The heat-kernel most-likely-path approximation

Jim Gatheral*, Tai-Ho Wang†‡

This Version: September 22, 2011

Abstract

In this article, we derive a new most-likely-path (MLP) approximation for implied volatility in terms of local volatility, based on time-integration of the lowest order term in the heat-kernel expansion. This new approximation formula turns out to be a natural extension of the well-known formula of Berestycki, Busca and Florent. Various other MLP approximations have been suggested in the literature involving different choices of most-likely-path; our work fixes a natural definition of the most-likely-path. We confirm the improved performance of our new approximation relative to existing approximations in an explicit computation using a realistic S&P500 local volatility function.

1 Introduction

Derivatives practitioners face two key problems that involve the parameterization of prices of vanilla options. The first problem is that, given the prices of vanilla options for a discrete set of strikes and expirations, the practitioner would like to be able to price options with other strikes and expirations. For this, the observed option prices need to be interpolated and extrapolated. This is typically achieved in practice by parameterizing the implied volatility surface; there is then a deterministic mapping (via the Black-Scholes formula) between implied volatilities returned by the parametric formula and option prices. The main issue here is that nobody has yet

*Department of Mathematics, Baruch College, CUNY. jim.gatheral@baruch.cuny.edu

†Department of Mathematics, Baruch College, CUNY. tai-ho.wang@baruch.cuny.edu

‡We are very grateful to Julien Guyon and Vladimir Lucic for sharing their results and for valuable discussions. The first author also thanks Nick Constanzino and his collaborators at Penn State University for sending an early version of their paper, inspiring one of our proofs. The second author is particularly grateful to Professor Gérard Ben Arous of Courant Institute for insightful and valuable comments and discussions. All errors are our responsibility.

devised an implied volatility surface parameterization that is arbitrage-free. More recently, there has been a trend to parameterize the Dupire local volatility surface directly – so long as local variance is positive, there can be no static arbitrage. The problem then becomes one of fast calibration, or how to quickly compute vanilla option prices given the parameterized local volatility function.

The second problem is how to price exotic options given the prices of vanilla options. In this case, the practitioner is aiming for fast calibration of a complex model (typically involving stochastic volatility) with realistic dynamics to given option prices. Through techniques such as Markovian projection, this problem can again be reduced to one of fast calibration of a local volatility surface to given option prices.

It is possible to generate an entire volatility surface by solving the Dupire forward PDE using numerical PDE techniques (see [1] for example). However, numerical PDE solutions are often too slow for fast calibration and in higher dimensions, completely impractical. Ideally, practitioners would like a more direct and faster method for generating the implied volatility surface from a parameterized local volatility function, even better if this method provides insight into the connection between local volatility and implied volatility functions, both of which encode the same information. The most-likely-path approach (as in [5], [8] and [9]) is such a method, offering both fast computation and natural intuition. Moreover, the most-likely-path approach has a natural and straightforward generalization to higher dimensions.

In the foregoing, as in [8], we will assume that the local volatility function is smooth enough to give an estimator of the most-likely-path type a chance to be competitive with a PDE solver. For example local volatility functions used by practitioners as well as local volatility functions generated from more sophisticated models (such as stochastic volatility models) typically fall into this class.

To make this precise, we assume the underlying asset S satisfies

$$dS_t = S_t\,\sigma_L(S_t,t)\,dW_t = a(S_t,t)\,dW_t.$$

under the risk-neutral probability measure (zero interest rates and dividends are assumed). Also, we denote the log-spot by $x_t := \log(S_t/S_0)$ in terms of which the local volatility function becomes

$$\sigma(x_t,t) := \sigma_L\left(S_0\,e^{x_t},t\right).$$

For notational simplicity, the terms a, σ_L and σ will be used interchangeably. We define the log-strike k of a European call or put option with strike K by $k := \log(K/S_0)$.

As is well-known, the market price of a call option with strike K and expiration T is often quoted in terms of the Black-Scholes implied volatility $\sigma_{\rm BS}(K,T)$ defined through the implicit relation:

$$C_{\rm BS}(K,\sigma_{\rm BS}(K,T),T) = \mathbb{E}\left[(S_T-K)^+\right],$$

where $C_{\rm BS}(\cdot)$ denotes the Black-Scholes formula and the expectation is taken under the risk-neutral measure.

Our problem then is to find an approximate expression for the implied volatility $\sigma_{\mathrm{BS}}(K,T)$ in terms of the local volatility function $\sigma(x_t,t)$.

Local volatilities can be computed directly from option prices using the Dupire equation (see [5] for example). Re-expressing option prices in terms of the Black-Scholes formula and implied volatility, with $w(k,T) := \sigma_{\mathrm{BS}}(k,T)^2\,T$, the following simple formula follows directly from the Dupire equation:

$$\sigma^2(k,T) = \frac{\frac{\partial w}{\partial T}}{1 - \frac{k}{w}\frac{\partial w}{\partial k} + \frac{1}{4}\left(-\frac{1}{4} - \frac{1}{w} + \frac{k^2}{w^2}\right)\left(\frac{\partial w}{\partial k}\right)^2 + \frac{1}{2}\frac{\partial^2 w}{\partial k^2}}. \tag{1.1}$$

If it were possible to invert (1.1) exactly, our problem would be solved. In the present paper, we proceed as follows:

In Section 2, we use the Chapman-Kolmogorov equation to express the transition density as a convolution of transition densities over small time-steps, each of which is approximated by the lowest order term in the heat-kernel expansion. The resulting multidimensional integral is then approximated using Laplace's method which gives rise to a variational boundary value problem. The solution to this problem, termed the variational most-likely-path, leads to a simple approximate expression for the European call price. By matching exponents of this approximate European option price with the exponent in the corresponding Black-Scholes expression, we deduce an approximate expression for Black-Scholes implied volatility in terms of local volatility. In Section 3, we show how to solve the variational problem using a fixed-point iterative procedure. We further show that our approximation is a natural extension of the well-known formula due to Berestycki, Busca and Florent, an extension that moreover behaves correctly under a deterministic time-change. In Section 4, we use our approximation to compute implied volatility smiles in a realistic local volatility model. The accuracy of our new method is then demonstrated through comparison with numerical PDE and other existing approximations. Finally, in Section 5, we summarize and conclude.

1.1 Prior literature

In [2], Berestycki, Busca and Florent (BBF) derived an approximate formula for implied volatility in terms of the local volatility by approximating (1.1) for short times and explicitly solving the resulting nonlinear PDE. In our notation, the BBF formula reads

$$\frac{1}{\sigma_{\mathrm{BS}}(k,T)} \approx \frac{1}{k}\int_0^k \frac{dy}{\sigma(y,0)}\,. \tag{1.2}$$

By a simple change of variable, the BBF formula may be re-expressed suggestively for our purposes as

$$\frac{1}{\sigma_{\mathrm{BS}}(k,T)} \approx \frac{1}{T}\int_0^T \frac{dt}{\sigma\left(k\,\frac{t}{T},0\right)}. \tag{1.3}$$

In words, the BBF formula approximates implied volatility as the harmonic mean of local volatility along the straight-line path from the log-spot $x_0 = 0$ to the log-strike $x_T = k$. Note however that on the right-hand side of (1.3), the local volatility function is "frozen" at time $t = 0$.

On the other hand, the following formula again from [5] (equation (3.5), page 28) is exact:

$$\sigma_{\mathrm{BS}}^2(k,T) = \frac{1}{T}\int_0^T \frac{\mathbb{E}[\sigma_t^2\, S_t^2\, \Gamma_{\mathrm{BS}}]}{\mathbb{E}[S_t^2\, \Gamma_{\mathrm{BS}}]} dt, \tag{1.4}$$

where Γ_{BS} is the Black-Scholes gamma, *i.e.* $\Gamma_{\mathrm{BS}} = \frac{\partial^2 C_{\mathrm{BS}}}{\partial s^2}$, and the instantaneous volatility σ_t is some arbitrary stochastic process (subject to some general technical conditions). For the details of the derivation of (1.4), we refer the reader to Chapter 3 of [5]. For the purposes of the present paper, we choose σ_t to be the deterministic local volatility function $\sigma(x_t, t)$, in which case the expectations in the integrand of (1.4) may be rewritten in terms of integrations over the transition density $p(\cdot)$ for S_t as

$$\frac{\mathbb{E}[\sigma_t^2\, S_t^2\, \Gamma_{\mathrm{BS}}]}{\mathbb{E}[S_t^2\, \Gamma_{\mathrm{BS}}]} = \int \sigma^2(x,t)\, q(x,t;k,T)\, dx,$$

where, as explained in detail in [5],

$$q(x,t;k,T) = \frac{p(t,x)\, S_t^2\, \Gamma_{\mathrm{BS}}}{\mathbb{E}[S_t^2\, \Gamma_{\mathrm{BS}}]}$$

is the density of a pinned diffusion that peaks on a line joining the stock price at inception to the strike price at expiration. The problem with equation (1.4) is that though it is exact, it is implicit in the sense that the transition densities $p(\cdot)$ and the Black-Scholes gammas Γ_{BS} themselves depend on the implied volatilities $\sigma_{\mathrm{BS}}(\cdot)$. This leads us to look for approximations.

The idea of the *most-likely-path* (MLP) approach is to approximate the expectation in (1.4) by evaluating the random process at a specific path $\tilde{x}(t)$, an approximation reminiscent of the maximum likelihood estimator. The most-likely-path implied volatility approximation is then obtained by averaging the local volatility function along the most likely path:

$$\sigma_{\mathrm{BS}}^2(k,T) \sim \frac{1}{T}\int_0^T \sigma^2(\tilde{x}(t),t) dt.$$

In the literature, there is no agreement on the best choice of most-likely-path. To assess the relative accuracy of our new approximation, we will focus on two particular versions of the most likely path approximation: A popular choice in [9] due to Reghai and a naïve extension of equation (1.3).

1.2 Adil Reghai's most-likely-path approximation

Reghai's approximation in [9] may be written as follows:

$$\sigma_{\mathrm{BS}}^2(k,T) = \frac{1}{T}\int_0^T \sigma^2(\tilde{x}(t),t)\,dt \tag{1.5}$$

with

$$\tilde{x}(t) = \frac{w(t)}{w(T)}\,k + \frac{w(t)}{2}\left(1 - \frac{w(t)}{w(T)}\right)$$

and

$$w(t) = \int_0^t \sigma^2(\tilde{x}(s),s)\,ds\,.$$

In this case, the distribution of the underlying at time t is approximated as lognormal so that the most-likely-path $\tilde{x}(t)$ is the mode of the normal process followed by the log-spot pinned at $(0,0)$ and (k,T) with time-dependent volatility $\sigma(\tilde{x}(t),t)$. This definition of most-likely-path is still implicit but the fixed-point algorithm described in [9] for computing $\tilde{x}(t)$ starting from an initial guess (for example, the straight line $x_0(t) = k\,t/T$ from log-spot to log-strike) typically converges within very few iterations. Again, the approximation made here is that the underlying is lognormal, which it is clearly not in general.

1.3 A naïve extension of the BBF formula

A natural extension of (1.3) that also involves integration along a path from the current spot at inception to the strike at expiration is the following:

$$\frac{1}{\sigma_{\mathrm{BS}}(k,T)} \approx \frac{1}{T}\int_0^T \frac{dt}{\sigma\left(k\,\frac{t}{T},t\right)} = \int_0^1 \frac{d\alpha}{\sigma\left(\alpha\,k,\alpha\,T\right)}. \tag{1.6}$$

In this approximation, which we will refer to as *BBFe*, implied volatility is approximated as the harmonic mean of local volatility along the straight-line path (in log-space) between the log-spot at inception $(0,0)$ and the log-strike at expiration (k,T). Note that all we have done here is to "unfreeze" the local volatility function which in (1.3) is "frozen" at time $t=0$, as previously discussed.

2 Most-likely-path from the heat kernel expansion

Recall that the underlying asset S follows the local volatility model (zero interest rates and dividends)

$$dS_t = S_t\,\sigma_L(S_t,t)dW_t = a(S_t,t)\,dW_t = S_t\,\sigma(x_t,t)\,dW_t. \tag{2.1}$$

Let $p(t, S_t; u, S_u)$, $t < u$, be the transition probability density associated with the local volatility model (2.1). The price C of a European call struck at K, with expiry at T is given by the expectation under risk neutral probability

$$C(s_0, K, T) = \mathbb{E}[(S_T - K)^+ | S_0 = s_0] = \int p(0, s_0, T, s)(s - K)^+ ds. \tag{2.2}$$

In Section 2 of [7], the authors use the heat kernel expansion for the transition density p and then apply Laplace's asymptotic formula to obtain an asymptotic expansion for the implied volatility in local volatility models for small time to maturity. In this section we follow the same line of argument but refine it a bit as follows. Instead of approximating the transition density at maturity T in one step as in the prior literature, we apply the Chapman-Kolmogorov equation to decompose the transition density from current time to maturity into many tiny steps. We then approximate the transition density for each such consecutive tiny step by the lowest order term in its corresponding heat kernel expansion. The price we pay for such a decomposition is that the expression (2.2) for the call price now involves a multidimensional integral. Laplace's asymptotic formula is applied to the resulting multidimensional integral to generate an asymptotic expansion for the call price.

To be more precise, we partition the time horizon $[0, T]$ into disjoint subintervals as $t_0 < t_1 < t_2 < \cdots < t_n$ with $t_0 = 0$ and $t_n = T$. For simplicity we further assume that the time intervals between the sampling slots are equal, i.e., $t_i - t_{i-1} = \Delta t = \frac{T}{n}$, for $i = 1, \cdots, n$. According to the Chapman-Kolmogorov equation, the transition density $p(0, S_0; T, S_T)$ satisfies

$$p(0, S_0; T, S_T) = \int \cdots \int p(0, S_0; t_1, S_1) p(t_1, S_1; t_2, S_2) \cdots p(t_{n-1}, S_{n-1}; T, S_T)\, dS_1 \cdots dS_{n-1}.$$

The expression (2.2) for the call price is then transformed into the following n-dimensional integral:

$$C = \int \cdots \int \left\{ \prod_{i=1}^{n-1} p(t_{i-1}, S_{i-1}; t_i, S_i)\, dS_i \right\} p(t_{n-1}, S_{n-1}; t_n, S_n)\, (S_n - K)^+\, dS_n. \tag{2.3}$$

The following heat kernel expansion for transition density p is well-known (see for instance [7] and the references therein): for $t < t'$,

$$p(t, x; t', y) = \frac{e^{-\frac{d^2(x,y,t)}{2(t'-t)}}}{\sqrt{2\pi(t'-t)}a(y, t')} \left(\sum_{i=0}^{k} H_i(t, x, y)(t' - t)^i \right) [1 + \mathcal{O}_{k+1}(t, x; t', y)],$$

where

$$
\begin{aligned}
d(x,y,t) &= \int_x^y \frac{d\xi}{a(\xi,t)},\\
H_0(t,x,y) &= \sqrt{\frac{a(x,t)}{a(y,t)}}e^{-\int_x^y \frac{d_t(\eta,y,t)}{a(\eta,t)}d\eta} = \sqrt{\frac{a(x,t)}{a(y,t)}}e^{\int_x^y\int_\eta^y \frac{a_t(\xi,t)}{a^2(\xi,t)a(\eta,t)}d\xi d\eta},\\
H_1(t,x,y) &= \frac{H_0(t,x,y)}{d(x,y,t)}\int_x^y \frac{1}{H_0(t,\eta,y)}\left[\frac{a^2}{2}\frac{\partial^2 H_0}{\partial x^2}+\frac{\partial H_0}{\partial t}\right]\frac{d\eta}{a(\eta,t)}.
\end{aligned}
$$

Here we note that $d(x,y,t)$ is the (signed) geodesic distance between x and y in the Riemannian metric induced by $\frac{1}{a(\cdot,t)}$ and the H_i's are called heat kernel coefficients. It is known that under certain technical conditions, the heat kernel expansion converges to the transition density p uniformly on every compact interval (see Section 3.4 in [7] for a treatment in one dimension). For our purposes, the following condition is imposed:

- $\sigma(x,t)$ is smooth and bounded with bounded derivatives in x,t. (2.4)

For our argument, we shall only need the first two terms H_0 and H_1. Moreover, under the conditions (2.4), the error terms $\mathcal{O}_k$, $k=1,2$, in the heat kernel expansion are bounded in the sense that, for any x and y in a compact set, and for any t and t' with $t<t'$, there exists a constant $A_k>0$ such that $|\mathcal{O}_k(t,x;t',y)|\le A_k(t'-t)^k$. Define, for $n>1$,

$$
\begin{aligned}
p_n^0(t,x;t',y) &= \frac{1}{(2\pi\Delta t)^{\frac{n}{2}}}\int\left(\prod_{i=1}^n \frac{e^{-\frac{d^2(x_{i-1},x_i,t_{i-1})}{2\Delta t}}}{a(x_i,t_i)}H_0(t_{i-1},x_{i-1},x_i)\right)\mathcal{D}x, \quad (2.5)\\
\hat{p}_n(t,x;t',y) &= \frac{1}{(2\pi\Delta t)^{\frac{n}{2}}}\int\left(\prod_{i=1}^n \frac{e^{-\frac{d^2(x_{i-1},x_i,t_{i-1})}{2\Delta t}}}{a(x_i,t_i)}\left[H_0(t_{i-1},x_{i-1},x_i)\right.\right.\\
&\qquad \left.\left.+H_1(t_{i-1},x_{i-1},x_i)\Delta t\right]\right)\mathcal{D}x, \quad (2.6)
\end{aligned}
$$

where $\mathcal{D}x=\prod_{i=1}^{n-1}dx_i$.

Remark 2.1. *Note that $\hat{p}_n$ in (2.6) is an analogue of p_n^0 in (2.5) but to one order higher in Δt. In Lemma 2.1, we prove the convergence of the call price approximation using $\hat{p}_n(\cdot)$ to the true call price given by the local volatility model (2.1). Our results in later sections use only H_0 and neglect H_1 and so are formally only to lowest order in T and have the same asymptotic order of convergence as the BBF approximation. However, for a realistic local volatility function such as that considered in Section 4, our most-likely-path approximation is very much superior to the BBF approximation.*

Now if we use (2.5) and (2.6) respectively to approximate the transition density, we obtain (different) approximations to the call price C. The following lemma permits us to characterize the asymptotic behavior of these two approximations as $n\to\infty$.

Lemma 2.1. *For $k = 0, 1$, let*

$$\widetilde{C}_k = \frac{1}{(2\pi\Delta t)^{\frac{n}{2}}} \int_{\Omega} e^{-\frac{D(\mathbf{S},\mathbf{t})}{\Delta t}} W_k(\mathbf{S}, \mathbf{t}) d\mathbf{S}, \tag{2.7}$$

where $\mathbf{t}$ *refers to the partition* $\mathbf{t} = \{0 = t_0 < t_1 < \cdots < t_n = T\}$, $\mathbf{S} = (S_1, \cdots, S_n)$, *and*

$$\begin{aligned}
D(\mathbf{S}, \mathbf{t}) &= \frac{1}{2} \sum_{i=1}^{n} d^2(S_{i-1}, S_i, t_{i-1}), \\
W_k(\mathbf{S}, \mathbf{t}) &= (S_n - K) \prod_{i=1}^{n} \frac{\widetilde{H}_k(t_{i-1}, t_i, S_{i-1}, S_i)}{a(S_i, t_i)}, \\
\widetilde{H}_0 &= H_0, \quad \widetilde{H}_1 = H_0 + H_1 \Delta t, \\
\Omega &= \{\mathbf{S} = (S_1, \cdots, S_n) : S_n \geq K\}.
\end{aligned}$$

If (2.5) *is used to approximate the transition density, the following bounds on the call price are obtained.*

$$[1 - A_1 \Delta t]^n \, \widetilde{C}_0 \leq C \leq [1 + A_1 \Delta t]^n \, \widetilde{C}_0 \,. \tag{2.8}$$

On the other hand, if (2.6) *is used to approximate the transition density, the following bounds on the call price are obtained*

$$\left[1 - A_2 (\Delta t)^2\right]^n \widetilde{C}_1 \leq C \leq \left[1 + A_2 (\Delta t)^2\right]^n \widetilde{C}_1 \,. \tag{2.9}$$

Proof. Since the proofs for both sets of inequalities in (2.9) are similar, we shall prove (2.8) only. Recall that since

$$p(t, x; t', y) = \frac{e^{-\frac{d^2(x,y,t)}{2(t'-t)}}}{\sqrt{2\pi(t'-t)}} \frac{H_0(t, x, y)}{a(y, t')} \left[1 + \mathcal{O}_1(t, x; t', y)\right],$$

we have the following upper and lower bounds for p.

$$\begin{aligned}
p(t_{i-1}, S_{i-1}; t_i, S_i) &\geq \frac{e^{-\frac{d^2(S_{i-1}, S_i, t_{i1})}{2\Delta t}}}{\sqrt{2\pi\Delta t}} \frac{H_0(t_{i-1}, S_{i-1}, S_i)}{a(S_i, t_i)} \left[1 - A_1 \Delta t\right], \\
p(t_{i-1}, S_{i-1}; t_i, S_i) &\leq \frac{e^{-\frac{d^2(S_{i-1}, S_i, t_{i1})}{2\Delta t}}}{\sqrt{2\pi\Delta t}} \frac{H_0(t_{i-1}, S_{i-1}, S_i)}{a(S_i, t_i)} \left[1 + A_1 \Delta t\right].
\end{aligned}$$

The proof is immediate by replacing each of the transition densities in (2.3) with their upper and lower bounds and integrating over Ω[1]. □

[1]More precisely, we integrate over a compact subset of Ω where the heat kernel expansion and the corresponding bounds hold, such that the contribution to the integral from the complement of this subset is arbitrarily small. This principle of not feeling the boundary is explained in detail in for example Appendix A of [7].

Remark 2.2. *The difference between* (2.8) *and* (2.9)*, which correspond to approximating the transition density* p *by the heat kernel expansion up to zeroth order or first order respectively, is that* (2.9) *shows that the approximation* $\widetilde{C}_1$ *tends to the true option price* C *as* $n \to \infty$ *in contrast to* (2.8)*.*

Precisely, from (2.8)*, as* $n \to \infty$, $\lim_{n\to\infty} [1 \pm A_1 \Delta t]^n = e^{\pm A_1 T}$ *so*

$$e^{-A_1 T} \lim_{n\to\infty} \widetilde{C}_0 \le C \le e^{A_1 T} \lim_{n\to\infty} \widetilde{C}_0$$

whereas (2.9) *gives*

$$C = \lim_{n\to\infty} \widetilde{C}_1$$

since $\lim_{n\to\infty} [1 \pm A_2 (\Delta t)^2]^n = 1$*. For the most-likely-path approximation presented in this paper, we only need* $\widetilde{C}_0$*. Once again, there is no guarantee of asymptotic convergence of our approximation to the true implied volatility for large* T*.*

To compute $\widetilde{C}_0$, we need to perform the n-dimensional integration in (2.7). We approximate this n-dimensional integral as $\Delta t \to 0^+$ using the following Laplace asymptotic formula:

Lemma 2.2. *(Laplace asymptotic formula)*
Let Ω be a closed set in $\mathbb{R}^n$ with nonempty and smooth boundary $\partial\Omega$. Suppose ϕ is a continuous function in Ω and attains its minimum uniquely on the boundary $\partial\Omega$ at $x^* \in \partial\Omega$ and, given any $\epsilon > 0$, there exists $\delta > 0$ such that $\phi(x) \ge \phi(x^*) + \delta$ for all $x \in \Omega \setminus B_\epsilon(x^*)$, where $B_\epsilon(x^*) = \{x : |x - x^*| < \epsilon\}$ is the open ball of radius ϵ centered at x^*. Assume that f is an integrable function in Ω, i.e., $\int_\Omega |f(x)|dx < \infty$ and that f vanishes identically in Ω^c and on the boundary $\partial\Omega$ but the inward normal directional derivative of f at x^* is nonzero. Then we have the asymptotic expansion as $\tau \to 0^+$

$$\int_\Omega e^{-\frac{\phi(x)}{\tau}} f(x) dx = (2\pi)^{\frac{n-1}{2}} \tau^{\frac{n+3}{2}} e^{-\frac{\phi(x^*)}{\tau}} \left[\frac{\nabla f(x^*) \cdot \nabla \phi(x^*)}{\sqrt{\det \partial_{\mathbf{t}}^2 \phi(x^*)} |\nabla \phi(x^*)|^3} + \mathcal{O}(\tau) \right], \quad (2.10)$$

where $\partial_{\mathbf{t}}^2 \phi(x^*)$ is the Hessian of ϕ in the tangential direction to Ω at x^*.

Proof. See for instance Section 8.3 in Bleistein and Handelsman [3]. $\square$

With the aid of the Laplace asymptotic formula (2.10) in Lemma 2.2 applied to the multidimensional integral in (2.7), $k = 0$, the following theorem is immediate.

Theorem 2.1. *For any given positive integer* n*, the European out of the money call option price* C *struck at* K *with expiry* T *in the local volatility model (2.1) has the following upper and lower bounds:*

$$(1 - A_1 \Delta t)^n \, \widehat{C}(\mathbf{S}^*, \mathbf{t}) \le C(s_0, K, T) \le (1 + A_1 \Delta t)^n \, \widehat{C}(\mathbf{S}^*, \mathbf{t}), \quad (2.11)$$

where

$$\widehat{C}(\mathbf{S}^*,\mathbf{t}) := \frac{T^{\frac{3}{2}}}{\sqrt{2\pi}} e^{-\frac{n}{T}D(\mathbf{S}^*,\mathbf{t})} \frac{U(\mathbf{S}^*,\mathbf{t})}{n^{\frac{3}{2}}},$$

$\mathbf{S}^* = (S_1^*, \cdots, S_n^*)$ *is a solution to the constrained optimization problem*

$$\min_{\mathbf{S}} D(\mathbf{S},\mathbf{t}) = \min_{\mathbf{S}} \frac{1}{2}\sum_{i=1}^{n} d^2(S_{i-1}, S_i, t_{i-1}) \tag{2.12}$$

subject to the constraints $S_n = K$ *and* $S_i \geq 0$, *for* $i = 1, \cdots, n-1$, *and*

$$U(\mathbf{S},\mathbf{t}) = \frac{1}{\sqrt{\det \partial_{\mathbf{t}}^2 D(\mathbf{S}^*)}} \frac{a^2(S_n, t_{n-1})}{d^2(S_{n-1}, K, t_{n-1})} \prod_{i=1}^{n} \frac{H_0(t_{i-1}, S_{i-1}, S_i)}{a(S_i, t_i)},$$

$\partial_{\mathbf{t}}^2 D(\mathbf{S}^*)$ *is the Hessian of* D *in the tangential direction of the set* $\Omega = \{\mathbf{S} = (S_1, \cdots, S_n) : S_n \geq K\}$ *at* $S^* \in \partial\Omega$.

In the limit $n \to \infty$, or equivalently as $\Delta t = T/n \to 0$,

$$d(S_{i-1}, S_i, t_{i-1}) \sim \frac{\Delta S_i}{a(S_{i-1}, t_{i-1})}$$

where $\Delta S_i = S_i - S_{i-1}$, so in this limit

$$\frac{1}{2\Delta t}\sum_{i=1}^{n} d^2(S_{i-1}, S_i, t_{i-1}) \sim \frac{1}{2}\sum_{i=1}^{n}\left[\frac{\frac{\Delta S_i}{\Delta t}}{a(S_{i-1}, t_{i-1})}\right]^2 \Delta t.$$

Then, in the limit $\Delta t \to 0$, the minimization problem (2.12) gives rise to the following variational problem:

$$\min_{S(t)} \frac{1}{2}\int_0^T \left[\frac{S'(t)}{a(S(t), t)}\right]^2 dt$$

subject to $S(0) = s_0$ and $S(T) = K$. The Euler-Lagrange equation associated with the variational problem is given by the second order ordinary differential equation:

$$-\left[\frac{S'}{a(S,t)}\right]' + \frac{a_t(S,t)}{a(S,t)}\frac{S'}{a(S,t)} = 0.$$

Therefore, the optimal path satisfies the boundary value problem:

$$\begin{cases} -\left[\frac{S'}{a(S,t)}\right]' + \frac{a_t(S,t)}{a^2(S,t)} S' = 0, \\ \\ S(0) = s_0, \quad S(T) = K. \end{cases} \tag{2.13}$$

Definition 2.1. *(Variational most likely path)*
The unique solution of the boundary value problem in (2.13) is called the *variational most likely path (variational MLP)*.

Corollary 2.1. *The European call option price C struck at K with expiry T in the local volatility model (2.1) has the following approximation: there exists a function $A(T)$ converging to 1 as $T \to 0$ such that*

$$C(s_0, K, T) = A(T) \frac{T^{\frac{3}{2}}}{\sqrt{2\pi}} e^{-\frac{1}{2} \int_0^T \left[\frac{S'(t)}{a(S(t),t)}\right]^2 dt} U(s_0, K, T)[1 + \mathcal{O}(T)], \tag{2.14}$$

where $S(t)$ is the variational most likely path, i.e., the unique solution to the boundary value problem (2.13), and U is given by

$$U(s_0, K, T) = \lim_{n \to \infty} \frac{U(\mathbf{S}^*, \mathbf{t})}{n^{\frac{3}{2}}}.$$

Proof. The result follows by replacing the exponent in (2.11) by its limit as $n \to \infty$. □

We now state our main result.

Theorem 2.2. *Assume the dynamic of the underlying asset S_t under risk neutral probability measure is governed by the local volatility model* (2.1) *and the local volatility function $\sigma(x,t)$ satisfies the conditions* (2.4). *Then the implied volatility σ_{BS} has the following expansion as $T \to 0^+$,*

$$\sigma_{\mathrm{BS}}(K, T) = \sigma_{\mathrm{BS},0} \left(1 + \mathcal{O}(T)\right),$$

with

$$\sigma_{\mathrm{BS},0} = \left(\frac{\sqrt{T}}{|k|} \sqrt{\int_0^T \left[\frac{S'(t)}{a(S(t),t)}\right]^2 dt} \right)^{-1}, \tag{2.15}$$

where $S(t)$ is the variational most-likely-path (MLP) given by (2.13).

Proof. Recall the following asymptotic expansion for Black-Scholes formula (see for instance Lemma 2.5 of [7]):

$$\begin{aligned} C_{\mathrm{BS}}(s_0, K, T, \sigma_{\mathrm{BS}}) &= \sqrt{\frac{s_0 K}{2\pi}} e^{-\frac{(\log K - \log s_0)^2}{2\sigma_{\mathrm{BS}}^2 T}} \frac{\sigma_{\mathrm{BS}}^3 T^{\frac{3}{2}}}{(\log K - \log s_0)^2} [1 + \mathcal{O}(T)] \\ &= \sqrt{\frac{s_0 K}{2\pi}} e^{-\frac{k^2}{2\sigma_{\mathrm{BS}}^2 T}} \frac{\sigma_{\mathrm{BS}}^3 T^{\frac{3}{2}}}{k^2} [1 + \mathcal{O}(T)] \end{aligned} \tag{2.16}$$

and from Corollary 2.1, we have

$$C(s_0, K, T) = A(T) \frac{T^{\frac{3}{2}}}{\sqrt{2\pi}} e^{-\frac{1}{2} \int_0^T \left[\frac{S'(t)}{a(S(t),t)}\right]^2 dt} U(s_0, K, T)[1 + \mathcal{O}(T)], \tag{2.17}$$

Matching the exponential terms in (2.16) and (2.17) we obtain

$$\frac{1}{\sigma_{\mathrm{BS},0}} = \frac{\sqrt{T}}{|k|} \sqrt{\int_0^T \left[\frac{S'(t)}{a(S(t),t)}\right]^2 dt},$$

proving the result. □

3 Solution of the variational problem

Recall that the variational MLP is the unique solution of equation (2.13):

$$-\left[\frac{S'}{a(S,t)}\right]' + \frac{a_t(S,t)}{a^2(S,t)} S' = 0,$$

with boundary conditions

$$S(0) = s_0; \quad S(T) = K.$$

We begin by changing variables in (2.13) to $x(t) := \log S(t)/s_0$, $a(S(t),t) = S(t)\,\sigma(x(t),t)$ and rearranging to obtain

$$\frac{d}{dt}\left(\log\left(\frac{x'(t)}{\sigma(x(t),t)}\right)\right) = \partial_t \log \sigma(x,t)|_{x=x(t)} =: f(x(t),t) \tag{3.1}$$

with boundary conditions $x(0) = 0$, $x(T) = k$ and where once again the log-strike $k := \log K/s_0$.

Remark 3.1. *Note that for any given closed-form parameterization of the local volatility surface, $f(\cdot)$ is known in closed-form. $f(\cdot)$ is a measure of the time-inhomogeneity of the local volatility surface. In particular, if the local volatility function is time-independent, $f(\cdot) = 0$.*

Integrating (3.1), we obtain in the general case

$$\frac{x'(t)}{\sigma(x(t),t)} = A \, \exp\left\{\int_0^t f(x(s),s)\,ds\right\}. \tag{3.2}$$

The constant A may be determined by the condition

$$x(T) = k = \int_0^T x'(t)\,dt = A \int_0^T dt\,\sigma(x(t),t)\,\exp\left\{\int_0^t f(x(s),s)\,ds\right\},$$

so (3.2) may in turn be rearranged to give

$$x'(t) = k\,\frac{\sigma(x(t),t)\,\exp\left\{\int_0^t f(x(s),s)\,ds\right\}}{\int_0^T dt\,\sigma(x(t),t)\,\exp\left\{\int_0^t f(x(s),s)\,ds\right\}}. \tag{3.3}$$

Integrating (3.3) gives

$$x(t) = k\,\frac{\int_0^t du\,\sigma(x(u),u)\,\exp\left\{\int_0^u f(x(s),s)\,ds\right\}}{\int_0^T du\,\sigma(x(u),u)\,\exp\left\{\int_0^u f(x(s),s)\,ds\right\}}. \tag{3.4}$$

Remark 3.2. *Equation (3.4) leads to an efficient fixed-point algorithm for solving for the variational most likely path reminiscent of Reghai's algorithm in [9]. The natural choice of first guess is just the straight line*

$$x_0(t) = k\,\frac{t}{T}.$$

Remark 3.3. *Note that*

$$\frac{S'(t)}{a(S(t),t)} = \frac{x'(t)}{\sigma(x(t),t)} = A\,\exp\left\{\int_0^t f(x(s),s)\,ds\right\},$$

so the value of the variational problem may be expressed as

$$\frac{1}{2}\int_0^T \left[\frac{S'(t)}{a(S(t))}\right]^2 dt = \frac{A^2}{2}\int_0^T \exp\left\{2\int_0^t f(x(s),s)\,ds\right\}dt \tag{3.5}$$

with

$$A = \frac{k}{\int_0^T dt\,\sigma(x(t),t)\,\exp\left\{\int_0^t f(x(s),s)\,ds\right\}}.$$

By substituting (3.5) into (2.15) we obtain

Lemma 3.1. *In terms of the local volatility $\sigma(x,t)$, equation (2.15) may be rewritten as*

$$\sigma_{\mathrm{BS},0} = \frac{\frac{1}{T}\int_0^T \sigma(x(t),t)\,\exp\left\{\int_0^t f(x(s),s)\,ds\right\}dt}{\sqrt{\frac{1}{T}\int_0^T \exp\left\{2\int_0^t f(x(s),s)\,ds\right\}dt}} \tag{3.6}$$

where $x(t)$ is (the logarithmic version of) the variational most likely path.

3.1 Time-separable local volatility

Suppose the local volatility function is time-separable so that $\sigma(x,t) = \sigma(x)\,\theta(t)$ for some functions $\sigma(\cdot)$ and $\theta(\cdot)$. Then by the definition of $f(\cdot)$,

$$f(x,t) = \frac{\theta'(t)}{\theta(t)}.$$

In particular

$$\exp\left\{\int_0^t f(x(s),s)\,ds\right\} = \frac{\theta(t)}{\theta(0)}$$

and equation (3.2) becomes

$$\frac{x'(t)}{\sigma(x(t))\,\theta(t)} = A\,\frac{\theta(t)}{\theta(0)}.$$

With $\phi(x) = \int_0^x \frac{d\xi}{\sigma(\xi)}$, the solution is

$$x(t) = \phi^{-1}\left(\phi(k)\,\frac{\int_0^t \theta^2(s)\,ds}{\int_0^T \theta^2(s)\,ds}\right). \tag{3.7}$$

Thus, in the time-separable case, with this choice of $x(t)$, our lowest order approximation for the implied volatility reads

$$\sigma_{\mathrm{BS},0} = \frac{\frac{1}{T}\int_0^T \sigma(x(t))\,\theta(t)\,dt}{\sqrt{\frac{1}{T}\int_0^T \theta^2(t)\,dt}}. \tag{3.8}$$

In particular, if $\sigma(x,t) = \sigma(x)$ is time-homogeneous, $\theta(\cdot)$ is constant, and

$$\sigma_{\mathrm{BS},0} = \frac{1}{T}\int_0^T \sigma(x(t))\,dt. \tag{3.9}$$

At first sight, (3.9) seems to differ from the familiar BBF formula (1.2) of Berestycki, Busca and Florent derived in [2] and again in [7]. A straightforward computation confirms that the BBF formula and (3.9) are in fact equivalent.

Of course, the time-separable case can be reduced to the time-homogeneous case by the simple deterministic time-change

$$\tau(t) = \int_0^t \theta^2(s)\,ds.$$

Note that the variational MLP (3.7) in the time-separable case is just the time-changed version of the variational MLP in the time-homogeneous case. It follows that our approximation has the correct behavior under a deterministic time-change. Consequently, the numerical accuracy of our approximation is just as great in the time-separable case as it is in the time-homogeneous case. Specifically, were we to repeat the numerical experiments of [7] (where the time-inhomogeneous local volatility function chosen was time-separable), we would find that the accuracy of our new approximation was equal to that of the BBF formula in the time-homogeneous case. Moreover, as is well-known, the BBF formula is typically highly accurate in the time-homogeneous case.

4 Numerical experiment

In [7], the authors performed numerical experiments that demonstrated the high accuracy of the BBF formula when the local volatility function is time-homogeneous. In the previous section, we established that our new approximation is exactly time-changed BBF in the case of time-separable local volatility. Thus, the accuracy of the

BBF formula in the time-homogeneous case is inherited by our new formula in the time-separable case. This motivates us to test the variational MLP approximation with a local volatility function that is both much more realistic and more difficult than those considered in [7]; difficult in the sense that this volatility function explodes at $t = 0$. Specifically, consider the following SVI[2] parameterization of the local volatility surface:

$$\sigma^2(k,t) = a + b\left\{\rho\left(\frac{k}{\sqrt{t}} - m\right) + \sqrt{\left(\frac{k}{\sqrt{t}} - m\right)^2 + \sigma^2\, t}\right\}$$

with

$$\begin{aligned} a &= 0.0012, \\ b &= 0.1634, \\ \sigma &= 0.1029, \\ \rho &= -0.5555, \\ m &= 0.0439. \end{aligned} \tag{4.1}$$

Thus the local variance $\sigma^2(k,t)$ on each time slice (or expiration) t is parameterized as SVI, effectively a hyperbola in the log-strike k. In practice, to avoid numerical problems, we freeze the local volatility function for times $t \leq 1/250$ years, roughly one trading day.

Remark 4.1. *Because the local volatility surface* (4.1) *is singular at* $t = 0$*, small-time expansions such as those of [7] cannot work.*

In Figure 1, we plot implied volatilities from numerical PDE computations against actual SPX bid and offered implied volatilities as of September 15, 2005, the day before triple-witching[3]. We confirm by inspection that the SVI parameterization (4.1) of the local volatility surface generates a realistic implied volatility surface.

In Figures 2 and 3, we plot the variational MLP approximation $\sigma_{\text{BS},0}$ from (3.6) against the smile generated by a numerical PDE computation from (4.1) and the smiles corresponding to the Reghai and BBFe most-likely-path approximations from (1.5) and (1.6) respectively. We note that our new approximation dominates the alternative most-likely-path approximations; for each expiration, the smile generated by the variational MLP approximation is closer to the (true) smile generated by numerical PDE computation than the alternatives. We note that BBFe does better than Reghai for shorter expirations and Reghai does better than BBFe for longer expirations (but again, variational MLP does better than either).

On closer inspection of the smiles, we may observe that for all three most-likely-path approximations, there is consistently a large approximation error at the cusp of

[2]For details of this parameterization see [6]

[3]This is the same volatility surface as the one plotted in Figure 3.2 on page 36 of [5].

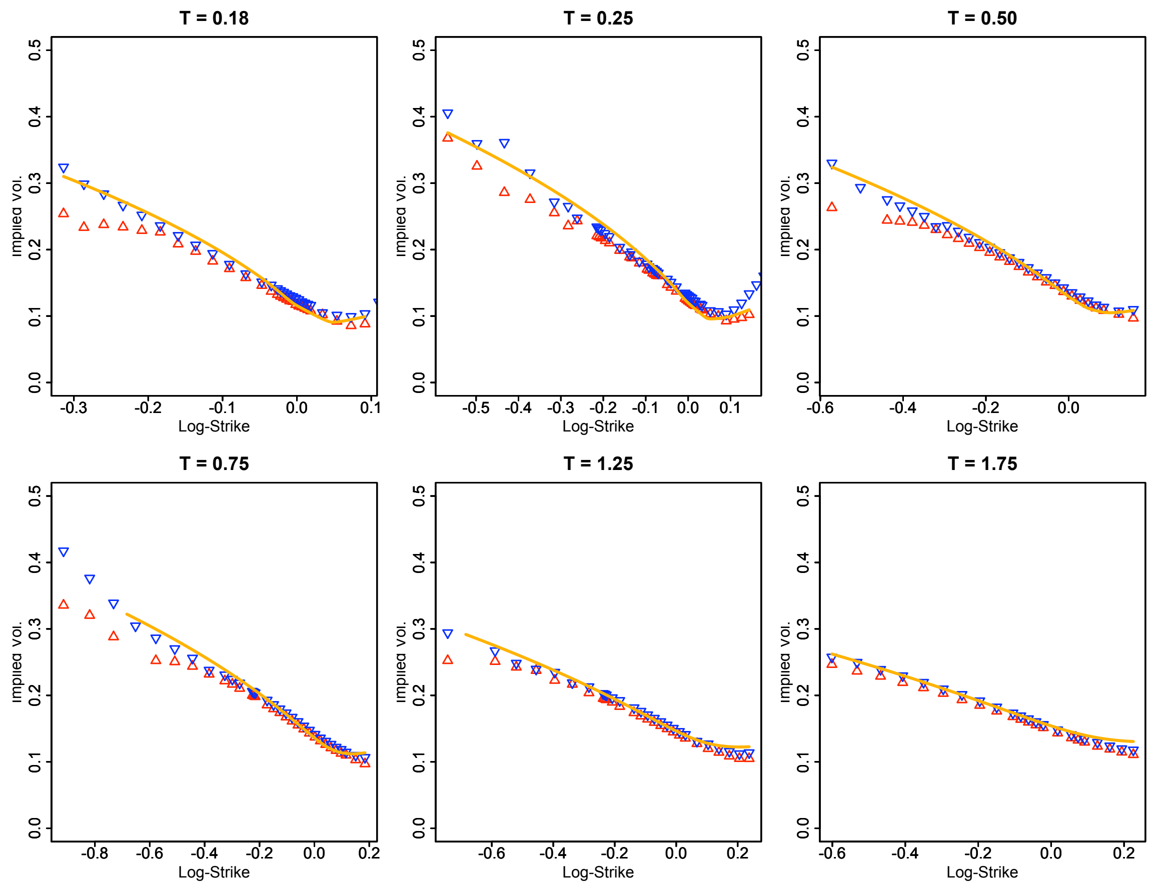

Figure 1: Solid lines are from PDE computations, triangles are market SPX bid and offered vols as of September 15, 2005. It is clear that our local volatility function generates a realistic implied volatility surface.

the smile. As noted in [8], this is because the most-likely-path technique fails when there is substantial curvature in the local volatility function. In that case, fixing a time t, one cannot reasonably approximate an integration over all possible prices of the underlying x_t by one point, the most-likely point $\tilde{x}(t)$. Appropriately approximating the integration over x_t may substantially improve accuracy. The price to be paid for this improved accuracy is in computational effort and indeed, a brute-force numerical PDE computation may often be faster (and of course even more accurate). In contrast, the fixed-point algorithm we proposed to compute our variational MLP estimate is very fast, typically converging within 3 or 4 iterations.

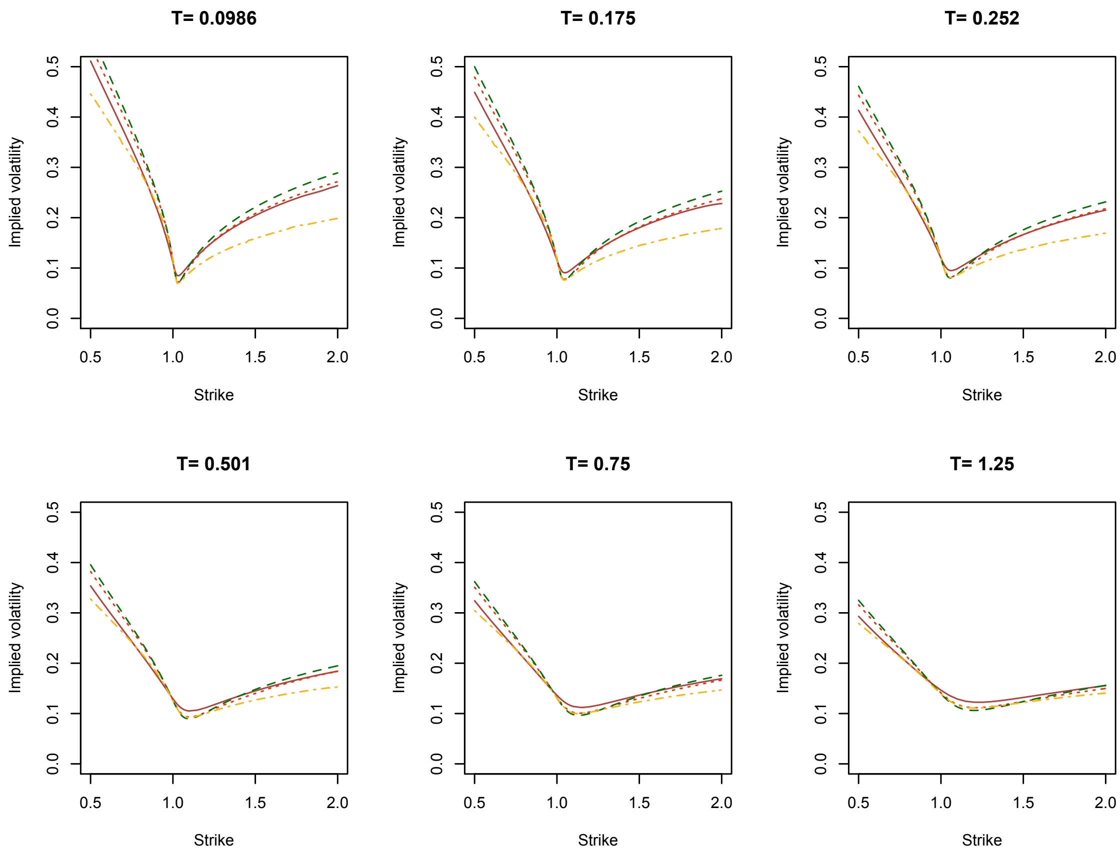

Figure 2: The solid lines are from PDE computations, the dotted lines are $\sigma_{\mathrm{BS},0}$, the dash-dotted lines are Reghai and the dashed lines are BBFe. Expirations shown are actual SPX market expirations as of September 15, 2005.

5 Summary and conclusion

We have derived a new most-likely-path estimate which we call *variational MLP* for approximating the implied volatility surface given a local volatility function. The variational MLP estimate is a natural extension of the BBF formula that moreover behaves correctly under a deterministic time-change, in contrast to prior definitions of most-likely-path. We have further shown by explicit computation using a realistic parameterization of the SPX local volatility function, that the variational MLP estimate outperforms two competing definitions of most-likely-path: a popular definition due to Adil Reghai and a naïve extension of the BBF formula.

How to improve the accuracy of our variational MLP estimate by for example better approximating the integration over x_t is left for future research.

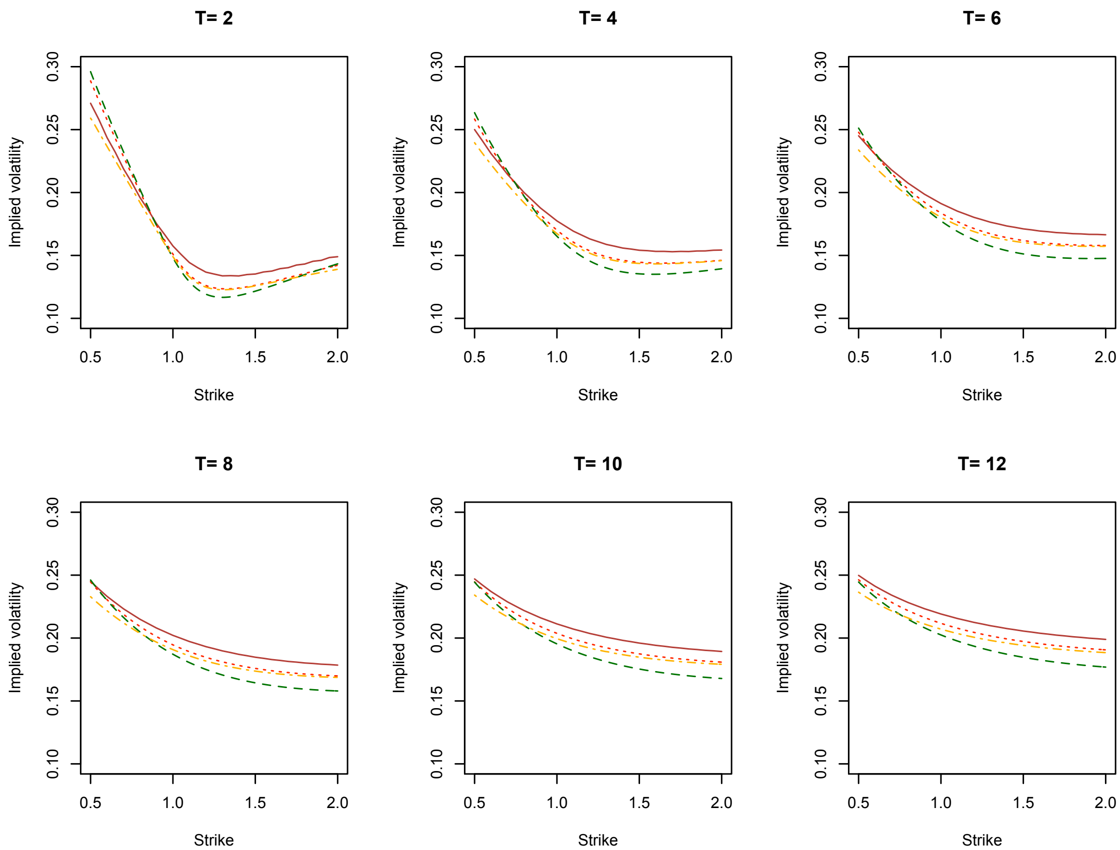

Figure 3: The solid lines are from PDE computations, the dotted lines are $\sigma_{\mathrm{BS},0}$, the dash-dotted lines are Reghai and the dashed lines are BBFe. Expirations are from 2 to 12 years.

Acknowledgments

We are very grateful to Julien Guyon and Vladimir Lucic for sharing their results and for valuable discussions. The first author also thanks Nick Constanzino and his collaborators at Penn State University for sending an early version of their paper. The second author is grateful to Professor Gérard Ben Arous of Courant Institute for insightful and valuable comments and discussions. All errors are our responsibility. Finally, we are particularly grateful to the referee whose valuable comments helped us correct some inaccuracies and improve our presentation.

References

[1] Achdou, R. and O. Pironneau, Computational methods for option pricing, SIAM

(2005).

[2] Berestycki, H., J. Busca, and I. Florent, Asymptotics and calibration of local volatility models, *Quantitative Finance* **2** (2002) 61–69.

[3] Bleistein, N. and R.A. Handelsman, Asymptotic expansions of integrals, Dover Publications (1986).

[4] Cheng, W., N. Costanzino, J. Liechty, A. Mazzucato and V. Nistor, Closed-form asymptotics and numerical approximations of 1D parabolic equations with applications to option pricing, *SIAM Journal on Financial Mathematics* (2011) forthcoming.

[5] Gatheral, J., The Volatility Surface: A Practitioner's Guide, Wiley Finance (2006).

[6] Gatheral, J., A parsimonious arbitrage-free implied volatility parameterization with application to the valuation of volatility derivatives, Presentation at *Global Derivatives & Risk Management, Madrid* (2004) available at `www.math.nyu.edu/fellows_fin_math/gatheral/madrid2004.pdf`.

[7] Gatheral, J., Hsu, E.P., Laurence, P., Ouyang, C., and Wang, T.-H., Asymptotics of implied volatility in local volatility models, *Mathematical Finance* (2011) forthcoming.

[8] Guyon, J. and P. Henry-Labordère, From local to implied volatilities, *SSRN preprint* (2010).

[9] Reghai, A., The hybrid most likely path, *Risk Magazine* April (2006).