case 2 ($$\rho \neq 1$$).** Let us choose $r = \rho \int^\sigma \frac{sds}{b(s)}$ and denote $\sigma = \sigma \sqrt{1-\rho^2}$ and $b = b \sqrt{1-\rho^2}$. Then we arrive at:

$$\sigma^2 \hat{p}_\sigma + b^2 \hat{p}_\sigma = 1.$$  (54)

with initial conditions:

$$\hat{p} \left( -\rho \int^\sigma \frac{sds}{b(s)} \right) = 0, \quad \hat{p}_\sigma \left( -\rho \int^\sigma \frac{sds}{b(s)} \right) = \frac{1}{\sigma f(0)}.$$  (55)

To get rid of squared derivatives we introduce a new function $\omega(y, \sigma)$ such that:

$$\sigma \hat{p}_\sigma = \cos \omega, \quad \hat{p}_\sigma = \sin \omega.$$  

with the initial condition

$$\omega \left( -\rho \int^\sigma \frac{sds}{b(s)} \right) = \arccos \left( \sqrt{1-\rho^2} \right).$$

Let us use the obvious relationship $\hat{p}_\sigma = \hat{p}_y$ to find a PDE for $\omega$. Indeed, we have:

$$\hat{p}_y = -\frac{\omega \sin \omega}{\sigma} - \frac{\cos \omega}{\sigma^2}, \quad \hat{p}_\sigma = \frac{\omega \cos \omega}{b}.$$
and the PDE immediately follows:

\[
\frac{\omega_y}{\bar{\sigma}} + \frac{1}{\bar{\sigma}^2} + \frac{\omega \tan \omega}{\bar{\sigma}} = 0.
\] (56)

Now let us assume a CEV type of vovol \( \bar{b}(\sigma) = \beta \sigma^\nu \), where \( \beta = \beta (1 - \rho^2)^{(1-\nu)/2} \), and guess that the right solution to (56) has the form:

\[
\omega(y, \sigma) = \phi(z), \quad \text{with} \quad z = \frac{\beta y}{\bar{\sigma}^\nu}.
\] (57)

After substitution (57) into (56) and subsequent cancellation we find:

\[
\frac{\phi'}{\sigma^{\nu + \delta}} + \frac{1}{\sigma^2} - \frac{\delta \phi' \tan \phi}{\sigma^2} = 0.
\]

It immediately follows that \( \delta = 2 - \nu \). Let us assume in the following that \( \gamma < 2 \). Given our assumption about the form of the vovol we have:

\[
-\rho \int_0^\sigma ds b(s) = -\frac{\rho a^{2-\nu}}{\beta (2-\nu)}.
\]

Then \( \phi(z) \) satisfies:

\[
\phi' + 1 - (2 - \nu) z \phi' \tan \phi = 0,
\] (58)

with initial condition:

\[
\phi \left( -\frac{\rho}{(2-\nu) \sqrt{1 - \rho^2}} \right) = \arccos \left( \sqrt{1 - \rho^2} \right).
\]

Function \( p(y, \sigma) \) can be obtained by integrating \( \cos \omega \) with respect to \( y \) and using the first initial condition in (55):

\[
\hat{p}(y, \sigma) = \frac{1}{\sigma} \int_0^y \cos \phi \left( \frac{\beta \eta}{\sigma^{2-\nu}} \right) d\eta.
\]

Recalling the change of variables \( y = \int_0^X \frac{ds}{f(s)} - r(\sigma) = \int_0^X \frac{ds}{f(s)} - \frac{\rho a^{2-\nu}}{\beta (2-\nu)} \) and changing the integration variable we have:

\[
p(x, \sigma) = \frac{x^{1-\nu}}{\beta} \int_0^x \frac{ds}{f(s)} - \frac{\rho}{\beta (2-\nu) \sqrt{1 - \rho^2}} \int_0^X \cos \phi(s) ds.
\]
From (58) it is straightforward to derive the PDE for $Q(s) = \cos \phi(s)$:

\[
Q' \left( \frac{(2 - \nu)s}{Q} + \frac{1}{\sqrt{1 - Q^2}} \right) + 1 = 0, \quad Q \left( -\frac{\rho}{(2 - \nu)\sqrt{1 - \rho^2}} \right) = \sqrt{1 - \rho^2}.
\]

To get rid of radicals we set $Q = \frac{1}{\sqrt{1 + G^2}}$ and obtain:

\[
G' = \frac{1 + G^2}{(2 - \nu)sG + 1}, \quad G \left( -\frac{\rho}{(2 - \nu)\sqrt{1 - \rho^2}} \right) = \frac{\rho}{\sqrt{1 - \rho^2}}.
\]

Finally we have

\[
p(x, \sigma) = \frac{X_0}{\sigma \zeta} \int_{-\frac{\rho}{(2 - \nu)\sqrt{1 - \rho^2}}}^{\frac{\rho}{(2 - \nu)\sqrt{1 - \rho^2}}} \frac{1}{\sqrt{1 + G^2(s)}} ds,
\]

and, consequently,

\[
\Upsilon(x, \sigma) = \sigma X \zeta \left[ \int_{0}^{X} ds f(s) \cdot \int_{-\frac{\rho}{(2 - \nu)\sqrt{1 - \rho^2}}}^{\frac{\rho}{(2 - \nu)\sqrt{1 - \rho^2}}} \frac{1}{\sqrt{1 + G^2(s)}} ds \right]^{-1}
\]

where $\zeta = \frac{\beta}{\sigma^{\alpha - \nu}} \int_{0}^{X} ds f(s)$.
<table>
<thead>
<tr>
<th>Authors</th>
<th>Model</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berestycki-Busca-Florent (2004)</td>
<td>Local volatility</td>
<td>Short term asymptotics</td>
</tr>
<tr>
<td>Berestycki-Busca-Florent (2002)</td>
<td>Local and stochastic volatility</td>
<td>Short term asymptotics</td>
</tr>
<tr>
<td>Fouque-Papanicolaou-Sircar (2000)</td>
<td>Stochastic volatility</td>
<td>Fast variation of volatility</td>
</tr>
<tr>
<td>Lee (2001)</td>
<td>Stochastic volatility</td>
<td>Slow variation of volatility</td>
</tr>
<tr>
<td>Lee (2001)</td>
<td>Stochastic volatility</td>
<td>Small variation of volatility</td>
</tr>
<tr>
<td>Lewis (2000)</td>
<td>Stochastic volatility</td>
<td>Long term asymptotics</td>
</tr>
</tbody>
</table>

Table 1. Approximations of implied volatilities
<table>
<thead>
<tr>
<th>K/S</th>
<th>Actual</th>
<th>HKL</th>
<th>A(1,1)</th>
<th>A(2,2)</th>
<th>A(3,3)</th>
<th>A(4,4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.</td>
<td></td>
<td>(\beta = 0.38, \rho = -0.19, \sigma = 0.15, \tau = 3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>0.234</td>
<td>0.241</td>
<td>0.231</td>
<td>0.232</td>
<td>0.232</td>
<td>0.232</td>
</tr>
<tr>
<td>0.6</td>
<td>0.193</td>
<td>0.197</td>
<td>0.192</td>
<td>0.192</td>
<td>0.192</td>
<td>0.192</td>
</tr>
<tr>
<td>0.8</td>
<td>0.167</td>
<td>0.168</td>
<td>0.167</td>
<td>0.166</td>
<td>0.166</td>
<td>0.166</td>
</tr>
<tr>
<td>1</td>
<td>0.153</td>
<td>0.153</td>
<td>0.154</td>
<td>0.153</td>
<td>0.153</td>
<td>0.153</td>
</tr>
<tr>
<td>1.2</td>
<td>0.152</td>
<td>0.152</td>
<td>0.152</td>
<td>0.152</td>
<td>0.152</td>
<td>0.152</td>
</tr>
<tr>
<td>1.4</td>
<td>0.158</td>
<td>0.159</td>
<td>0.157</td>
<td>0.157</td>
<td>0.157</td>
<td>0.157</td>
</tr>
<tr>
<td>1.6</td>
<td>0.167</td>
<td>0.168</td>
<td>0.165</td>
<td>0.165</td>
<td>0.165</td>
<td>0.165</td>
</tr>
<tr>
<td>B.</td>
<td></td>
<td>(\beta = 0.33, \rho = -0.22, \sigma = 0.15, \tau = 5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>0.223</td>
<td>0.231</td>
<td>0.220</td>
<td>0.221</td>
<td>0.220</td>
<td>0.220</td>
</tr>
<tr>
<td>0.6</td>
<td>0.187</td>
<td>0.192</td>
<td>0.186</td>
<td>0.186</td>
<td>0.186</td>
<td>0.186</td>
</tr>
<tr>
<td>0.8</td>
<td>0.165</td>
<td>0.167</td>
<td>0.165</td>
<td>0.164</td>
<td>0.164</td>
<td>0.164</td>
</tr>
<tr>
<td>1</td>
<td>0.153</td>
<td>0.154</td>
<td>0.153</td>
<td>0.153</td>
<td>0.153</td>
<td>0.153</td>
</tr>
<tr>
<td>1.2</td>
<td>0.151</td>
<td>0.152</td>
<td>0.151</td>
<td>0.150</td>
<td>0.150</td>
<td>0.150</td>
</tr>
<tr>
<td>1.4</td>
<td>0.154</td>
<td>0.155</td>
<td>0.154</td>
<td>0.153</td>
<td>0.153</td>
<td>0.153</td>
</tr>
<tr>
<td>1.6</td>
<td>0.160</td>
<td>0.162</td>
<td>0.159</td>
<td>0.158</td>
<td>0.158</td>
<td>0.158</td>
</tr>
<tr>
<td>C.</td>
<td></td>
<td>(\beta = 0.25, \rho = -0.34, \sigma = 0.15, \tau = 10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>0.204</td>
<td>0.214</td>
<td>0.204</td>
<td>0.203</td>
<td>0.203</td>
<td>0.203</td>
</tr>
<tr>
<td>0.6</td>
<td>0.177</td>
<td>0.183</td>
<td>0.178</td>
<td>0.177</td>
<td>0.176</td>
<td>0.176</td>
</tr>
<tr>
<td>0.8</td>
<td>0.160</td>
<td>0.164</td>
<td>0.162</td>
<td>0.160</td>
<td>0.159</td>
<td>0.159</td>
</tr>
<tr>
<td>1</td>
<td>0.149</td>
<td>0.152</td>
<td>0.152</td>
<td>0.150</td>
<td>0.149</td>
<td>0.149</td>
</tr>
<tr>
<td>1.2</td>
<td>0.144</td>
<td>0.146</td>
<td>0.147</td>
<td>0.144</td>
<td>0.144</td>
<td>0.144</td>
</tr>
<tr>
<td>1.4</td>
<td>0.143</td>
<td>0.145</td>
<td>0.145</td>
<td>0.143</td>
<td>0.142</td>
<td>0.142</td>
</tr>
<tr>
<td>1.6</td>
<td>0.145</td>
<td>0.147</td>
<td>0.146</td>
<td>0.144</td>
<td>0.143</td>
<td>0.143</td>
</tr>
</tbody>
</table>

**Table 2. Implied volatilities and their approximations under log-normal version of SABR model.**

HKL refers to the approximation obtained in Hagan et al. (2002). A(k,k) is given by \(Y(\Lambda(\tau)) \cdot \Pi(\tau)\) with \(\Pi\) and \(\Lambda\) substituted by their expansions up to \(\tau^k\). In particular, A(1,0) coincides with HKL.
The order of expansion $N$ means that the asymptotics of $i$ is built up to $t^{N/2}$.

Figure 1 Time required for building the short term asymptotics of $i$. The expansion was constructed by running a MAPLE code in Pentium IV, 2.66 MG, 512 RAM.
Figure 2 Implied volatilities of 1-month options and their series expansions.

Here the series expansion of $I$ corresponds to the short term asymptotics of $i$ up to $\tau^{9/2}$.
Figure 3 Implied volatilities of 1-month options and their approximations based on the transformation of $\Upsilon$. $A(k,k)$ is given by $\Upsilon(X\Lambda(\tau)) \cdot \Pi(\tau)$ with $\Pi$ and $\Lambda$ substituted by their expansions up to $\tau^k$. 

(a) Heston model: $d\sigma_t^2 = 5.0(0.045 - \sigma_t^2) \cdot dt + 0.48\sigma_t \cdot (-0.77dW_t^{(1)} + 0.64dW_t^{(2)})$

(b) GARCH model: $d\sigma_t^2 = 1.6(0.075 - \sigma_t^2) \cdot dt + 2.2\sigma_t^2 \cdot (-0.75dW_t^{(1)} + 0.66dW_t^{(2)})$

(c) CEV model: $d\sigma_t^2 = 0.69(0.052 - \sigma_t^2) \cdot dt + 1.38\sigma_t^{0.85} \cdot (-0.77dW_t^{(1)} + 0.64dW_t^{(2)})$
\( \sigma = 0.1 \)

(a) Heston model: 
\[
d\sigma_t^2 = 5.0(0.045 - \sigma_t^2) \cdot dt + 0.48\sigma_t \cdot \left( -0.77dW_t^{(1)} + 0.64dW_t^{(2)} \right)
\]

\( \sigma = 0.2 \)

(b) GARCH model: 
\[
d\sigma_t^2 = 1.6(0.075 - \sigma_t^2) \cdot dt + 2.2\sigma_t^2 \cdot \left( -0.75dW_t^{(1)} + 0.66dW_t^{(2)} \right)
\]

\( \sigma = 0.3 \)

(c) CEV model: 
\[
d\sigma_t^2 = 0.69(0.052 - \sigma_t^2) \cdot dt + 1.38\sigma_t^{0.85} \cdot \left( -0.77dW_t^{(1)} + 0.64dW_t^{(2)} \right)
\]

Figure 4 Implied volatilities of 2-month options and their approximations based on the transformation of \( \Upsilon \).

\( \Lambda(k,k) \) is given by \( \Upsilon(X\Lambda(\tau)) \cdot \Pi(\tau) \) with \( \Pi \) and \( \Lambda \) substituted by their expansions up to \( \tau^k \).
Figure 5. Comparison of approximations of the price density.
The model is the lognormal version of SABR with $\beta = 1$, $\rho = -0.7$, $\sigma = 0.3$, $\tau = 3$. HKL refers to the approximation obtained in Hagan et al. (2002). $A(2,2)$ is given by $Y(X\Lambda(\tau)) \cdot \Pi(\tau)$ with $\Pi$ and $\Lambda$ substituted by their expansions up to $\tau^2$. 