

Computing the Implied Volatility in Stochastic Volatility Models

HENRI BERESTYCKI

École des Hautes Études en Sciences Sociales

JÉRÔME BUSCA

CNRS and Université Paris Dauphine

AND

IGOR FLORENT

HSBC CCF

1 Introduction and Main Results

The Black-Scholes model [6, 23] has gained wide recognition on financial markets. One of its shortcomings, however, is that it is inconsistent with most observed option prices. Although the model can still be used very efficiently, it has been proposed to relax its assumptions, and, for instance, to consider that the volatility of the underlying asset S is no longer a constant but rather a stochastic process. There are two well-known approaches to achieve this goal. In the first class of models, the volatility is assumed to depend on the variables t (time) and S , giving rise to the so-called *local volatility* models. The second one, conceptually more ambitious, considers that the volatility has a stochastic component of its own. In the latter, the number of factors is increased by the amount of stochastic factors entering the volatility modeling. Both models are of practical interest.

In these contexts, it is relevant to express the resulting prices in terms of *implied volatilities*. Given a price, the Black-Scholes *implied volatility* is determined, for each given product (that is for each given strike and expiry date defining, say, the call option) as the unique value of the volatility parameter for which the Black-Scholes pricing formula agrees with that given price. Actually, it is common practice on trading floors to quote and to observe prices in this way. A great advantage of having prices expressed in such dimensionless units is to provide easy comparison between products with different characteristics.

In principle, the implied volatility can be inferred from computed options prices by inverting the Black-Scholes formula. It is more convenient, however, to directly analyze the implied volatility. Indeed, this approach allows us to shed light on qualitative properties that would otherwise be more difficult to establish. In particular, we derive here several asymptotic formulae that are of practical interest, for example, in the calibration problem. The latter—an inverse problem that consists

in determining model parameters from the observation of market instruments—is typically computationally intensive.

This paper addresses precisely these questions for stochastic volatility models. In an earlier paper, we had carried out a similar program in the framework of local volatility models (see [4, 5]). There, the asymptotic behavior near expiry was given explicitly by an ordinary differential equation. The situation here is more involved. We introduce a new method to determine the implied volatility in the framework of stochastic volatility models for European call or put options. It is described in Section 1.2, where we show that the implied volatility is obtained by solving a quasi-linear parabolic partial differential equation, the initial condition of which is given by the solution of an eikonal (first-order) Hamilton-Jacobi equation. We establish a uniqueness result for this equation that is new and of independent interest.

The proof takes up Section 3 and involves the auxiliary notion of “effective volatility.” Its definition and some useful results about the effective volatility are given in Section 2. This notion is due to Derman and Kani [12]. Using a result of Varadhan and a large-deviation approach, the behavior of the effective volatility near expiry has been obtained by Avellaneda et al. [2, 3] in the context of basket options, and we apply the same methodology here. This is described in Section 2. We give original proofs of all these various results related to the effective volatility introducing a PDE type approach to this topic. These take up Sections 4 and 5. Furthermore, we establish here a new characterization of the Varadhan geodesic distance that is involved in the limiting theorem near expiry as a viscosity solution to a Hamilton-Jacobi equation.

Lastly, in Section 6 we show how our general approach applies to a variety of specific popular stochastic volatility models. There we derive closed-form approximate formulae. Finally, numerical computations illustrate the high-order accuracy that is achieved through the approximation with only one- or two-term expansions.

1.1 Stochastic Volatility Models

We assume that the volatility of the underlying asset S_t is given as a function of $n - 1$ stochastic factors $y_t = (y_t^1, \dots, y_t^{n-1})$ that follow diffusion processes. Specifically, we consider the following S.D.E.

$$(1.1) \quad \begin{cases} \frac{dS_t}{S_t} = r dt + \sigma(S_t, y_t, t) dW_t \\ dy_t = \theta(y_t, t) dt + v(y_t, t) dZ_t \end{cases}$$

where $W_t \equiv Z_t^0$, $Z_t = (Z_t^1, \dots, Z_t^{n-1})$, are standard Wiener processes. We define the correlation matrix $\Omega = (\rho_{ij})_{0 \leq i, j \leq n-1}$ by $\langle Z_t^i, Z_t^j \rangle = \rho_{ij} dt$. In (1.1), $\theta_t = (\theta_t^1, \dots, \theta_t^{n-1})$ are drift coefficients and $v(y_t, t) = (v_{ij})_{1 \leq i, j \leq n-1}$ is a diffusion matrix. Precise regularity and growth conditions on these coefficients will be stated below. We assume that $S_t \in (0, +\infty)$ and that y_t belongs to a domain $\mathcal{O} \subset \mathbb{R}^{n-1}$ a.s. In applications the domain \mathcal{O} will typically be the whole space or a

half-space. This framework includes such popular stochastic volatility models as the Heston model and the log-normal volatility model. We assume for simplicity in the sequel that $\mathcal{O} = \mathbb{R}^{n-1}$ (an example with \mathcal{O} a half-space is discussed later on in Section 6).

As is classical, we will assume that the fair value of any option is the expectation of its discounted payoff at maturity T under the probability for which (1.1) holds, i.e.,

$$(1.2) \quad C(S, y, t; K, T) = \mathbb{E}(e^{-r(T-t)}(S_T - K)_+ | \mathcal{F}_t),$$

where \mathcal{F}_t is the natural filtration. A discussion of this property as well as the choice of the probability measure are outside the scope of this paper. Equivalently, from the Feynman-Kac relation, C can be obtained as the solution of the linear parabolic partial differential equation in n space dimensions

$$(1.3) \quad \begin{cases} C_t + LC = 0 \\ C(S, y_t, t = T; K, T) = (S - K)_+ \end{cases}$$

in the domain $\{S > 0, y \in \mathbb{R}^{n-1}\}$, where the pricing operator L is defined by

$$(1.4) \quad \begin{aligned} L\varphi = & \frac{1}{2}\sigma^2(S, y, t)S^2\frac{\partial^2\varphi}{\partial S^2} \\ & + \sigma(S, y, t)S \sum_{1 \leq j, k \leq n-1} \rho_{0j}v_{k,j}(y, t)\frac{\partial^2\varphi}{\partial S\partial y_k} \\ & + \frac{1}{2} \sum_{1 \leq i, j, k, l \leq n-1} \rho_{ij}v_{ik}(y, t)v_{jl}(y, t)\frac{\partial^2\varphi}{\partial y_k\partial y_l} \\ & + rS\frac{\partial\varphi}{\partial S} + \sum_{1 \leq i \leq n-1} \theta_i\frac{\partial\varphi}{\partial y_i} - r\varphi. \end{aligned}$$

The *implied volatility* function $\Sigma(S, y, t; K, T)$ is uniquely defined by the relation

$$(1.5) \quad C(S, y, t; K, T) = C^{\text{BS}}(S, t; K, T; \Sigma(S, y, t; K, T)),$$

where $C^{\text{BS}}(S, t; K, T; \Sigma)$ is the price of call options in the Black-Scholes model [6] for a given volatility (constant) parameter $\Sigma > 0$. Recall that it is given by

$$(1.6) \quad C^{\text{BS}}(S, t; K, T; \Sigma) = SN(d_1) - Ke^{-r(T-t)}N(d_2)$$

with

$$d_1 = \frac{\ln(Se^{r(T-t)}/K)}{\Sigma\sqrt{T-t}} + \frac{1}{2}\Sigma\sqrt{T-t}, \quad d_2 = d_1 - \Sigma\sqrt{T-t}.$$

1.2 Main Results

Henceforth we use the following notation corresponding to indices being shifted by one digit. The correlation matrix is now $\Omega = (\omega_{ij})$ with $\omega_{ij} = \rho_{i-1,j-1}$, $1 \leq i, j \leq n$, the diffusion matrix is given by $M = (m_{ij})$ with $m_{11} = \sigma$, $m_{1k} = m_{k1} = 0$ if $k \neq 1$, $m_{ij} = v_{i-1,j-1}$ if $2 \leq i, j \leq n$. The drift terms are defined by $q_1 = rS - \frac{1}{2}\sigma^2(S, y, t)$, $q_i(x, \tau) = \theta_{i-1}(y, t)$, $i = 2, \dots, n$. For later purposes we also introduce a modified drift coefficient $\tilde{\theta}_1 = 0$ and $\tilde{\theta}_i = q_i + \omega_{1i}\sigma v_{i-1}$ for $i = 2, \dots, n$.

Let us introduce the reduced variables

$$\tau = T - t, \quad x_1 = \ln\left(\frac{S}{K}\right) + r(T - t), \quad x_i = y_{i-1} \text{ for } i \geq 2.$$

From now on, the ‘‘space’’ variables are $x = (x_1, \dots, x_n)$. With a slight abuse, we keep the notation $\sigma(x, \tau) = \sigma(S, y, t)$ and likewise for other functions. Last, we introduce a normalized price

$$u(x, \tau) \equiv u(x, \tau; K, T) = e^{r\tau} \frac{C(S, y, t; K, T)}{K}.$$

Throughout the paper, we shall make the following technical assumptions on the diffusion coefficients:

$$(H) \quad \begin{cases} q \in C^{\alpha,\alpha/2} \\ M\Omega M^T \in C^{\alpha,\alpha/2} \\ C(1 + |x|)^{-2}|\xi|^2 \leq \langle M\Omega M^T(x, \tau)\xi, \xi \rangle \leq C(1 + |x|)^2|\xi|^2 \end{cases}$$

for all $x, \xi \in \mathbb{R}^n$, $\tau \in (0, T)$. Here $C^{\alpha,\alpha/2}$ denotes the space of functions having uniformly bounded partial Hölder differential quotients with exponent α (respectively, $\alpha/2$) in the space (respectively, time) variables; see [15].

A straightforward computation shows that the pricing PDE in the reduced variables is

$$(1.7) \quad \begin{cases} u_\tau = \frac{1}{2} \text{Tr}(M\Omega M^T D^2 u) + q(x, \tau) \cdot Du \\ u(x, 0) = (e^{x_1} - 1)_+, \end{cases}$$

the equation being satisfied in $\mathbb{R}^n \times (0, T)$. Note that the matrix $M = M(x, \tau)$ depends on the variables x and τ .

Throughout the paper, D denotes the gradient vector $(\partial/\partial x_1, \dots, \partial/\partial x_n)$, D^2 the Hessian matrix, M^T the transpose of the matrix M , and Tr the trace.

The following classical result asserts the solvability of the pricing equation.

PROPOSITION 1.1 *Under assumption (H) there is a unique classical solution $u \in C_{\text{loc}}^{2,\alpha,\alpha/2}(\mathbb{R}^n \times (0, T)) \cap C(\mathbb{R}^n \times [0, T])$ that satisfies the arbitrage bounds*

$$(1.8) \quad (e^{x_1} - 1)_+ < u(x, \tau) < e^{x_1}$$

for all $x \in \mathbb{R}^n$, $\tau \in (0, T)$.

Our main result here is to derive a degenerate quasi-linear parabolic equation satisfied by the implied volatility, together with its asymptotic behavior as $\tau \rightarrow 0$.

THEOREM 1.2 *Under assumption (H) the implied volatility function is the unique solution $\Sigma(x, \tau) \in W_{loc}^{2,1,\infty}(\mathbb{R}^n \times (0, T))$ of the following well-posed nonlinear degenerate parabolic initial value problem*

$$(1.9) \quad \begin{cases} (\tau \Sigma^2)_\tau = H(x, \tau, \Sigma, D\Sigma) + \tau \mathcal{L}(x, \tau, \Sigma, D^2\Sigma) \\ \Sigma(x, 0) = \Sigma^0(x). \end{cases}$$

Here the operators in the equation are given by

$$(1.10) \quad H(x, \tau, \Sigma, D\Sigma) = \text{Tr} \left[M \Omega M^T \Sigma^2 \left(D \left(\frac{x_1}{\Sigma} \right) \otimes D \left(\frac{x_1}{\Sigma} \right) - \frac{1}{4} \tau^2 \Sigma^2 D\Sigma \otimes D\Sigma \right) \right] + \tau \tilde{\theta} \cdot \Sigma D\Sigma$$

and

$$(1.11) \quad \mathcal{L}(x, \tau, \Sigma, D^2\Sigma) = \Sigma \text{Tr}(M \Omega M^T D^2\Sigma)$$

in \mathbb{R}^n . The initial condition is determined by $\Sigma^0 = x_1/\psi^0(x, y)$ where ψ^0 is the unique solution of the eikonal equation

$$(1.12) \quad \begin{cases} H^0(x, D\psi^0) \equiv \text{Tr}[M \Omega M^T(x, 0) D\psi^0 \otimes D\psi^0] = 1 \\ \psi^0 = 0 \quad \text{on } \{x_1 = 0\} \\ \psi^0(x) > 0 \quad \text{for } x_1 > 0. \end{cases}$$

Furthermore, this function can be represented as

$$(1.13) \quad \psi^0(x) = \inf_{\{\xi_1=0\}} G(x, \xi),$$

where $G \in W_{loc}^{1,\infty}(\mathbb{R}^N \times \mathbb{R}^N)$ is the unique viscosity solution of the Hamilton-Jacobi equation

$$(1.14) \quad \begin{cases} H^0(x, D_x G) = 1 \quad \text{in } \mathbb{R}^N \setminus \{\xi\} \\ G(x = \xi, \xi) = 0 \\ \lim_{|x| \rightarrow \infty} G(x, \xi) = +\infty \end{cases}$$

for all $\xi \in \mathbb{R}^N$.

In formulae (1.10)–(1.12) and throughout the paper, we use the notation $A \otimes B$ to denote the matrix $(A_i B_j)_{1 \leq i, j \leq n}$ for any two vectors A and B .

Clearly, Theorem 1.2 provides a way to *directly* compute the implied volatility relating it to the parameters of the model (the coefficients of the SDE (1.1)). Moreover, it is a key to establishing some delicate qualitative properties. Obviously, Theorem 1.2 yields an approximation formula near expiry for the implied volatility. It is worth emphasizing the fact that to get this approximation, one has to solve a Hamilton-Jacobi equation. But furthermore, equation (1.9) can be used to systematically expand Σ in powers of τ , Σ^0 being the leading-order term. Indeed, the

higher-order terms are obtained by solving transport equations. This can be done explicitly in several cases; see Section 6 for an example. These computations are quite useful in practice. We will see in Section 6 that, in practical cases, they yield remarkably accurate approximations.

We wish to emphasize that the existence and uniqueness results for problem (1.12) in the theorem above are new and rather delicate. Indeed, the problem is set in the whole space with an “interior Dirichlet condition” on a hyperplane. (Thus, in practice, the problem is set in two half-spaces with a boundary condition on the limiting hyperplane.) There is a vast literature devoted to the Dirichlet problem for Hamilton-Jacobi equations in bounded domains (see, in particular, [10, 19, 22]), but comparatively little is known for unbounded domains, apart from the work of O. Alvarez [1]. There an existence and uniqueness result for Hamilton-Jacobi equations of type (1.12) is established but under the assumption that the Hamiltonian $H^0(x, p)$ has linear growth in x as $|x| \rightarrow \infty$. Note that, here, in view of assumption (H), the Hamiltonian has quadratic growth in x . We therefore require an extension of the result of Alvarez with some new ingredients. Incidentally, this raises the question of the range of validity of such a uniqueness result in a more general setting. In other words, how much can the result of Alvarez be extended to allow for more general growth conditions?

2 Effective Volatility and Limiting Behavior near Expiry

In the proof of Theorem 1.2, we require the auxiliary notion of *effective volatility*, which we now describe.

It has been known for some time that from the point of view of valuing all European options at $t = t_0$, it is equivalent to solving the n -dimensional PDE (1.3) for all values of K , $T > 0$ or to solve a *one-dimensional* PDE with a well-chosen diffusion coefficient. The latter is called *effective local volatility*, following the terminology in [12]. This result is due to Derman and Kani [12]. The precise statement is the following:¹

THEOREM 2.1 [12] *Let $C = C(S_0, y_0, t_0; K, T)$ be the solution of (1.3), i.e., the family of all call option prices at $t = t_0$ for an initial spot S_0 and initial value of the volatility factors $y_0 = (y_0^1, \dots, y_0^{n-1})$. The function C is also the unique solution of the parabolic problem in the variables (K, T)*

$$(2.1) \quad \begin{cases} C_T = \frac{1}{2}\bar{\sigma}^2(S_0, y_0, t_0; K, T)K^2C_{KK} - rKC_K \\ C(S_0, y_0, t_0; K, T = t_0) = (S_0 - K)_+, \end{cases}$$

where the effective local volatility $\bar{\sigma} \in C([0, +\infty)^2)$ is given by

$$(2.2) \quad \bar{\sigma}^2(S_0, y_0, t_0; K, T) = \frac{\int \pi(S_0, y_0, t_0; K, y, T)\sigma^2(K, y, T)dy}{\int \pi(S_0, y_0, t_0; K, y, T)dy},$$

¹The authors are thankful to M. Avellaneda for pointing out the importance of the result in Theorem 2.1.

and $\pi(S, y, t; S', y', t')$ is the Green function of the pricing problem, i.e., the solution of

$$(2.3) \quad \begin{cases} \pi_t + L\pi = 0 \\ \pi(S, y, t = t'; S', y', t') = \delta_{(S,y)=(S',y')}, \end{cases}$$

where L is defined in (1.4), defined for $S, S' > 0, y, y' \in \mathbb{R}^{n-1}, t < t'$.

Let us note that, with this definition of the Green function, the solution C of (1.3) is given by

$$(2.4) \quad C(S, y, t; K, T) = \int \pi(S, y, t; S', y', T)(S' - K)_+ dS' dy'.$$

This can be seen as a generalization of a well-known formula regarding local volatility of B. Dupire [14] for the following reason: Consider the case when σ is independent of y : $\sigma(S, y, t) \equiv \sigma(S, t)$ —the so-called *local* or *deterministic volatility model*; we find that (2.2) reduces to $\bar{\sigma}(S_0, y_0, t_0; K, T) \equiv \sigma(K, T)$, that is, Dupire’s result. It is also related to the papers of Derman and Kani [11], Dupire [14], and the formula of Breeden and Litzenberger [7]. For a review of those results, see [21]. The latter states that state price densities are recovered from options prices by differentiating twice with respect to the strike variable. It reads

$$(2.5) \quad \pi(S, y, t; S', y', t') = \frac{\partial^2 C}{\partial K^2} C(S, y, t; S', t').$$

That equation (2.1) defines a well-posed problem, however, is far from obvious. It relies on the fact that the effective volatility remains locally bounded as $T \rightarrow t_0$; this is very much linked to the assumption that the stochastic volatility factors follow a diffusion process. Indeed, this property may be violated for other kinds of stochastic processes. One of the difficulties here in proving that the effective volatility remains locally bounded as $T \rightarrow t_0$ is that the volatility of the diffusion process is not necessarily bounded.

For these reasons we give here, in Section 4, a new and complete proof of Theorem 2.1 relying on PDE methods.

Note that in view of expression (2.2), the equivalent volatility $\bar{\sigma}$ can be thought of as the conditional expectation

$$(2.6) \quad \bar{\sigma}^2(S_0, y_0, t_0; K, T) = \mathbb{E}^{S_0, y_0, t_0} (\sigma^2(K, y_T, T) \mid S_T = K).$$

This formula results from a formal application of Ito’s formula on the terminal payoff of the call, $(S_T - K)_+$ (see [12]).

In using the results of Theorem 2.1, an essential role is played by the limiting behavior close to expiry. This is the object of the following result, in which we identify the limit of the equivalent local volatility for short time to maturity:

THEOREM 2.2 *When $T - t_0 \rightarrow 0$ the effective volatility converges locally uniformly to $\bar{\sigma}(S, y, t_0 = T; K, T)$, which satisfies the limiting first-order problem*

$$(2.7) \quad \begin{cases} \text{Tr}(M\Omega M^T Dd \otimes D\bar{\sigma}) = 0 & \forall S > 0, \forall y \in \mathbb{R}^{n-1}, \\ \bar{\sigma}(S = K, y, T; K, T) = \sigma(K, y, T) & \forall y \in \mathbb{R}^{n-1}, \end{cases}$$

where $d = d(S, y; K)$ is Varadhan’s signed geodesic distance to the hyperplane $\{S = K, y \in \mathbb{R}^{n-1}\}$ associated to the elliptic operator L defined in [24].

This result is analogous to that in [2, 3], and we apply the same methods. However, it is stated here in the framework of stochastic volatility models. In Section 5 below, we give an original and rigorous proof of the result with PDE arguments.

We will make use of another characterization of the geodesic distance. In the next statement, we derive a way of actually computing the function d .

THEOREM 2.3 *The signed geodesic distance d defined above is the unique viscosity solution of*

$$(2.8) \quad \begin{cases} \text{Tr}(M\Omega M^T Dd \otimes Dd) = 1 & \forall S > 0, \forall y \in \mathbb{R}^{n-1}, \\ d(S = K, y; K) = 0 & \forall y \in \mathbb{R}^{n-1}. \end{cases}$$

Note that d is nothing else than ψ^0 defined in (1.12) in the variables (S, y) . This result yields a simple means of computing $\bar{\sigma}$ at $t_0 = T$ by applying the *method of rays*. Indeed, given $K > 0$, one can solve the system of ODEs $(\dot{S}(\tau), \dot{y}(\tau)) = M\Omega M^T \nabla d(S(\tau), y(\tau))$, $(S(0), y(0)) = (S_0, y_0)$ where ∇d is smooth, that is, at least in a neighborhood² of the hyperplane $\{S = K, y \in \mathbb{R}^{n-1}\}$. It is easily seen that

$$(2.9) \quad \bar{\sigma}(S_0, y_0, T; K, T) = \sigma(K, y(\tau^*), T),$$

where τ^* is the first time $S(\tau)$ hits the hyperplane $\{S = K, y \in \mathbb{R}^{n-1}\}$. An interpretation of (2.9) is that the averaging that takes place in (2.2) is replaced in the asymptotic regime $t_0 \rightarrow T$ by an evaluation of the stochastic volatility at a point $y = y(\tau^*)$ that is determined by the geodesic distance d . Intuitively, the higher the “volatility of the volatility factors” is, the farther $y(\tau^*)$ will deviate from y_0 , thus creating a more pronounced *smile*.

An important consequence of Theorem 2.1 is that the implied volatility remains bounded and bounded away from 0 as $T - t_0 \rightarrow 0$ —again an essential feature of diffusion models, which will be used in the proof of Theorem 1.2. We state this property in the following proposition:

PROPOSITION 2.4 *Given $\tau_0 > 0$, there exists a continuous function Λ , $\Lambda = \Lambda(R, \tau_0)$, such that*

$$(2.10) \quad 0 < \Lambda(R, \tau_0)^{-1} \leq \Sigma(x, \tau) \leq \Lambda(R, \tau_0)$$

for all $x \in B_R = \{x \in \mathbb{R}^n : |x| < R\}$, $\tau \in (0, \tau_0)$.

² Actually, this ODE is solvable in the classical sense up to the cut locus of (K, y_0) .

3 Determining the Implied Volatility

The proof of Theorem 1.2 will take up all this section. It is divided into several parts.

3.1 Derivation of the PDE for the Implied Volatility

We first show that the nonlinear PDE in (1.9) follows from the pricing equation (1.3) in a straightforward way.

Let us denote by U the solution to (1.7) corresponding to the Black-Scholes model with unit volatility, that is, $M\Omega M^T \equiv e_1 \otimes e_1$. Then U solves

$$(3.1) \quad \begin{cases} U_\tau = \frac{1}{2}(U_{x_1x_1} - U_{x_1}), & \tau > 0, x_1 \in \mathbb{R}, \\ U(x, 0) = (e^{x_1} - 1)_+. \end{cases}$$

The explicit solution to (3.1) is readily seen to be

$$(3.2) \quad U(x, \tau) \equiv U(x_1, \tau) = e^{x_1} N\left(\frac{x_1}{\sqrt{\tau}} + \frac{1}{2}\sqrt{\tau}\right) - N\left(\frac{x_1}{\sqrt{\tau}} - \frac{1}{2}\sqrt{\tau}\right),$$

where

$$(3.3) \quad N(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{y^2}{2}} dy.$$

Note that this is nothing else but the classical Black-Scholes formula [6] written in reduced variables. Next, owing to the obvious scaling properties of (3.1), we observe that the implied volatility (1.5) is equivalently defined by

$$(3.4) \quad u(x, \tau) = U(x_1, \Sigma(x_1, \tau)^2 \tau).$$

Next, we make use of the following relations, which hold for the solution of (3.1):

$$(3.5) \quad \frac{U_{x_1\tau}}{U_\tau}(x_1, s) = \frac{1}{2} - \frac{x_1}{s}, \quad \frac{U_{\tau\tau}}{U_\tau}(x_1, s) = -\frac{1}{2s} + \frac{x_1^2}{2s^2} - \frac{1}{8}.$$

With these, a direct computation allows us to derive the nonlinear PDE in (1.9) from the pricing equation (1.7).

3.2 Construction of the Limiting Solution

We are going to construct a solution of the eikonal equation (1.12). It is clearly equivalent to solving first $H(x, 0, \Sigma, D\Sigma) = \Sigma^2$ and then setting $\psi = x_1/\Sigma$. For this, we first localize the problem in a ball B_R and solve the localized problem by the *vanishing viscosity* method [22]. Then we prove uniqueness for the limiting problem and let $R \rightarrow \infty$. A difficulty in the first step is the lack of coercivity of the Hamiltonian on account of the $D(x_1/\Sigma)$ term in (1.10). This makes it difficult to construct a solution by the usual Perron method. However, since

$$(3.6) \quad \Sigma\left(\frac{x_1}{\Sigma}\right)_i = \delta_{1,i} - x_1(\ln \Sigma)_i,$$

we are naturally led to the natural change of variable $w = \ln \Sigma$.

For a fixed $R > 0$ we now show that the localized problem in the new variables can actually be solved.

LEMMA 3.1 *Let $\Lambda > 0, \epsilon > 0$. There exists a unique classical solution $\Sigma^\epsilon \in C_{loc}^{2,\alpha}(B_R) \cap C(\overline{B_R})$ of*

$$(3.7) \quad \begin{cases} H(x, 0, \Sigma^\epsilon, D\Sigma^\epsilon) + \epsilon \Delta(\ln \Sigma^\epsilon) - (\Sigma^\epsilon)^2 = 0 \\ \Sigma^\epsilon|_{\partial B_R} = \Lambda \end{cases}$$

(where H is defined in (1.10)) that satisfies the bound

$$(3.8) \quad 0 < \min \left\{ \Lambda, \inf_{x \in \overline{B_R}} \sigma(x, 0) \right\} \leq \Sigma^\epsilon(x, 0) \leq \max \left\{ \Lambda, \sup_{x \in \overline{B_R}} \sigma(x, 0) \right\}$$

uniformly in $\epsilon > 0$.

When $\epsilon \rightarrow 0$, for Λ and R fixed, the function Σ^ϵ converges locally uniformly in B_R to the unique viscosity solution Σ^0 of

$$(3.9) \quad H(x, 0, \Sigma^0, D\Sigma^0) = (\Sigma^0)^2$$

in \mathbb{R}^n . The function $\psi^0 = x_1/\Sigma^0$ is the solution of (1.12).

PROOF OF LEMMA 3.1: Setting $w^\epsilon = -\ln \Sigma^\epsilon$, equation (3.7) can be rewritten as

$$(3.10) \quad \text{Tr} \left(M\Omega M^T(x, 0) \left((\delta_{1,i} + x_1 w_i^\epsilon) (\delta_{1,j} + x_1 w_j^\epsilon) \right)_{i,j} \right) - \epsilon \Delta w^\epsilon - e^{-2w^\epsilon} = 0.$$

This equation is clearly uniformly elliptic and coercive and hence enjoys a comparison principle. Now, noting that $\text{Tr}(M\Omega M^T(x, 0)e_1 \otimes e_1) = \sigma^2(x, 0)$, we see that the constants

$$-\ln \left(\min \left(\Lambda, \inf_{x \in \overline{B_R}} \sigma(x, 0) \right) \right) \quad \text{and} \quad -\ln \left(\max \left(\Lambda, \sup_{x \in \overline{B_R}} \sigma(x, 0) \right) \right)$$

are respectively sub- and supersolutions of (3.10). Perron’s method [9] thus applies in a standard way to give the existence of a viscosity solution satisfying the bounds (3.8), uniqueness being a trivial consequence of the comparison principle. The $C^{2,\alpha}$ regularity of solutions follows from (H) and from the general theory of Hölder regularity for elliptic equations [16]. □

The uniqueness property for the limiting equation (3.9) is fairly unusual given that there is no apparent boundary condition. As a matter of fact, (3.9) is degenerate on the hyperplane $\{x_1 = 0\}$, which turns out to imply a boundary condition there that is automatically satisfied by any solution.

PROPOSITION 3.2 *There exists a unique viscosity solution of (2.8); it is given by the representation formula (1.13).*

PROOF: The proof has three steps.

Step 1: Construction of the fundamental solution. Let us fix $\xi \in \mathbb{R}^N$. By the results of Kruřkov [20], it is classical to construct (using the vanishing viscosity method) a viscosity solution $G(x, \xi)$ of $H^0(x, D_x G) = 1$ satisfying $G(\xi, \xi) = 0$ and $G > 0$ for $|x - \xi| > 0$. If one defines the radial envelopes of $M\Omega M^T(x, 0)$ by

$$\overline{A}(r) = \sup_{|x|=r} M\Omega M^T(x, 0), \quad \underline{A}(r) = \inf_{|x|=r} M\Omega M^T(x, 0),$$

together with the solutions of the radial problems

$$\langle \overline{A}(r)D\underline{G}, D\underline{G} \rangle \equiv \overline{A}(r)\underline{G}_r^2 = 1, \quad \langle \underline{A}(r)D\overline{G}, D\overline{G} \rangle = 1, \quad \underline{G}(0) = \overline{G}(0) = 0,$$

it is classical (by applying a comparison principle when constructing G) that $\underline{G} \leq G \leq \overline{G}$. In view of assumption (H), one gets the behavior at infinity

$$(3.11) \quad C_0^{-1}(1 + (\ln|x - \xi|)_+) \leq G(x, \xi) \leq C_0(1 + |x - \xi|^2), \quad i = 1, 2,$$

so that in particular

$$(3.12) \quad \lim_{|x| \rightarrow \infty} G(x, \xi) = \infty.$$

Step 2: Uniqueness of the fundamental solution. The solution we have just constructed is unique. The proof of this fact here follows the approach of Alvarez [1]. Recall that [1] is restricted to linear growth of the Hamiltonian, while the growth here is quadratic. Therefore some new estimates are needed.

Let us consider two solutions G_1 and G_2 in $W^{1,\infty}$ of (1.14). Since the first-order equation is satisfied a.e., each G_i satisfies (3.11) and (3.12). Let us now perturb G_1 into a strict supersolution by introducing $\tilde{G}_1 = \Phi^{-1}(G_1)$ with $\Phi(z) = z$ for all $z \in [0, M]$, $\Phi(z) = 1/(3C_0) \ln(ze^{6MC_0}/(2M))$ in $z \in (2M, +\infty)$, and Φ increasing and C^1 in $(0, +\infty)$. This is a supersolution of the eikonal equation in (1.14) since we clearly have

$$\text{Tr}(M\Omega M^T(x, 0)D\tilde{G}_1 \otimes D\tilde{G}_1) = \frac{1}{\Phi'^2(\tilde{G}_1)} \geq 1.$$

A key observation is now that $\tilde{G}_1 \geq G_2$ for $|x| \gg 1$. Indeed, in view of (3.12) one has $\tilde{G}_1(x, \xi) \geq C|x|^3 \geq G_2(x, \xi)$ for $|x| \geq R = R(M, C_0, \xi)$. We can then apply the classical comparison principle [9] in the bounded domain $B_R \setminus \{\xi\}$ to infer that $\tilde{G}_1 \geq G_2$ there. Sending now M to ∞ , one gets $G_1 \geq G_2$ everywhere.

Step 3: Representation formula. It is classical that, given any set S , the function $\psi = \inf_{\xi \in S} G(x, \xi)$ is a viscosity solution of $H^0(x, D_x) = 1$ that vanishes on S . Choosing $S = \partial\omega_R$, $\omega_R = \{x = (x_1, x') \mid x_1 > 0, |x'| \leq R\}$, and noting that $\psi \geq 0$ in $\{x_1 \geq 0\}$, we can apply the classical comparison principle in ω_R to infer that $\psi \geq \inf_{\xi \in \partial\omega_R} G(x, \xi)$ in ω_R for all $R > 0$. Since $G(x, \xi) \rightarrow \infty$ when $|x - \xi| \rightarrow \infty$, we have in the limit $R \rightarrow \infty$

$$(3.13) \quad \psi(x) \geq \inf_{\{\xi=(\xi_1, \xi') \mid \xi_1=0\}} G(x, \xi) \quad \text{in } \{x_1 > 0\}.$$

To get the reverse inequality, we use a similar argument to that in step 2 which we only briefly sketch. We define $\varphi_n(x) = \inf_{\{\xi=(\xi_1, \xi') \mid \xi_1=0, |\xi'| \leq n\}} G(x, \xi)$ and modify it at infinity into a strict supersolution with the help of Φ defined in step 2. Since ψ satisfies (1.12), by (H) we have $|D_x \psi| \leq C(1 + |x - \xi|^2)^{1/2}$, so that ψ can grow at most quadratically at ∞ . We infer from here that $\psi \leq \varphi_n$ for $|x| \geq R = R(n)$. Applying the classical comparison principle in $\omega_R \setminus \{\xi = (\xi_1, \xi') \mid \xi_1 = 0, |\xi'| \leq n\}$, one gets $\psi \leq \varphi_n$ there. Sending n to ∞ yields the reverse inequality in (3.13). \square

3.3 A Comparison Principle

LEMMA 3.3 *Suppose $\underline{\Sigma} \in C(\overline{B_R} \times (0, \tau_0))$ (respectively, $\overline{\Sigma} \in C(\overline{B_R} \times (0, \tau_0))$) satisfies*

$$(3.14) \quad (\tau \underline{\Sigma}^2)_\tau \leq H(x, \tau, D \underline{\Sigma}) + \tau \mathcal{L}(x, \tau, D^2 \underline{\Sigma})$$

(respectively, $\overline{\Sigma}, \geq$) in $B_R \times (0, \tau_0)$ in the viscosity sense, together with the lateral comparison

$$(3.15) \quad \underline{\Sigma} \leq \overline{\Sigma} \quad \text{on } \partial B_R \times (0, \tau_0)$$

and the growth condition at $\tau = 0$

$$(3.16) \quad \lim_{\tau \rightarrow 0} \tau \underline{\Sigma}^2(x, \tau) = \lim_{\tau \rightarrow 0} \tau \overline{\Sigma}^2(x, \tau) = 0 \quad \forall x \in B_R.$$

Then

$$(3.17) \quad \underline{\Sigma}(x, \tau) \leq \overline{\Sigma}(x, \tau) \quad \forall x \in B_R, \forall \tau \in (0, \tau_0).$$

PROOF: By a change of function this boils down to applying a classical maximum principle for solutions to linear parabolic PDEs. Setting $\underline{u}(x, \tau) = U(x, \underline{\Sigma}^2(x, \tau))$ and $\overline{u}(x, \tau) = U(x, \overline{\Sigma}^2(x, \tau))$, differentiating, and using relations (3.5) show that \underline{u} is a subsolution (respectively, \overline{u} is a supersolution) of (1.7), and $\underline{u} \leq \overline{u}$ on the lateral boundary as well as at initial time. It should be noted that (3.16) has been used to ensure that $\underline{u}(x, 0) = \overline{u}(x, 0) = u(x, 0) = (e^{x_1} - 1)_+$.

We now apply the comparison principle (Proposition 3.3) to infer that $\underline{u}(x, \tau) \leq u(x, \tau) \leq \overline{u}(x, \tau)$ everywhere, that is,

$$U(x, \underline{\Sigma}(x, \tau)^2 \tau) \leq U(x, \Sigma(x, \tau)^2 \tau) \leq U(x, \overline{\Sigma}(x, \tau)^2 \tau).$$

Since U is increasing as a function of its time argument, this in turn gives

$$(3.18) \quad \underline{\Sigma}(x, \tau) \leq \Sigma(x, \tau) \leq \overline{\Sigma}(x, \tau)$$

for all $x \in B_R, \tau \in (0, T)$. \square

3.4 Sub- and Supersolutions

We are going to define sub- and supersolutions of the time-dependent equation (1.9) from a relaxed version of problem (3.7). For this we introduce for $\delta, \varepsilon > 0$ the solution $\overline{\Sigma}^{\varepsilon, \delta}$ (respectively, $\underline{\Sigma}^{\varepsilon, \delta}$) of

$$(3.19) \quad \begin{cases} H(x, 0, \overline{\Sigma}^{\varepsilon, \delta}, D\overline{\Sigma}^{\varepsilon, \delta}) + \varepsilon \Delta(\ln \overline{\Sigma}^{\varepsilon, \delta}) - (\overline{\Sigma}^{\varepsilon, \delta})^2 = -\delta \\ \overline{\Sigma}^{\varepsilon, \delta}|_{\partial B_R} = \Lambda(R, 1) + 1 \equiv \Lambda(R) \end{cases}$$

(respectively, $\underline{\Sigma}^{\varepsilon, \delta}, \delta \rightarrow -\delta, \Lambda(R, 1) + 1 \rightarrow (\Lambda(R, 1) + 1)^{-1}$), where $\Lambda(R, 1)$ is defined in Proposition 2.4.

Ultimately we intend to treat (1.9) for small times as a perturbation of (3.7). The difficulty that arises in this task is the lack of a uniform estimate on the Hessian of the solutions of (3.19), which will eventually be needed to control the time-dependent terms in (1.9). A standard way to get around this is to regularize by sup and inf convolution [9], i.e., define for $\eta > 0$ the function

$$(3.20) \quad \overline{\Sigma}^{\varepsilon, \delta, \eta}(x) = \inf_{\xi \in B_R} \left(\overline{\Sigma}^{\varepsilon, \delta}(\xi) + \frac{1}{2\eta} |\xi - x|^2 \right),$$

and, respectively,

$$(3.21) \quad \underline{\Sigma}^{\varepsilon, \delta, \eta}(x) = \sup_{\xi \in B_R} \left(\underline{\Sigma}^{\varepsilon, \delta}(\xi) - \frac{1}{2\eta} |\xi - x|^2 \right).$$

For the reader’s convenience we list some of the well-known properties of sup and inf convolution that we shall use below. We refer to [9] for the definition of the semijets $J^{2, \pm}$ as well as the notion of pointwise twice differentiability.

LEMMA 3.4 *The following properties hold for all $\varepsilon, \delta, \eta > 0$:*

- (i) *The functions $\overline{\Sigma}^{\varepsilon, \delta, \eta}$ and $\underline{\Sigma}^{\varepsilon, \delta, \eta}$ are locally Lipschitz and a.e. twice differentiable.*
 - (ii) *$\overline{\Sigma}^{\varepsilon, \delta, \eta} \rightarrow \overline{\Sigma}^{\varepsilon, \delta}$ as $\eta \rightarrow 0$ (respectively, $\underline{\Sigma}^{\varepsilon, \delta, \eta}, \underline{\Sigma}^{\varepsilon, \delta}$) uniformly in B_R .*
 - (iii) *$\overline{\Sigma}^{\varepsilon, \delta, \eta}$ (respectively, $\underline{\Sigma}^{\varepsilon, \delta, \eta}$) is semiconcave with $D^2 \overline{\Sigma}^{\varepsilon, \delta, \eta} \leq (1/\eta)I$ (respectively, $D^2 \underline{\Sigma}^{\varepsilon, \delta, \eta} \geq -(1/\eta)I$) a.e.*
 - (iv) *If $(p, N) \in J^{2,+}(\underline{\Sigma}^{\varepsilon, \delta, \eta}(x_0))$ (respectively, $J^{2,-}, \overline{\Sigma}^{\varepsilon, \delta, \eta}$) for $x_0 \in B_R$, then $(p, N) \in J^{2,+}(\underline{\Sigma}^{\varepsilon, \delta}(x_0 + \eta p))$ (respectively, $J^{2,-}, \overline{\Sigma}^{\varepsilon, \delta, \eta}, x_0 - \eta p$) and*
- $$(3.22) \quad \underline{\Sigma}^{\varepsilon, \delta, \eta}(x_0) + \frac{1}{2}\eta |p|^2 = \underline{\Sigma}^{\varepsilon, \delta}(x_0 + \eta p),$$
- (respectively, $\overline{\Sigma}^{\varepsilon, \delta, \eta}, -\eta$).*
- (v) *$|D \overline{\Sigma}^{\varepsilon, \delta, \eta}|_{L^\infty(B_R)} \leq \min(|D \overline{\Sigma}^{\varepsilon, \delta}|_{L^\infty(B_R)}, \frac{1}{\eta}C(R))$ (respectively, $\underline{\Sigma}^{\varepsilon, \delta, \eta}$).*

For the proof of this lemma, see [9] for (i) through (iv) and [8] for (v).

We claim that these functions are super- and subsolutions of the time-dependent equation (1.9) for sufficiently small times. More precisely, we have the following result:

LEMMA 3.5 *For all $R > 0$ and $\delta > 0$, there exist $\eta_0 = \eta_0(R, \delta) > 0$ such that for all $0 < \eta < \eta_0(R, \delta)$, all $0 < \varepsilon < \varepsilon_0(R, \delta, \eta)$, and $0 < \tau < \tau_0(R, \delta, \eta)$, the function $\overline{\Sigma}^{\varepsilon, \delta, \eta}$ (respectively, $\underline{\Sigma}^{\varepsilon, \delta, \eta}$) is a supersolution (respectively, subsolution) of the equation in (1.9) in the domain $B_R \times (0, \tau_0)$.*

Let us first make clear how the last result (1.12) in Theorem 1.2 follows from this lemma. Applying the comparison principle (Lemma 3.3) we get that

$$(3.23) \quad \underline{\Sigma}^{\varepsilon, \delta, \eta}(x) \leq \Sigma(x, \tau) \leq \overline{\Sigma}^{\varepsilon, \delta, \eta}(x)$$

in $B_R \times (0, \tau_0)$. Now this implies

$$(3.24) \quad \underline{\Sigma}^{\varepsilon, \delta, \eta}(x) \leq \liminf_{\tau \rightarrow 0} \Sigma(x, \tau) \leq \limsup_{\tau \rightarrow 0} \Sigma(x, \tau) \leq \overline{\Sigma}^{\varepsilon, \delta, \eta}(x).$$

We now send first $\varepsilon \rightarrow 0$ and observe that by the arguments in the proof of Lemma 3.1, it is easy to see that $\overline{\Sigma}^{\varepsilon, \delta}$ (respectively, $\underline{\Sigma}^{\varepsilon, \delta}$) converges locally uniformly in B_R to $\overline{\Sigma}^\delta$ (respectively, $\underline{\Sigma}^\delta$) solution of

$$(3.25) \quad H(x, 0, \overline{\Sigma}^\delta, D\overline{\Sigma}^\delta) - (\overline{\Sigma}^\delta)^2 = -\delta$$

(respectively, $\underline{\Sigma}^\delta, \delta \rightarrow -\delta$). Now, using (3.20)–(3.21) we see that when $\varepsilon \rightarrow 0$ one has $\overline{\Sigma}^{\varepsilon, \delta, \eta} \rightarrow \overline{\Sigma}^{\delta, \eta}$ (respectively, $\underline{\Sigma}$), the inf (respectively, sup) convolution of the above-defined $\overline{\Sigma}^\delta$ ($\underline{\Sigma}^\delta$). Using the uniqueness result in Section 3.2 we infer that $\overline{\Sigma}^\delta$ and $\underline{\Sigma}^\delta$ converge as $\delta \rightarrow 0$ to Σ^0 defined in (1.12). We finally deduce that $\Sigma(x, \tau)$ converges as $\tau \rightarrow 0$ to the Σ^0 solution of the eikonal equation (1.12). This will eventually complete the proof of Theorem 1.2.

To conclude this section we now prove Lemma 3.5.

PROOF: We observe that for all $\varphi > 0$

$$|H(x, \tau, \varphi, D\varphi) - H(x, 0, \varphi, D\varphi)| \leq C(|\varphi|_{W^{1,\infty}} + |\ln \varphi|_{W^{1,\infty}})\tau$$

and, by using (iii) in Lemma 3.4, that

$$\tau \mathcal{L}(x, \tau, \overline{\Sigma}^{\varepsilon, \delta, \eta}, D^2 \overline{\Sigma}^{\varepsilon, \delta, \eta}) \leq \tau \frac{\Lambda(R)}{\eta}$$

(respectively, $\tau \mathcal{L}(x, \tau, \underline{\Sigma}^{\varepsilon, \delta, \eta}, D^2 \underline{\Sigma}^{\varepsilon, \delta, \eta}) \geq -\tau \Lambda(R)^{-1}/\eta$), together with

$$(3.26) \quad \varepsilon \Delta(\ln \overline{\Sigma}^{\varepsilon, \delta, \eta}) \leq \varepsilon C(R, |\overline{\Sigma}^{\varepsilon, \delta, \eta}|_{W^{1,\infty}}) \frac{1}{\eta}$$

(respectively, $\varepsilon \Delta(\ln \underline{\Sigma}^{\varepsilon, \delta, \eta}) \geq -\varepsilon C(R, |\underline{\Sigma}^{\varepsilon, \delta, \eta}|_{W^{1,\infty}})/\eta$). Finally, let us note that $|D\overline{\Sigma}^{\varepsilon, \delta}|_{L^\infty} \leq C(R)$ (respectively, $\underline{\Sigma}$) [22].

Using (3.26) and Lemma 3.4(v) leads to the estimates $|\underline{\Sigma}^{\varepsilon, \delta, \eta}|_{W^{1,\infty}} \leq C(R, \eta)/\eta$ and $|\overline{\Sigma}^{\varepsilon, \delta, \eta}|_{W^{1,\infty}} \leq C(R)/\eta$ uniformly in ε and δ . This implies that $\overline{\Sigma}^{\varepsilon, \delta, \eta}$ satisfies

$$\begin{aligned} (\tau(\overline{\Sigma}^{\varepsilon, \delta, \eta})^2)_\tau &\geq H(x, \tau, \overline{\Sigma}^{\varepsilon, \delta, \eta}, D\overline{\Sigma}^{\varepsilon, \delta, \eta}) + \tau \mathcal{L}(x, \tau, \overline{\Sigma}^{\varepsilon, \delta, \eta}, D^2 \overline{\Sigma}^{\varepsilon, \delta, \eta}) \\ &\quad + \delta - \varepsilon C(R, \eta) - \tau C(R, \eta) - C(R)\eta \end{aligned}$$

(respectively, $\underline{\Sigma}^{\varepsilon, \delta, \eta}$, \leq), these last terms accounting for the shift in the space variable due to the sup convolution; see Lemma 3.4(iv). For suitable ε , τ , and η , the remainder term is positive (respectively, negative); hence the statement in Lemma 3.5. \square

4 Effective Volatility: Existence and Uniqueness

PROOF: Let us introduce the generalized Dupire differential operator

$$(4.1) \quad \bar{L} = \frac{1}{2} \bar{\sigma}^2(S_0, y_0, t_0; K, T) K^2 \partial_K^2 - rK \partial_K.$$

In order to establish (2.1), we intend to show by applying the maximum principle that the function $w = (\partial_T - \bar{L})C$ vanishes identically for the choice of diffusion coefficient (2.2). For this we compute the Lie bracket (defined by $[A, B] = AB - BA$) of the pricing operator (1.4) with (4.1); we find

$$(4.2) \quad [\partial_T - \bar{L}, \partial_t + L](C) = \frac{1}{2} K^2 (\partial_t + L) \left(\frac{\partial^2 C}{\partial K^2} \bar{\sigma}^2 \right).$$

Differentiating the pricing equation (1.4) twice with respect to K and using the definition of the Green function π leads to

$$(4.3) \quad \frac{\partial^2 C}{\partial K^2}(S, y_0, t; K, T) = \int dy \int \pi(S, y_0, t; S', y, T) \delta_{S'=K} dS'$$

$$(4.4) \quad = \int \pi(S, y_0, t; K, y, T) dy.$$

Hence by (4.2) if one sets $\bar{\sigma}$ as

$$(4.5) \quad \left(\int \pi(S, y, t; K, y', T) dy' \right) \bar{\sigma}^2(S, y, t; K, T) = \int dS' \int \pi(S, y, t; S', y', T) \phi(S', y'; K, T) dy'$$

for any function distribution ϕ such that (4.5) is integrable, we see that

$$(4.6) \quad (\partial_t + L) \left(\frac{\partial^2 C}{\partial K^2} \bar{\sigma}^2 \right) = 0,$$

so that the function $w = (\partial_T - \bar{L})C$ satisfies the backward parabolic equation

$$(4.7) \quad (\partial_t + L)w = 0.$$

In a final step, we show that the choice $\phi \equiv \delta_{S'=K} \sigma^2(K, y', T)$ ensures the terminal condition $w_{t=T} = 0$. This allows us to apply the maximum principle to (4.7) to infer that $w \equiv 0$, which establishes (2.1) and (2.2).

We write the pricing equation (1.3) in the distribution sense at $t = T$ to find that

$$(4.8) \quad C_t + \frac{1}{2} \sigma^2(S, y, T) S^2 \delta_{S=K} + rS \mathbb{I}_{S>K} - r(S - K)_+ = 0.$$

Since $C(S, y, \xi; K, \xi) = (S - K)_+$ for all $\xi > 0$, we find that $C_t(S, y, t = T; K, T) = -C_T(S, y, t = T; K, T)$, so that we can rewrite (4.8) as

$$(4.9) \quad -C_T + \frac{1}{2}\sigma^2(K, y, T)K^2C_{KK} - rKC_K = 0$$

in the distribution sense at $t = T$. Now, (4.5) evaluated at $t = T$ yields (in the distribution sense)

$$(4.10) \quad \delta_{S=K}\bar{\sigma}^2(S, y, t = T; K, T) = \phi(S, y; K, T).$$

Hence $\phi(S, y; K, T) = \delta_{S=K}\sigma^2(K, y, T)$ is the only choice for which

$$(4.11) \quad w|_{t=T} \equiv (\partial_T - \bar{L})C|_{t=T} = 0$$

is satisfied. In view of (4.7) and (4.11), the maximum principle implies $w \equiv 0$. \square

5 Asymptotic Behavior of the Effective Volatility

The derivation of the asymptotic behavior rests on a formula of Varadhan that is essentially a large-deviation result. We first show that the function $\bar{\sigma}^2$ converges to a limit as $t_0 \rightarrow T$ and then prove that this limit satisfies the first-order Hamilton-Jacobi equation (2.7).

Varadhan’s formula [24], which we are going to use, states that

$$(5.1) \quad \lim_{t_0 \rightarrow T} (T - t_0) \ln(\pi(S_0, y_0, t_0; K, y, T)) = -d^2(S_0, y_0; K, y).$$

Here d is the geodesic distance defined by

$$(5.2) \quad d^2(S_0, y_0; K, y) = \inf \left\{ \int \langle M\Omega M^T(X(s), 0)\dot{X}(s), \dot{X}(s) \rangle ds \right\},$$

where the infimum is computed over all smooth paths $(X(s))_{s \in [0,1]}$ valued in $(0, +\infty) \times \mathbb{R}^{n-1}$ such that $X(0) = (S_0, y_0)$ and $X(1) = (K, y)$. We now note that

$$(5.3) \quad \left(\int \pi(S_0, y_0, t_0; K, y, T) dy \right)^{-1} \pi(S_0, y_0, t_0; K, y, T)$$

can be viewed as a family of positive functions of y_0 indexed by $\tau = T - t_0$ with unit L^1 mass. Applying standard stationary-phase arguments, one sees clearly from (5.1) that this family of kernels converges in $\mathcal{D}'(\mathbb{R}^{n-1})$ to a measure $\mu_{S_0, y_0, K}$ with unit mass supported in the set

$$(5.4) \quad \mathcal{S}_{S_0, y_0, K} = \{y \in \mathbb{R}^{n-1} : d(S_0, y_0; K, y) = d_\Gamma(S_0, y_0; K)\},$$

where d_Γ is the geodesic distance to the hyperplane $\{S = K\}$, i.e.,

$$(5.5) \quad d_\Gamma(S_0, y_0; K) = \inf_{y \in \mathbb{R}^{n-1}} d(S_0, y_0; K, y).$$

This argument shows that

$$(5.6) \quad \lim_{t_0 \rightarrow T} \bar{\sigma}^2(S_0, y_0, t_0; K, T) = \int \sigma^2(K, y, T) \mu_{S_0, y_0, K}(dy).$$

Henceforth this limit will be denoted by $\bar{\sigma}^2(S_0, y_0, T; K, T)$. Note that for S_0 sufficiently close to K , the cut locus of d_Γ is not reached, which implies that the set $\mathcal{S}_{S_0, y_0, K}$ is a singleton, and, in turn, that $\mu_{S_0, y_0, K}$ is a Dirac mass at that point.

PROOF: We now intend to establish that $\bar{\sigma}(S_0, y_0, T; K, T)$ satisfies the first-order Hamilton-Jacobi equation (2.7). For this we expand (4.6) to find

$$(5.7) \quad (\partial_t + L)(\bar{\sigma}^2) + \langle M\Omega M^T D(\ln \tilde{\pi}), D(\bar{\sigma}^2) \rangle = 0,$$

where $\tilde{\pi} = C_{KK} = \int \pi(S_0, y_0, t_0; K, y, T) dy$. We now multiply (5.7) by a test function $\phi \equiv \phi(S, y) \in C_0^\infty((0, +\infty) \times \mathbb{R}^{n-1})$, integrate it in $t \in (t_0, T)$, $S > 0$, $y \in \mathbb{R}^{n-1}$, and integrate by parts in the space variables the first two terms. The time derivative term gives rise to boundary values that cancel out when $t_0 \rightarrow T$ thanks to (5.6). Integrating by parts in the space variables, the term involving L disappears as well due to the local boundedness of $\bar{\sigma}$. Finally, the bracket term can be rewritten using the identity

$$(5.8) \quad \int_{t_0}^T \ln \tilde{\pi} dt = (T - t_0) \int_0^1 ds \ln \left(\int \pi(S_0, y_0, t_0 + s(T - t_0); K, y, T) dy \right).$$

Thanks to (5.1) this term clearly establishes (2.7) in the distribution sense; now by (2.2) $\bar{\sigma}$ is Lipschitz, so that (2.7) is actually satisfied a.e. in $(0, +\infty) \times \mathbb{R}^{n-1}$. \square

PROOF OF PROPOSITION 2.4: It follows from the above proof that the effective local volatility $\bar{\sigma}$ is locally bounded in (S_0, y_0, K, T) uniformly in $t_0 < T$. We shall now prove that the result in Proposition 2.4 follows from here. For this we observe that since the prices of call options satisfy (2.1), we have reduced the problem, for each (S_0, y_0, t_0) , to a pricing equation in *one space dimension*. Indeed, this was the main point of introducing the notion of effective volatility. This allows us to use the setting of one-dimensional local volatility models studied in [4, 5]. Setting $\tilde{\sigma}(x, \tau) = \bar{\sigma}(S_0, y_0, t; K, T)$ with $x = (x_1, y_0)$, $x_1 = \ln(S_0/K) + r(T - t_0)$, and $\tau = T - t_0$ as usual, we know that the implied volatility $\Sigma(x, \tau)$ is the solution of the *one-dimensional* nonlinear parabolic equation

$$(5.9) \quad \begin{cases} \frac{1}{\bar{\sigma}^2(x, \tau)} (\tau \Sigma^2)_\tau = \Sigma^2 \left(\frac{x_1}{\Sigma} \right)_{x_1}^2 + \tau \Sigma \Sigma_{x_1 x_1} - \frac{1}{4} \tau^2 \Sigma^2 \Sigma_{x_1}^2 \\ \Sigma(x, 0) = \left(\frac{1}{x} \int_0^x \frac{d\xi}{\bar{\sigma}(\xi, 0)} \right)^{-1} \end{cases}$$

that one may term an effective nonlinear PDE on the implied volatility. In [5] we established that (5.9) satisfies a comparison principle (i.e., all sub- and supersolutions must be ordered). Consequently, all we need do is construct a sub- and a supersolution of (5.9) in order to prove (2.10).

Let us recall (see [5]) that on account of the degenerate feature of (5.9), it is not necessary to check that the sub and supersolutions are ordered at $\tau = 0$; indeed, this is automatically satisfied since the initial condition simply reflects the limiting equation at $\tau = 0$.

We set, for $x_1 \in (-1, 1)$, $\bar{\Sigma}(x_1, x') = \Lambda(1 + x_1^2)$ with

$$\Lambda = \max\left(2 \sup_{\substack{|x| \leq 1 \\ 0 < \tau < 1}} \bar{\sigma}(x, \tau), 1\right),$$

and set $\bar{\Sigma}(x) = 2\Lambda|x_1|$ in $|x_1| > 1$. Note that this is a C^1 function. For $|x_1| > 1$ the right-hand side in (5.9) is negative while the left-hand side is positive. In the set $|x_1| \leq 1$, one computes

$$(5.10) \quad \left(\frac{x_1}{\bar{\Sigma}}\right)_{x_1} = \frac{1}{\Lambda} \left(\frac{1 - x_1^2}{(1 + x_1^2)^2}\right),$$

whereas $\bar{\Sigma}_{x_1 x_1} / \bar{\Sigma} \leq 2$. From here we clearly see that $\bar{\Sigma}$ is a supersolution for $0 < \tau < 1/(2\Lambda^2)$.

As for the subsolution we set, for $A > 1$, $\underline{\Sigma}^A(x) = \lambda(A)(1 - (x_1/A)^2)$ for $x_1 \in (-A, A)$ with $\lambda(A) = \min(1, \frac{1}{2} \inf_{\substack{|x| \leq A \\ 0 < \tau < 1}} \bar{\sigma}(x, \tau))$, and set $\underline{\Sigma}^A = 0$ in $\{|x_1| \geq A\}$. To correctly interpret (5.9) when $\underline{\Sigma}$ vanishes, one should keep in mind that (5.9) is but the rephrasing of the pricing equation in terms of the implied volatility; thus setting $\Sigma \equiv 0$ in an interval corresponds to considering a price function identically equal to the payoff function $(e^x - 1)_+$ there. A direct computation shows that the payoff function is a subsolution of the pricing equation [5], whatever $\tilde{\sigma}$ is. Hence $\underline{\Sigma}$ appear as the maximum of 0, which is a subsolution, and $\lambda(A)(1 - (x_1/A)^2)$. It thus remains to show that the latter function is a subsolution in $\{|x_1| < A\}$. For this we note that $(1 - x \underline{\Sigma}_{x_1}^A / \underline{\Sigma}^A)^2 \geq 1$, $|\underline{\Sigma}^A \underline{\Sigma}_{x_1 x_1}^A| \leq 2\lambda(A)^2/A^2 \leq 2$, $|\underline{\Sigma}^A \underline{\Sigma}_{x_1}^A| \leq 2\lambda(A)/A \leq 2$, so that the left-hand side in (5.9) is no greater than the right-hand side for $0 < \tau < \frac{1}{8}$. This shows that, for each $A > 0$, the function $\underline{\Sigma}^A$ is a subsolution of (5.9), and thus so is $\underline{\Sigma}(x) = \sup_{A > 1} \underline{\Sigma}^A(x)$. Note that $\underline{\Sigma}(x_1, x') \geq \frac{1}{4} \inf_{|\tilde{x}| < |x_1|} \tilde{\sigma}(\tilde{x}, 0)$ for all $x_1 \in \mathbb{R}$.

Applying the comparison principle easily yields the local boundedness of the implied volatility function (2.10). □

6 Examples and Further Remarks

In these models the space dimension is $n = 2$ (one stochastic volatility factor). For convenience we adopt the notation $x = x_1 = \ln(S_0/K) + r(T - t_0)$, $\tau = T - t_0$, and $y = x_2$.

6.1 A Log-Normal, Mean-Reverting Stochastic Volatility Model

In this model the dynamics of the underlying asset is given by

$$(6.1) \quad \begin{cases} dS_t = rS_t dt + y_t S_t \sigma(S_t) dW_t^1 \\ dy_t = a(\theta - y_t) dt + \kappa y_t dW_t^2 \end{cases}$$

with $\langle dW_t^1, dW_t^2 \rangle = \rho dt$, where a, θ , and κ are nonnegative constants. We simply rephrase our main result for this model. The signed geodesic distance d is the

unique solution of

$$(6.2) \quad \begin{cases} \sigma(x)^2 y^2 d_x^2 + 2\rho\kappa y^2 \sigma(x) d_x d_y + \kappa^2 y^2 d_y^2 = 1 \\ d(x=0, y) = 0 \\ d(x, y) > 0 \quad \text{for } x > 0 \end{cases}$$

with the usual notational abuse on σ . It turns out that d is given by the explicit expression

$$(6.3) \quad d(x, y) = \frac{1}{\kappa} \ln \left(\frac{\kappa z - \rho + \sqrt{1 - 2\rho\kappa z + \kappa^2 z^2}}{1 - \rho} \right),$$

where $z = \hat{x}/y, \hat{x} = \int_0^x d\zeta/\sigma(\zeta)$. From Theorem 1.2 one readily gets

$$(6.4) \quad \Sigma^0(x, y) \equiv \Sigma(x, y, \tau = 0) = \frac{x}{d(x, y)}.$$

As for the equivalent local volatility $\bar{\sigma}$, we know that it is the unique solution of (2.7). It can be computed explicitly by noting that, using (5.9), one has

$$(6.5) \quad \bar{\sigma}(x, y, 0) = \frac{1}{d_x(x, y)} = y\sqrt{1 - 2\rho\kappa z + \kappa^2 z^2}.$$

6.2 The Heston Model

In this model the driving process is given by

$$(6.6) \quad \begin{cases} dS_t = rS_t dt + \sqrt{y_t} S_t dW_t^1 \\ dy_t = a(\theta - y_t) dt + \kappa \sqrt{y_t} dW_t^2 \end{cases}$$

where a, θ , and κ are nonnegative constants. As was mentioned earlier, all of our results, which are stated for a process y_t valued in \mathbb{R} , extend in a straightforward way to the case $y_t \in (0, +\infty)$. Hence the signed geodesic distance d is the unique solution of

$$(6.7) \quad \begin{cases} yd_x^2 + 2\rho\kappa y d_x d_y + \kappa^2 y d_y^2 = 1 \\ d(x=0, y) = 0 \\ d(x, y) > 0 \quad \text{for } x > 0. \end{cases}$$

Once d is (numerically) computed, using Theorem 1.2 and 2.2 one easily infers the value of the asymptotic implied volatility $\Sigma^0(x, y) \equiv \Sigma(x, y, \tau = 0) = x/d(x, y)$ and the local volatility $\bar{\sigma}(x, y, 0) = 1/d_x(x, y)$.

6.3 Further Results

The main purpose of Theorem 1.2 was twofold. First, we determined the limiting value of the implied volatility function Σ in the limit $\tau \rightarrow 0$; second, we showed that Σ solves the nonlinear evolution problem (1.9). Following this methodology, it becomes clear that one can get further terms in a Taylor expansion in powers of τ . Indeed, this task is much simpler than determining the zero-order term, since any singularity is now removed from the problem. For this reason we

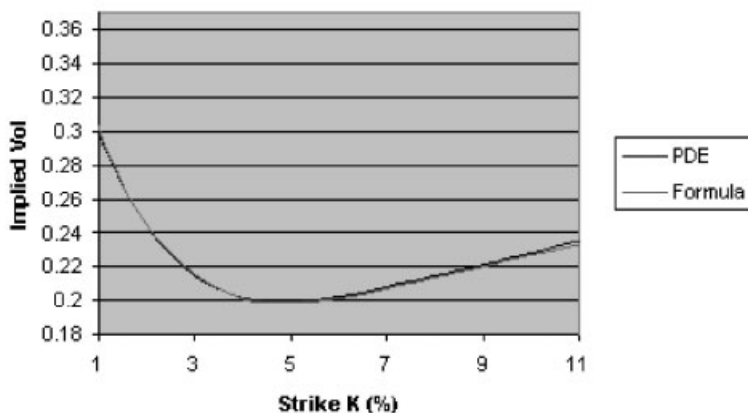


FIGURE 6.1. Implied volatility smile, formula (6.4) vs. PDE solution for model (6.1) with $r = 0$, $S_0 = 4\%$, $T = 2$ years, $\sigma \equiv 1$, $y_0 = 0.2$, $a = 0$, $\kappa = 0.3$, $\rho = -0.2$.

will not undertake this task in full generality; rather, we give an example in the case (6.1) with $a = \rho = 0$. If one makes the first-order time to maturity expansion

$$(6.8) \quad \Sigma(x, y, \tau) = \Sigma^0(x, y)(1 + \tau \Sigma^1(x, y) + O(\tau^2)),$$

it is not difficult to see that Σ^1 must satisfy the transport equation

$$(6.9) \quad 2\Sigma^1 + y^2\sigma^2(x)d d_x \Sigma_x^1 + \kappa^2 y^2 d d_y \Sigma_y^1 = \frac{1}{2}\sigma^2(x)y^2 \frac{\Sigma_{xx}^0}{\Sigma^0} + \frac{1}{2}\kappa^2 y^2 \frac{\Sigma_{yy}^0}{\Sigma^0},$$

where d is defined in (6.3). This equation happens to be explicitly solvable; we find

$$(6.10) \quad \Sigma^1(x, y) = -\frac{1}{d^2} \ln \left(\frac{x}{d(x, y)} \frac{1}{y\sqrt{\sigma(0)\sigma(x)}} \frac{1}{(1 + \kappa^2 z^2)^{1/4}} \right),$$

where the notation is defined in Section 6.1 above.

Remark. Following a completely different and interesting approach based on formal asymptotics and educated guesses, Hagan et al. [17] have independently addressed in the question of asymptotic expansions on implied volatilities in stochastic volatility models. That paper is in the same spirit as an earlier work [18] on local volatility models. They derive in particular two-term expansions analogous to (6.8). It should be noted, however, that formula (A65) in [17, p. 101] appears to be somewhat different from ours.

6.4 Numerical Examples

We examine the accuracy of the asymptotic formula in (1.9) by comparing it to benchmark prices computed by solving the pricing problem (1.7) on a refined finite difference grid. We address the example of the model described in Section 6.1 with

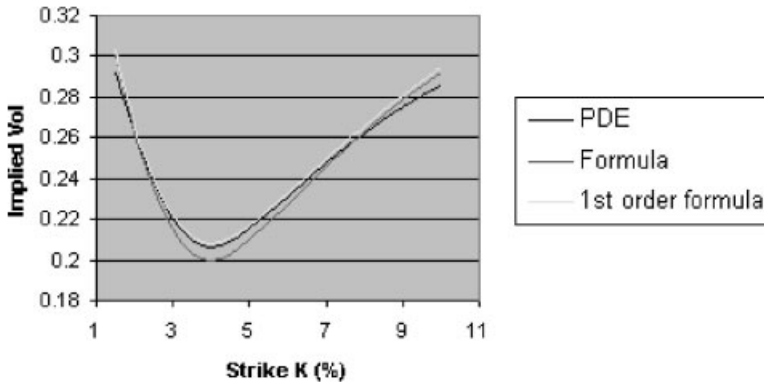


FIGURE 6.2. Implied volatility smile, formulae (6.4) and (6.10) vs. PDE solution for model (6.1) with $r = 0$, $S_0 = 4\%$, $T = 2$ years, $\sigma \equiv 1$, $\gamma_0 = 0.2$, $a = 0$, $\kappa = 0.5$, $\rho = 0$.

a negative asset/volatility correlation that leads to a characteristic “skew” phenomenon (Figure 6.1) and with a zero correlation that leads to a symmetric smile in log variables (Figure 6.2). Moreover, we illustrate the gain in accuracy provided by the two-term expansion (6.8) in Figure 6.2. We observe a satisfactory agreement between the asymptotic formula and the numerically computed smile.

Acknowledgments. This paper was written as part of a research program at HSBC CCF Derivatives Research. The authors thank M. Avellaneda for many useful discussions about this paper. The authors also thank J. M. Roquejoffre for helpful comments about Proposition 3.2 and O. Alvarez for reference [1].

Bibliography

- [1] Alvarez, O. Bounded-from-below viscosity solutions of Hamilton-Jacobi equations. *Differential Integral Equations* **10** (1997), no. 3, 419–436.
- [2] Avellaneda, M.; Boyer-Olson, D.; Busca, J.; Friz, P. Reconstruction of volatility: Pricing index options using the steepest-descent approximation. *Risk* (2002), October, 87–91.
- [3] Avellaneda, M.; Boyer-Olson, D.; Busca, J.; Friz, P. Application of large deviation methods to the pricing of index options in finance. *C. R. Math. Acad. Sci. Paris* **336** (2003), no. 3, 263–266.
- [4] Berestycki, H.; Busca, J.; Florent, I. An inverse parabolic problem arising in finance. *C. R. Acad. Sci. Paris Sér. I Math.* **331** (2000), no. 12, 965–969.
- [5] Berestycki, H.; Busca, J.; Florent, I. Asymptotics and calibration of local volatility models. *Quant. Finance* **2** (2002), no. 1, 61–69.
- [6] Black, F.; Scholes, M. The pricing of option and corporate liabilities. *J. Political Economy* **81** (1973), no. 3, 637–654.
- [7] Breeden, D. T.; Litzenberger, R. H. Prices of state contingent claims implicit in option prices. *J. Business* **51** (1978), no. 4, 621–651.
- [8] Cabré, X. Topics in regularity and qualitative properties of solutions of nonlinear elliptic equations. *Discrete Contin. Dyn. Syst.* **8** (2002), no. 2, 331–359.

- [9] Crandall, M.G.; Ishii, H.; Lions, P.-L. User's guide to viscosity solutions of second-order partial differential equations. *Bull. Amer. Math. Soc.* **27** (1992), no. 1, 1–67.
- [10] Crandall, M. G.; Lions, P.-L. Remarks on the existence and uniqueness of unbounded viscosity solutions of Hamilton-Jacobi equations. *Illinois J. Math.* **31** (1987), no. 4, 665–688.
- [11] Derman, E.; Kani, I. Riding on a smile. *Risk* **7** (1994), no. 2, 32–39.
- [12] Derman, E.; Kani, I. Stochastic implied trees: arbitrage pricing with stochastic term and strike structure of volatility. *Internat. J. Theoret. Appl. Finance* **1** (1998), no. 1, 61–110.
- [13] Derman, E.; Kani, I.; Chriss, N. Implied trinomial trees of the volatility smile. *J. Derivatives* **3** (1996), no. 4, 7–22.
- [14] Dupire, B. Pricing with a smile. *Risk* **7** (1994), no. 1, 18–20.
- [15] Friedman, A. *Partial differential equations of parabolic type*. Prentice-Hall, Englewood Cliffs, N.J., 1964.
- [16] Gilbarg, D.; Trudinger, N. S. *Elliptic partial differential equations of second order*. 2nd ed. Springer, Berlin–Heidelberg, 1998.
- [17] Hagan, P. S.; Kumar, D.; Lesniewski, A. S.; Woodward, D. E. Managing smile risk. *Wilmott* **18** (2002), no. 11, 84–108.
- [18] Hagan, P.; Woodward, D. E. Equivalent black volatilities. *Appl. Math. Finance* **6** (1999), 147–157.
- [19] Ishii, H. A simple, direct proof of uniqueness for solutions of the Hamilton-Jacobi equations of eikonal type. *Proc. Amer. Math. Soc.* **100** (1987), no. 2, 247–251.
- [20] Kružkov, S. N. Generalized solutions of Hamilton-Jacobi equations of eikonal type. I. Statement of the problems; existence, uniqueness and stability theorems; certain properties of the solutions. *Mat. Sb. (N.S.)* **98(140)** (1975), no. 3(11), 450–493.
- [21] Lee, R. W. Implied and local volatilities under stochastic volatility. *Internat. J. Theoret. Appl. Finance* **4** (2001), no. 1, 45–89.
- [22] Lions, P.-L. *Generalized solutions of Hamilton-Jacobi equations*. Research Notes in Mathematics, 69. Pitman, Boston–London, 1982.
- [23] Merton, R. C. Theory of rational pricing. *Bell. J. Economics* **4** (1973), no. 1, 141–183.
- [24] Varadhan, S. R. S. On the behavior of the fundamental solution of the heat equation with variable coefficients. *Comm. Pure Appl. Math.* **20** (1967), 431–455.

HENRI BERESTYCKI
 École des Hautes Études
 en Sciences Sociales
 C.A.M.S.
 54 Boulevard Raspail
 75270 Paris Cedex 06
 FRANCE
 E-mail: hb@ehess.fr

JÉRÔME BUSCA
 CNRS and Université Paris Dauphine
 CEREMADE
 Place Maréchal de Lattre de Tassigny
 75775 Paris Cedex 16
 FRANCE
 E-mail: jerome.busca@normalesup.org

IGOR FLORENT
 HSBC CCF
 103 Avenue des Champs-Élysées
 75419 Paris Cedex 08
 FRANCE
 E-mail: igor.florent@ccf.com

Received August 2003.