MARKOVIAN PROJECTION METHOD FOR VOLATILITY CALIBRATION

VLADIMIR PITERBARG

Abstract: We present the "Markovian projection method", a method to obtain closed-form approximations to European option prices on various underlyings that, in principle, is applicable to any (diffusive) model. Successful applications of the method have already appeared in the literature, in particular for interest rate models (short rate and forward Libor models with stochastic volatility), and interest rate/FX hybrid models with FX skew. The purpose of this note is thus not to present other instances where the Markovian projection method is applicable (even though more examples are indeed given) but to distill the essence of the method into a conceptually simple "plan of attack", a plan that anyone who wants to obtain European option approximations can follow.

First version: 25 May 2006
This version: 25 May 2006

1. Introduction

European-style options are usually the most liquid options available in any market. More often than not, they are the only options that are liquid enough to be used for model calibration. Thus, efficient methods for valuing European-style options are a critical requirement for any model. In this paper we develop the *Markovian projection* (MP) method, a very general and powerful approach to deriving accurate approximations to European option prices in a wide range of models.

Various ideas related to the Markovian projection method have already been used, and very successfully so, in a wide range of contexts. In Piterbarg (2004) and Piterbarg (2005b), closed-form approximations to European swaptions prices in a forward Libor model with stochastic volatility have been obtained. In Andreasen (2005), that same has been accomplished for a short rate model with stochastic volatility. Piterbarg (2006a) and Piterbarg (2005a) have introduced an FX skew in the interest rate/FX hybrid model and demonstrated how closed-form approximations for FX options can be derived. Same techniques can easily be adapted to interest rate/equity hybrids, for example. In this note, we formalize the ideas from these papers as the MP method.

The Markovian projection method is based on combining pioneering ideas of Dupire (see Dupire (1997)) on local and stochastic volatility (and recently rediscovered, in the quantitative finance context, results of Gyöngy, see Gyöngy (1986)) with parameter-averaging techniques that have been developed for stochastic volatility models relatively

---

Key words and phrases. Local volatility, stochastic volatility, Markovian projection, parameter averaging, Dupire's local volatility, index options, basket options, spread options.

I would like to thanks my colleagues at Barclays Capital, as well as Leif Andersen and Martin Forde, for valuable comments and discussions.

recently. Both Dupire and Gyöngy show that if two underlyings, each governed by its own SDEs, are given, and the expected values of the SDE coefficients conditioned on the appropriate underlying (the so-called "Dupire's local volatilities") are equal, then European options prices on (or, equivalently, one-dimensional marginal distributions of) the two underlyings are the same. This observation is the key tool to replace a model with complicated dynamics for the underlying with a much simpler one. Once that is accomplished, the (local or stochastic volatility) model that is obtained is further simplified, by applying the parameter averaging techniques, to a form that admits closed form solutions.

The MP method proceeds in four steps.

Step 1: For the underlying of interest, its SDE, driven by a single Brownian motion, is written down by computing its quadratic variance (and combining all dt terms if exist).

Step 2: In this SDE, the diffusion and drift (if exists) coefficients are replaced with their expected values conditional on the underlying. As shown in Section 2, this does not affect the values of European options.

Step 3: The conditional expected values from Step 2 are computed or, more commonly, approximated. Methods based on, or related to, Gaussian approximations are often used for this step, see Section 4 below.

Step 4: Parameter averaging techniques (see Piterbarg (2005c), Piterbarg (2005b)) are used to relate the time-dependent coefficients of the SDE obtained in Step 3 to time-independent ones. This, typically, allows for a quick and direct computation of European option values.

2. One-dimensional distributions of Ito processes

Suppose a model is given. Let $S(t)$ be the quantity of interest, European options on which are to be valued. For example, it could be a forward price of an equity share, a forward FX rate, a Libor or a swap rate, and so on. The market observable may or may not be a "primitive" variable in the model. For example, equity models are most often expressed in terms of the share price, and then the dynamics are imposed on $S(\cdot)$ directly. On the other hand, in the interest rate modeling, the dynamics are often imposed on short rates or on Libor rates. If $S(\cdot)$ is the swap rate, then it is not a "primitive" variable in such a model; the dynamics of $S(\cdot)$ are implicitly defined by the model expressed in terms of other variables. In either case, an Ito process for $S(\cdot)$, given by the model, can always be written down. It is of the general form

$$dS(t) = \Sigma(t, \dots)\, dW(t). \tag{1}$$

Here $\Sigma(t, \dots)$ is, in general, a stochastic process that may depend on $S(\cdot)$ and other sources of randomness. Note that we do not include the dt term. It will be clear from what follows that this is not a limitation, as the drifts can be handled as well. We write the process without the dt term largely by convention, as most often there is measure under which $S(\cdot)$ is a martingale, and it is precisely the measure that we are interested in studying the dynamics of $S(\cdot)$ under (eg the swap measure for a swap rate, or the forward measure for a forward FX rate).

Another feature of the way (1) is written is the fact that the Brownian motion $dW(\cdot)$ is one-dimensional. Again, this is not a limitation of the method. Even if $S(\cdot)$ is driven by a multi-dimensional Brownian motion, a one-dimensional Brownian motion, and the diffusion coefficient, can always be found such that the law of the original process and the new one are exactly the same. Let us elaborate. Suppose S follows

$$dS(t) = \sum_{i=1}^{d} \Sigma_i(t,\dots)\, dW_i(t), \tag{2}$$

where $(\Sigma_1(t,\dots),\dots,\Sigma_d(t,\dots))$ is a d-dimensional process and $(W_1(t),\dots,W_d(t))$ is a d-dimensional Brownian motion with the correlation matrix $\{\rho_{ij}\}_{i,j=1}^{d}$. Then, define $\Sigma(t,\dots)$ by

$$\Sigma^2(t,\dots) = \sum_{i,j=1}^{d} \Sigma_i(t,\dots)\Sigma_j(t,\dots)\rho_{ij}$$

and $dW(t)$ by

$$dW(t) = \Sigma^{-1}(t,\dots)\left(\sum_{i=1}^{d} \Sigma_i(t,\dots)\, dW_i(t)\right).$$

Simple quadratic variation calculations confirm that dW is indeed a (one-dimensional) Brownian motion and the law of the process $S(\cdot)$ given by (2) is the same as the one given by (1).

To price a European option on $S(\cdot)$ with expiry T and strike K, we need to know the one-dimensional distribution of $S(\cdot)$ (at time T). On the other hand, the knowledge of prices of all European options for all strikes K for a given expiry T is equivalent to knowing the (one-dimensional) distribution of $S(\cdot)$ at time T. In view of these remarks, the importance of the following result is evident ((Gyöngy, 1986, Theorem 4.6)).

Theorem 2.1 (Gyongy 1986)**.** *Let $X(t)$ be given by*

$$dX(t) = \alpha(t)\, dt + \beta(t)\, dW(t), \tag{3}$$

where $\alpha(\cdot), \beta(\cdot)$ are adapted bounded stochastic processes such that (3) admits a unique solution. Define $a(t,x), b(t,x)$ by

$$a(t,x) = \mathsf{E}\left(\alpha(t)\middle|X(t)=x\right),$$
$$b^2(t,x) = \mathsf{E}\left(\beta^2(t)\middle|X(t)=x\right).$$

Then the SDE

$$\begin{aligned} dY(t) &= a(t,Y(t))\, dt + b(t,Y(t))\, dW(t), \\ Y(0) &= X(0), \end{aligned} \tag{4}$$

admits a weak solution $Y(t)$ that has the same one-dimensional distributions as $X(t)$.

Remark 2.1. *The process $Y(\cdot)$ follows what is known in financial mathematics as a "local volatility" process. The function $b(t,x)$ is often called "Dupire's local volatility" for $X(\cdot)$, after Dupire who directly linked $\mathsf{E}\left(\beta^2(t)\middle|X(t)=x\right)$ to prices of European options on $X(\cdot)$ when $\alpha(\cdot)\equiv 0$ (see Dupire (1997)), a link that can be seen to lead to the same result.*

Remark 2.2. *Since $X(\cdot)$ and $Y(\cdot)$ have the same one-dimensional distributions, the prices of European options on $X(\cdot)$ and $Y(\cdot)$ for all strikes K and expiries T will be the same. Thus, for the purposes of European option valuation and/or calibration to European options, a potentially very complicated process $X(\cdot)$ can be replaced with a much simpler Markov process $Y(\cdot)$, the Markovian projection*[1] *of $X(\cdot)$.*

Remark 2.3. *Efficient valuation and approximation methods for European options in local volatility models, ie models of the type (4), are available, see for example Andersen and Brotherton-Ratcliffe (2001).*

Proof. The original proof in Gyöngy (1986) is fairly involved and we do not reproduce it here. Instead, we present an outline of the proof inspired by Savine (2000), a much more financially-motivated approach. As explained above, the case $\alpha(t) = 0$ is most common in financial applications, so that is what we consider.

Denote

$$c(t,K) = \mathsf{E}_0\left((X(t) - K)^+\right).$$

Here $\{c(t,K)\}_{t,K}$ are the values of European call options on $X(\cdot)$ for expiries t and strikes K. Dupire in Dupire (1994) was the first to find the expression for the local volatility in a local volatility model that reproduces all European options $\{c(t,K)\}_{t,K}$ for all expiries t and strikes K. The key insight of this proof is to note that the same technique can be applied to European option prices *obtained from another model*, instead of the market option prices as originally done by Dupire.

As explained above, matching all European option prices (for all expiries and strikes) is equivalent to matching all one-dimensional distributions. "Dupire local volatility" $b(t,x)$ in the model

$$dY(t) = b(t, Y(t))\, dW(t)$$

for the set of option prices $\{c(t,K)\}$ is given by

$$b^2(t,K) = 2\frac{\frac{\partial}{\partial t}c(t,K)}{\frac{\partial^2}{\partial K^2}c(t,K)}, \tag{5}$$

see Dupire (1994). To compute the right-had side, we first write (the use of delta-functions in the integrands can be justified by Tanaka's formula, see Karatzas and Shreve (1997))

$$d(X(t) - K)^+ = 1_{\{X(t)>K\}}dX(t) + \frac{1}{2}\delta_{\{X(t)=K\}}d\langle X(t)\rangle,$$

and, since $X(t)$ is a martingale under the measure considered,

$$\mathsf{E}(X(t) - K)^+ - (X(0) - K)^+ = \frac{1}{2}\int_0^t \mathsf{E}\left(\delta_{\{X(s)=K\}}d\langle X(s)\rangle\right).$$

Clearly

$$\mathsf{E}\left(\delta_{\{X(t)=K\}}d\langle X(t)\rangle\right) = \mathsf{E}\left(\delta_{\{X(t)=K\}}\right) \times \mathsf{E}\left(d\langle X(t)\rangle | X(t) = K\right)$$

and

$$\mathsf{E}\left(\delta_{\{X(t)=K\}}\right) = \frac{\partial^2}{\partial K^2}\mathsf{E}(X(t) - K)^+ = \frac{\partial^2}{\partial K^2}c(t,K).$$

[1] Often we will use the same letter to denote both, the original process and its Markovian projection.

From (3)

$$d\langle X(t)\rangle = \beta^2(t)\, dt,$$

so that

$$\mathsf{E}\left(\delta_{\{X(t)=K\}} d\langle X(t)\rangle\right) = \frac{\partial^2}{\partial K^2} c(t,K)^2 \times \mathsf{E}\left(\beta^2(t)\middle| X(t)=K\right) dt.$$

In particular,

$$\begin{aligned}\frac{\partial}{\partial t} c(t,K) &= \frac{\partial}{\partial t}\left(\mathsf{E}(X(t)-K)^+ - (X(0)-K)^+\right)\\ &= \frac{1}{2} \times \frac{\partial^2}{\partial K^2} c(t,K) \times \mathsf{E}\left(\beta^2(t)\middle| X(t)=K\right).\end{aligned}$$

Substituting this equality into (5) we obtain

$$b^2(t,K) = \mathsf{E}\left(\beta^2(t)\middle| X(t)=K\right),$$

and the theorem is proved.

∎

For the variable $S(\cdot)$ considered above, the theorem immediately implies that we can replace (1) with the Markovian projection

$$dS(t) = \tilde{\Sigma}(t,S(t))\, dW(t) \tag{6}$$

to price European options, where $\tilde{\Sigma}(t,x)$, the deterministic function of time and state, is obtained by

$$\tilde{\Sigma}^2(t,x) = \mathsf{E}\left(\Sigma^2(t,\dots)\middle| S(t)=x\right). \tag{7}$$

We emphasize that the Markovian projection is exact for European options but, of course, does not preserve the dependence structure of $S(\cdot)$ at different times or, equivalently, the marginal distributions of the process of orders higher than one. Thus, the prices of securities dependent on sampling of $S(\cdot)$ at multiple times, such as barriers, American options and so on, are different between the original model (1) and the projected model (6).

In practice, the conditional expected value in (7) needs to be computed/approximated. This is often impossible to do exactly, and various approximations need to be employed. We discuss common approaches later in the paper. Next, however, we consider another ways that Theorem 2.1 can be used.

3. Applications to stochastic volatility models

Suppose a stochastic volatility model is given. Let $S(\cdot)$ be the variable of interest as before, governed in the model by the equation (1). The results of Theorem 2.1 allow us to replace the dynamics with local volatility dynamics for the purposes of pricing European options. However, substituting a local volatility model for a stochastic volatility model is not always ideal. The two types of models are very different, and that makes computing the conditional expected value in (7) particularly difficult. This renders various possible approximations rather inaccurate. As a general rule, when one model is approximated with another, it is always best to keep as many features as possible the same in the two models. For example, a stochastic volatility model should be approximated with a (potentially simpler) stochastic volatility model, a model with jumps by a (simpler) model with jumps, and so on.

Upon some reflection, it should be clear that Theorem 2.1 provides us with a tool that is more general than just a way of approximating any model with a local volatility one. The following trivial Corollary to Theorem 2.1 provides the required framework.

Corollary 3.1. *If two processes have the same Dupire's local volatility, the European option prices on both are the same for all strikes and expiries.*

Consider a stochastic volatility model. Let $X_1(t)$ follow

$$dX_1(t) = b_1(t, X_1(t))\sqrt{z_1(t)}\,dW(t),$$

where $z_1(t)$ is some variance process. Suppose we would like to match the European option prices on $X_1(\cdot)$ (for all expiries and strikes) in a model of the form

$$dX_2(t) = b_2(t, X_2(t))\sqrt{z_2(t)}\,dW(t),$$

where $z_2(t)$ is a different, and potentially simpler, variance process. Then Corollary 3.1 and Theorem 2.1 imply that to achieve that, we need to choose $b_2(t,x)$ such that

$$\mathsf{E}\left(b_2^2(t,x)z_2(t)\middle| X_2(t) = x\right) = \mathsf{E}\left(b_1^2(t,x)z_1(t)\middle| X_1(t) = x\right),$$

or, rearranging the terms,

$$b_2^2(t,x) = b_1^2(t,x)\frac{\mathsf{E}(z_1(t)|X_1(t) = x)}{\mathsf{E}(z_2(t)|X_2(t) = x)}. \tag{8}$$

To apply this formula, the conditional expected values need to be computed. Note that it is beneficial to choose z_2 as similar to z_1 as possible. Then, various approximations employed to compute both conditional expected values will tend to cancel each other, thus increasing the accuracy of calculating $b_2(t,x)$.

It is generally the case that a realistic stochastic volatility model with a local volatility component should have a local volatility that does not have a lot of curvature (such as a linear or a power function of the underlying). Moreover, such a choice makes many calculations and approximations much easier, as then local volatility functions can just be described by two numbers, a value and a slope (at the forward value of the underlying, for example). In practice it means that when a stochastic volatility model is constructed to fit implied volatility smiles (be that market volatilities or implied volatilities from a different model), it is better to choose the stochastic volatility that explains as much of the *curvature* of the implied volatility smile as possible. Then the local volatility function will be "almost linear", and the benefits mentioned above will apply. For the concrete example of equation (8), one should typically try to find $z_2(\cdot)$ that has the same term variance as $z_1(\cdot)$, as it is the volatility of stochastic variance that defines the curvature of the smile.

Another typical application of Corollary 3.1 proceeds as follows. Suppose

$$dX_1(t) = \beta(t)\sqrt{z(t)}\,dW(t),$$

where $z(t)$ is the stochastic variance process, but $\beta(t)$ is not deterministic, but a stochastic process in its own right (eg a complicated function of the state variables in a term structure model of interest rates, with $S(\cdot)$ being a swap rate). We would like to replace the SDE with a local-stochastic volatility model,

$$dX_2(t) = b(t, X_2(t))\sqrt{z(t)}\,dW(t).$$

Then, according to Corollary 3.1, we need to set

$$b^2(t,x) = \frac{\mathrm{E}\left(\beta^2(t)z(t)\middle|X_1(t)=x\right)}{\mathrm{E}\left(z(t)\middle|X_2(t)=x\right)}. \tag{9}$$

This formula can be simplified when, for example, $\beta(\cdot)$ and $z(\cdot)$ are (approximately) conditionally independent given $X_1(\cdot)$. Then

$$b^2(t,x) \approx \mathrm{E}\left(\beta^2(t)\middle|X_1(t)=x\right)\frac{\mathrm{E}\left(z(t)\middle|X_1(t)=x\right)}{\mathrm{E}\left(z(t)\middle|X_2(t)=x\right)}. \tag{10}$$

In many real-world situations, it can be safely assumed that $\mathrm{E}(z(t)|X_1(t)=x) \approx \mathrm{E}(z(t)|X_2(t)=x)$, in which case the formula simplifies further,

$$b^2(t,x) \approx \mathrm{E}\left(\beta^2(t)\middle|X_1(t)=x\right). \tag{11}$$

4. GAUSSIAN APPROXIMATIONS FOR CONDITIONAL EXPECTED VALUE CALCULATIONS

Recall the set up of Section 2, namely the general equation of the type (1) for an underlying of interest. Calculations of conditional expected values, as in (7) (or (8)) are often the most difficult part of the MP method. Tractable formulas are hard to find for any type of process, with one notable exception. If in (7) the process $\left(S(t), \Sigma^2(t,\dots)\right)$ is jointly Gaussian, then the calculations are straightforward. While this is the case in only the trivial situation of $\Sigma^2(t,\dots)$ being deterministic, the availability of closed-form solutions in the Gaussian case gives rise to the approximation in which the actual dynamics of $S(t)$ and $\Sigma^2(t)(=\Sigma^2(t,\dots))$ are replaced by the Gaussian ones. While the details vary from case to case, the general approach is given by the following proposition.

Proposition 4.1. *Let the dynamics of $\left(S(t), \Sigma^2(t)\right)$ be written in the following form,*

$$\begin{aligned} dS(t) &= \Sigma(t)\,dW(t), \\ d\Sigma^2(t) &= \eta(t)\,dt + \varepsilon(t)\,dB(t), \end{aligned} \tag{12}$$

where $\Sigma(t)$, $\eta(t)$ and $\varepsilon(t)$ are adapted stochastic processes (that may depend on $S(\cdot)$, $\Sigma^2(\cdot)$, etc) and $W(t), B(\cdot)$ are both Brownian motions. Then the conditional expected value in (7) can be approximated by

$$\begin{aligned} \tilde{\Sigma}^2(t,x) &\triangleq \mathrm{E}\left(\Sigma^2(t)\middle|S(t)=x\right) \\ &= \overline{\Sigma^2}(t) + r(t)(x - S_0), \\ r(t) &= \frac{\int_0^t \bar{\varepsilon}(s)\bar{\Sigma}(s)\bar{\rho}(s)\,ds}{\int_0^t \bar{\Sigma}^2(s)\,ds}, \end{aligned} \tag{13}$$

where $\bar{\varepsilon}(t)$, $\overline{\Sigma^2}(t)$, $\bar{\rho}(t)$, $\bar{\Sigma}(t)$ *are, respectively, deterministic approximations to* $\varepsilon(t)$, $\Sigma^2(t)$, $\rho(t) \triangleq \langle dW(t), dB(t)\rangle / dt$, *and* $\Sigma(t)$. *In particular, one can take*

$$\begin{aligned}
\bar{\varepsilon}(t) &= \mathsf{E}\varepsilon(t), \\
\overline{\Sigma^2}(t) &= \int_0^t (\mathsf{E}\eta(s))\, ds, \\
\bar{\rho}(t) &= \mathsf{E}(\langle dW(t), dB(t)\rangle / dt), \\
\bar{\Sigma}(t) &= \sqrt{\overline{\Sigma^2}(t)}.
\end{aligned}$$

Proof. If (X,Y) is a Gaussian vector, then

$$\mathsf{E}(X|Y) = \mathsf{E}X + \frac{\mathsf{covar}(X,Y)}{\mathsf{Var}(Y)}(Y - \mathsf{E}Y). \tag{14}$$

The result follows by computing the required moments in a Gaussian approximation to the dynamics of (12).■

Remark 4.1. *If more is known about the form of the dynamics (12), other deterministic approximations to* $\varepsilon(t)$ *(and other parameters) can be more useful.*

Remark 4.2. *The approximation (13) is linear in* x. *The Gaussian approximation, hence, works best when the local volatility functions that are used are either linear or close to linear. This provides another angle to the comment at the end of Section 3, namely that the most practical technique for computing conditional expected values, the Gaussian approximation, is not going to work well when the curvature in the volatility smile is explained by the local volatility function, rather than by the stochastic volatility component.*

Sometimes it is beneficial to apply approximations that are not Gaussian, but linked to them. For example, a lognormal approximation can be used in the place of a Gaussian one. Then, the approximation is based on the known expression for $\mathsf{E}(X|Y)$ when (X,Y) are jointly lognormal. Other twists involving "mapping functions" $Y \to f(Y)$ are also possible.

5. Parameter averaging

Parameter averaging methods are extensively described elsewhere (see Piterbarg (2005c), Piterbarg (2005b)). Here, we present a short summary of the results.

Consider a stochastic volatility model of the form

$$\begin{aligned}
dz(t) &= \theta(1 - z(t))\, dt + \gamma(t)\sqrt{z(t)}\, dV(t), \\
dS(t) &= (\beta(t)S(t) + (1-\beta(t))S(0))\,\sigma(t)\sqrt{z(t)}\, dW(t), \\
z(0) &= 1, \quad S(0) = S_0.
\end{aligned} \tag{15}$$

with time-dependent volatility $\sigma(t)$, skew $\beta(t)$, and volatility of variance $\gamma(t)$. Fix a particular $T > 0$.

Theorem 5.1. *Prices of European options in the model (15) with expiry $T > 0$ are well-approximated by the European option prices in the following model with constant coefficients,*

$$dz(t) = \theta (1 - z(t))\, dt + \eta \sqrt{z(t)}\, dV(t), \tag{16}$$
$$dS(t) = (bS(t) + (1-b) S(0))\, \lambda \sqrt{z(t)}\, dW(t),$$

where

$$\eta^2 = \frac{\int_0^T \gamma^2(t) \rho(t)\, dt}{\int_0^T \rho(t)\, dt},$$
$$b = \frac{\int_0^T \beta(t) v^2(t) \sigma^2(t)\, dt}{\int_0^T v^2(t) \sigma^2(t)\, dt},$$

and λ is a solution to

$$\mathbf{E} \exp\left(\frac{g''(\zeta)}{g'(\zeta)} \int_0^T \sigma^2(t) z(t)\, dt \right) = \mathbf{E} \exp\left(\lambda^2 \frac{g''(\zeta)}{g'(\zeta)} \int_0^T z(t)\, dt \right).$$

Here

$$\rho(r) = \int_r^T ds \int_s^T dt\, \sigma^2(t) \sigma^2(s) e^{-\theta(t-s)} e^{-2\theta(s-r)},$$
$$v^2(t) = \int_0^t \sigma^2(s)\, ds + \eta^2 e^{-\theta t} \int_0^t \sigma^2(s) \frac{e^{\theta s} - e^{-\theta s}}{2\theta} ds,$$
$$g(x) = \frac{S_0}{b} \left(2\Phi\left(b\sqrt{x}/2 \right) - 1 \right),$$
$$\zeta = \int_0^T \sigma^2(t)\, dt,$$

and $\Phi(\cdot)$ is the standard Gaussian cumulative distribution function.

Proof. See Piterbarg (2005c), Piterbarg (2005b). ∎

Remark 5.1. *Efficient numerical methods to value European options in the model (16) are available, see Andersen and Andreasen (2002).*

Remark 5.2. *Since the parameter averaging methods are especially well developed for the model of the type (15), in the applications of the Markovian projection method it is beneficial to try to reduce the dynamics of the quantity of interest, $S(\cdot)$, to this particular form. One should realize, however, that the only restriction imposed by (15) is the particular form of the dynamics of the stochastic variance $z(\cdot)$ (a mean reverting square root process). The form of the local volatility component of $S(\cdot)$ can be seen to be just a linearization of whatever the actual local volatility function is used.*

Remark 5.3. *The averaging results have been derived in the case of zero asset-volatility correlation, $\langle dW, dV \rangle = 0$. Experience shows, however, that the approximations work very well even if the correlation is not zero.*

6. Examples

In this section we consider a couple of examples of varying degrees of complexity, to demonstrate typical applications of the method we developed. The reader will note that in the examples, the use of the parameter averaging methods is not required, as the approximate equations already have constant coefficients. This is done on purpose, as the applications of the PA methods have been extensively exposed elsewhere, and we try to simplify the presentation on the account of clarity.

6.1. **Example 1: Options on an index in local volatility models.** Consider a collection of N assets $\mathbf{S}(t) = (S_1(t), \ldots, S_N(t))^\top$ each driven by its own local volatility model

$$dS_n(t) = \varphi_n(S_n(t))\, dW_n(t),$$
$$n = 1, \ldots, N,$$

with a vector of Brownian motions $(W_1(t), \ldots, W_N(t))$ with an $N \times N$ correlation matrix

$$\langle dW_i, dW_j \rangle = \rho_{ij}, \quad i, j = 1, \ldots, N.$$

We restrict the set of local volatility functions $\varphi_n(x)$ to the set of functions well-described by its value and the first order derivative at the forward. In effect, we require $\varphi_n(x)$ to be linear, or close to linear (such as CEV). As always, this is not really a restriction as using more complicated local volatility functions is not a good idea anyway.

Define an *index* $S(t)$ by

$$S(t) = \sum_{n=1}^{N} w_n S_n(t) = \mathbf{w}^\top \mathbf{S}(t).$$

An index is a weighted average of the assets. We would like to derive an approximation formula to (European) options on the index $S(t)$. This problem was considered first in Avellaneda et al. (2002).

Applying Ito's lemma to $S(t)$ we obtain that

$$dS(t) = \sum_{n=1}^{N} w_n \varphi_n(S_n(t))\, dW_n(t).$$

If we define

$$\sigma^2(t) = \sum_{n,m=1}^{N} w_n w_m \varphi_n(S_n(t))\, \varphi_m(S_m(t))\, \rho_{nm},$$
$$dW(t) = \frac{1}{\sigma(t)} \sum_{n=1}^{N} w_n \varphi_n(S_n(t))\, dW_n(t),$$

then

$$dS(t) = \sigma(t)\, dW(t),$$

and, by computing its quadratic variation, we obtain that dW is a standard Brownian motion. The process $\sigma(t)$ is, of course, a complicated function of $\mathbf{S}(t)$. By the results of

Section 2 and, in particular, Theorem 2.1, for the purposes of European option valuation, we can replace this SDE with a simple one,

$$dS(t) = \varphi(t, S(t))\, dW(t)$$

where

$$\varphi^2(t,x) = \mathsf{E}\left(\sigma^2(t) \middle| S(t) = x\right).$$

We have,

$$\varphi^2(t,x) = \sum_{n,m=1}^{N} w_n w_m \rho_{nm} \mathsf{E}\left(\varphi_n(S_n(t))\,\varphi_m(S_m(t)) | S(t) = x\right).$$

By expanding each term $\varphi_n(S_n(t))\,\varphi_m(S_m(t))$ to the first order around the forward (a good approximation by near-linearity assumption) we obtain,

$$\begin{aligned}\varphi_n(S_n(t))\,\varphi_m(S_m(t)) \approx\ & p_n p_m \\ & + p_n q_m (S_m(t) - S_m(0)) \\ & + p_m q_n (S_n(t) - S_n(0)),\end{aligned}$$

so that

$$\varphi^2(t,x) \approx \sum_{n,m=1}^{N} w_n w_m \rho_{nm} p_n p_m (1 + \chi_n(x) + \chi_m(x)), \tag{17}$$

where

$$\chi_n(x) = \frac{q_n}{p_n} \mathsf{E}\left(S_n(t) - S_n(0) | S(t) = x\right), \tag{18}$$

and

$$p_n = \varphi_n(S_n(0)), \quad q_n = \varphi_n'(S_n(0)).$$

To compute the conditional expected value $\mathsf{E}\left(S_n(t) - S_n(0) | S(t) = x\right)$ we apply the Gaussian approximation,

$$\mathsf{E}\left(S_n(t) - S_n(0) | S(t) = x\right) \approx \mathsf{E}\left(\bar{S}_n(t) - S_n(0) \middle| \bar{S}(t) = x\right), \tag{19}$$

where

$$\begin{aligned} d\bar{S}_n(t) &= p_n\, dW_n(t), \\ d\bar{S}(t) &= p\, d\bar{W}(t), \end{aligned}$$

and

$$p^2 = \sum_{n,m=1}^{N} w_n w_m p_n p_m \rho_{nm}, \tag{20}$$

$$d\bar{W}(t) = \frac{1}{p} \sum_{n=1}^{N} w_n p_n\, dW_n(t).$$

In this model,

$$\mathsf{E}\left(\bar{S}_n(t) - S_n(0) \middle| \bar{S}(t) = x\right) = \frac{\left\langle \bar{S}_n(t), \bar{S}(t) \right\rangle}{\left\langle \bar{S}(t), \bar{S}(t) \right\rangle}(x - S(0)) \\ = \rho_n \frac{p_n}{p}(x - S(0)), \tag{21}$$

where

$$\rho_n \triangleq \langle d\bar{W}(t), dW_n(t) \rangle / dt \\ = \frac{1}{p} \sum_{m=1}^{N} w_m p_m \rho_{nm}. \tag{22}$$

Combining (18), (19), (21) we obtain

$$\chi_n(x) = q_n \frac{\rho_n}{p}(x - S(0)).$$

Substituting it into (17), we obtain the following proposition.

Proposition 6.1. *For the purposes of European option pricing, the dynamics of the index $S(t)$ can be approximated with the following SDE,*

$$dS(t) = \varphi(S(t))\, dW(t),$$

where $\varphi(x)$ is such that

$$\varphi(S(0)) = p, \quad \varphi'(S(0)) = q,$$

with p given by (20),

$$q \triangleq \frac{1}{p^2} \sum_{n,m=1}^{N} w_n w_m p_n p_m \rho_{nm} \frac{q_n \rho_n + q_m \rho_m}{2}, \tag{23}$$

and $\{\rho_n\}$'s given by (22).

Remark 6.1. *Note that in our approximation, $\varphi(x)$ does not depend on t; this will not always be the case. If the local volatility function depends on t, then parameter averaging methods can be used to find the "'aggregate" volatility and skew.*

Remark 6.2. *The parameter q is the slope of the local volatility function for the index. Note that the expression (23) represents it as a weighted average of the slopes of the local volatility functions of the index's components q_n, and correlations ρ_n between the Brownian motions driving the components and the index. Another expression can be easily derived from (23) by manipulating the order of summation,*

$$q \triangleq \frac{1}{p} \sum_{n=1}^{N} w_n p_n \rho_n^2 q_n. \tag{24}$$

Remark 6.3. *The approximations we develop work best when all q_n's are "not very different". This would be the case when, for example, a swap rate is represented as a weighted average of contiguous Libor rates, and we can expect their skews to be similar. In other cases, when q_n's can be significantly different, as for example, in an FX rate basket, non-linear approximations of Avellaneda et al. (2002) have better accuracy (but are more complicated).*

6.2. **Example 2: Spread options with stochastic volatility.** Consider the problem of valuing spread options, ie (European) options on $S(t) \triangleq S_1(t) - S_2(t)$, where each of the assets S_1, S_2 follows *its own* stochastic volatility process (as pointed out in Piterbarg (2006b), spread options are very sensitive to stochastic variance decorrelation, an often overlooked fact). In particular we define

$$
\begin{aligned}
(25) \qquad & dS_i(t) = \varphi_i(S_i(t))\sqrt{z_i(t)}\,dW_i(t), \\
& dz_i(t) = \theta(1 - z_i(t))\,dt + \eta_i\sqrt{z_i(t)}\,dW_{2+i}(t), \quad z_i(0) = 1, \\
& \quad i = 1, 2,
\end{aligned}
$$

with the correlations given by

$$\left\langle dW_i(t), dW_j(t)\right\rangle = \rho_{ij}, \quad i, j = 1, \ldots, 4.$$

The spread $S(t)$ is the underlying of interest; the first step is to write an SDE for it. We have,

$$dS(t) = \sigma(t)\,dW(t), \tag{26}$$

where

$$
\begin{aligned}
(27) \qquad \sigma^2(t) &= (\varphi_1(S_1(t))u_1(t))^2 - 2(\varphi_1(S_1(t))u_1(t))(\varphi_2(S_2(t))u_2(t))\rho_{12} \\
&\quad + (\varphi_2(S_2(t))u_2(t))^2, \\
dW(t) &= \frac{1}{\sigma(t)}(\varphi_1(S_1(t))u_1(t)\,dW_1(t) - \varphi_2(S_2(t))u_2(t)\,dW_2(t)), \\
u_i(t) &= \sqrt{z_i(t)}, \quad i = 1, 2.
\end{aligned}
$$

Using the same notations as in the previous example, we set

$$p_i = \varphi_i(S_i(0)), \quad q_i = \varphi_i'(S_i(0)), \quad i = 1, 2.$$

As explained previously, directly replacing $\sigma^2(t)$ (in (26)) with its expected value conditional on $S(\cdot)$ is not a good idea when stochastic volatility is involved. Per Corollary 3.1 and the discussion in Section 3, we shall try to find a stochastic volatility process $z(\cdot)$ such that the curvature of the smile of the spread $S(\cdot)$ is explained by it, and the local volatility function is only used to induce the volatility skew.

To identify a suitable candidate for $z(\cdot)$, let us consider what the expression for $\sigma^2(t)$ would be if the local volatility functions for the components, $\varphi_i(x)$, $i = 1, 2$, were constant functions. We see that in this case, the expression for $\sigma^2(t)$ in (27) would not involve the processes $S_i(\cdot)$, $i = 1, 2$ and thus would be a good candidate for the stochastic variance process. Hence we define (the division by $p^2 \triangleq \sigma^2(0)$ is to preserve the scaling $z(0) = 1$)

$$z(t) = \frac{1}{p^2}\left(p_1^2 z_1(t) - 2p_1p_2\rho_{12}\sqrt{z_1(t)z_2(t)} + p_2^2 z_2(t)\right), \tag{28}$$

where

$$p = \sigma(0) = \left(p_1^2 - 2p_1p_2\rho_{12} + p_2^2\right)^{1/2}. \tag{29}$$

Proposition 6.2. *For the purposes of European option pricing, the dynamics of the spread $S(t)$ can be approximated with the following SDE,*

$$dS(t) = \varphi(S(t))\sqrt{z(t)}\,dW(t),$$

where $z(\cdot)$ given by (28), the function $\varphi(\cdot)$ satisfies

$$\varphi(S(0)) = p, \quad \varphi'(S(0)) = q,$$

with

$$q \triangleq \frac{1}{p}\left(p_1\rho_1^2 q_1 - p_2\rho_2^2 q_2\right),$$

and ρ_1, ρ_2 are defined by

$$\begin{aligned} \rho_1 &= \frac{1}{p}(p_1 - p_2\rho_{12}), \\ \rho_2 &= \frac{1}{p}(p_1\rho_{12} - p_2). \end{aligned} \tag{30}$$

Proof. By Corollary 3.1, the right $\varphi(x)$ id given by

$$\varphi^2(t,x) = \frac{\mathsf{E}\left(\sigma^2(t)\middle|\, S(t)=x\right)}{\mathsf{E}\left(z(t)|\, S(t)=x\right)}. \tag{31}$$

The expression for $\mathsf{E}\left(\sigma^2(t)\middle|\, S(t)=x\right)$ is a linear combinations of the conditional expected values of the terms

$$\varphi_i(S_i(t))\,\varphi_j\left(S_j(t)\right)u_i(t)\,u_j(t).$$

Each one is approximated to the first order by

$$p_i p_j\left(1 + \frac{q_i}{p_i}(S_i(t) - S_i(0)) + \frac{q_j}{p_j}\left(S_j(t) - S_j(0)\right) + (u_i(t) - 1) + \left(u_j(t) - 1\right)\right).$$

As in Example 1, the conditional expected values $\mathsf{E}(S_i(t) - S_i(0)|\, S(t) = x)$ can be easily computed in the Gaussian approximation. The same idea can be applied to $\mathsf{E}(u_i(t) - 1|\, S(t) = x)$. In particular,

$$\begin{aligned} \mathsf{E}(S_i(t) - S_i(0)|\, S(t) = x) &\approx \mathsf{E}\left(\bar{S}_i(t) - \bar{S}_i(0)\middle|\, \bar{S}(t) = x\right) \\ \mathsf{E}(u_i(t) - 1|\, S(t) = x) &\approx \mathsf{E}\left(\bar{u}_i(t) - 1|\, \bar{S}(t) = x\right), \end{aligned}$$

where (dt terms are ignored, although they may be included for more accurate approximations)

$$\begin{aligned} d\bar{S}(t) &= p\,d\bar{W}(t), \\ d\bar{S}_i(t) &= p_i\,dW_i(t) \\ d\bar{u}_i(t) &= \frac{\eta_i}{2}\,dW_{2+i}(t), \\ d\bar{W}(t) &= \frac{1}{p}(p_1\,dW_1(t) - p_2\,dW_2(t)), \end{aligned} \tag{32}$$

and

$$\mathsf{E}\left(\bar{S_i}(t)-\bar{S_i}(0)\middle|\bar{S}(t)=x\right)=\frac{p_i\rho_i}{p}(x-S(0)),$$
$$\mathsf{E}\left(\bar{u_i}(t)-1\middle|\bar{S}(t)=x\right)=\frac{\eta_i\rho_{2+i}}{2p}(x-S(0)),$$
$$i=1,2.$$

Here, as in the previous example, we have denoted $\rho_i\triangleq\langle d\bar{W}(t),dW_i(t)\rangle/dt$, $i=1,\dots,4$, so that (see (30)),

$$\rho_i=\frac{1}{p}(p_1\rho_{i1}-p_2\rho_{i2}),\quad i=1,\dots,4. \tag{33}$$

Under the linear approximations

$$\begin{aligned}
&\mathsf{E}\left(\sigma^2(t)\middle|S(t)=x\right)\\
&=p_1^2\left(1+2q_1\rho_1\frac{x-S(0)}{p}+\eta_1\rho_3\frac{x-S(0)}{p}\right)\\
&-2p_1p_2\rho_{12}\left(1+(q_1\rho_1+q_2\rho_2)\frac{x-S(0)}{p}+\left(\frac{\eta_1\rho_3+\eta_2\rho_4}{2}\right)\frac{x-S(0)}{p}\right)\\
&+p_2^2\left(1+2q_2\rho_2\frac{x-S(0)}{p}+\eta_2\rho_4\frac{x-S(0)}{p}\right)
\end{aligned}$$

Then,

$$\mathsf{E}\left(\sigma^2(t)\middle|S(t)=x\right)=p^2\left(1+2q_u\frac{x-S(0)}{p}\right), \tag{34}$$

where

$$\begin{aligned}
q_u\;=\;&\frac{1}{2p^2}\left[p_1^2(2q_1\rho_1+\eta_1\rho_3)\right.\\
&\left.-2p_1p_2\rho_{12}\left((q_1\rho_1+q_2\rho_2)+\left(\frac{\eta_1\rho_3+\eta_2\rho_4}{2}\right)\right)+p_2^2(2q_2\rho_2+\eta_2\rho_4)\right].
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\mathsf{E}(z(t)|S(t)=x)=p_1^2&\left(1+\eta_1\rho_3\frac{x-S(0)}{p}\right)/p^2\\
&-2p_1p_2\rho_{12}\left(1+\left(\frac{\eta_1\rho_3+\eta_2\rho_4}{2}\right)\frac{x-S(0)}{p}\right)/p^2\\
&+p_2^2\left(1+\eta_2\rho_4\frac{x-S(0)}{p}\right)/p^2,
\end{aligned}$$

so that

$$\begin{aligned}
\mathsf{E}(z(t)|S(t)=x)&=1+2q_l\frac{x-S(0)}{p},\\
q_l&=\frac{1}{2p^2}\left(p_1^2\eta_1\rho_3-2p_1p_2\rho_{12}\left(\frac{\eta_1\rho_3+\eta_2\rho_4}{2}\right)+p_2^2\eta_2\rho_4\right).
\end{aligned} \tag{35}$$

The quantities $\varphi(S(0))$, $\varphi'(S(0))$ are computed from the equation (31) and the expressions (34), (35) for the numerator and the denominator, as we note that

$$q = q_u - q_l.$$

∎

To effectively apply the result of Proposition 6.2, we still need to derive a simple form for the stochastic variance process $z(\cdot)$, defined by (28). Ideally, we should try to write the SDE for it in the same form as the SDEs for z_i's in (25). This cannot be done exactly; so we approximate.

Proposition 6.3. *For the purposes of European option pricing, the dynamics of the spread $S(t)$ can be approximated with the following SDE,*

$$\begin{aligned} dS(t) &= \varphi(S(t))\sqrt{z(t)}\,dW(t), \\ dz(t) &= \theta\left(1+\frac{\gamma}{\theta}-z(t)\right)dt+\bar{\eta}\sqrt{z(t)}d\bar{B}(t), \end{aligned}$$

where $\varphi(x)$ is given in Proposition 6.2,

$$\begin{aligned} \bar{\eta}^2 &= \frac{1}{p^2}\left((p_1\eta_1\rho_1)^2-2(p_1\eta_1\rho_1)(p_2\eta_2\rho_2)\rho_{34}+(p_2\eta_2\rho_2)^2\right), \\ d\bar{B}(t) &= \frac{1}{\bar{\eta}p}(p_1\eta_1\rho_1\,dW_3(t)-p_2\eta_2\rho_2\,dW_4(t)), \end{aligned}$$

and γ is given by

$$\gamma \triangleq \frac{p_1p_2\rho_{12}}{4p^2}\left(\eta_1^2-2\eta_1\eta_2\rho_{34}+\eta_2^2\right). \tag{36}$$

The (approximation of the) correlation between dW and $d\bar{B}$ is given by

$$\langle dW, d\bar{B}\rangle = \frac{1}{p^2\bar{\eta}}\left(p_1^2\eta_1\rho_1\rho_{13}-p_1p_2(\eta_1\rho_1\rho_{23}+\eta_2\rho_2\rho_{14})+p_2^2\eta_2\rho_2\rho_{24}\right).$$

Proof. For $z(\cdot)$ defined by (28)

$$dz(t) = \delta_1(t)\,dt+\delta_2(t)\,dt+\delta_3(t)\,dt+\eta(t)\,dB(t),$$

where

$$\begin{aligned} \delta_1(t) &= \theta\frac{p_1^2}{p^2}\left(1-\frac{p_2}{p_1}\rho_{12}\sqrt{\frac{z_2(t)}{z_1(t)}}\right)(1-z_1(t)), \\ \delta_2(t) &= \theta\frac{p_2^2}{p^2}\left(1-\frac{p_1}{p_2}\rho_{12}\sqrt{\frac{z_1(t)}{z_2(t)}}\right)(1-z_2(t)), \\ \delta_3(t) &= \frac{p_1p_2\rho_{12}}{4p^2}\left(\sqrt{\frac{z_2(t)}{z_1(t)}}\eta_1^2-2\eta_1\eta_2\rho_{34}+\sqrt{\frac{z_1(t)}{z_2(t)}}\eta_2^2\right), \end{aligned}$$

and

$$\begin{aligned}
\eta^2(t) &= v_1^2(t) + 2\rho_{34} v_1(t) v_2(t) + v_2^2(t), \\
dB(t) &= \frac{1}{\eta(t)} \left(v_1(t)\, dW_3(t) + v_2(t)\, dW_4(t) \right), \\
v_1(t) &= \eta_1 \frac{p_1^2}{p^2} \left(\sqrt{z_1(t)} - \frac{p_2}{p_1} \rho_{12} \sqrt{z_2(t)} \right), \\
v_2(t) &= \eta_2 \frac{p_2^2}{p^2} \left(\sqrt{z_2(t)} - \frac{p_1}{p_2} \rho_{12} \sqrt{z_1(t)} \right).
\end{aligned}$$

The expressions are quite complicated, and also not “closed” in $z(\cdot)$. Theorem 2.1 can be applied again, now to the process for $z(\cdot)$. The motivation for that is that the curvature of the volatility smile (of options on $S(\cdot)$) is driven by the variance of the stochastic variance, and that is preserved under the Markovian projection of $z(\cdot)$. Were we do that, we would need to compute conditional expected values of the type $\mathsf{E}\left(\sqrt{z_i(t) z_j(t)} \,\middle|\, z(t) = x\right)$ and $\mathsf{E}\left(\sqrt{z_i(t)/z_j(t)} \,\middle|\, z(t) = x\right)$, for which we would apply the Gaussian approximation method. For the purposes of this example we, however, employ a much more ad-hoc (and simpler) approximations, in which we

- replace $\sqrt{z_1(t)}$, $\sqrt{z_2(t)}$ with $\sqrt{z(t)}$;
- replace $\sqrt{\frac{z_2(t)}{z_1(t)}}$, $\sqrt{\frac{z_1(t)}{z_2(t)}}$ with 1.

Under this approximations,

$$\begin{aligned}
\delta_1(t) + \delta_2(t) &= \theta(1 - z), \\
\delta_3(t) &= \gamma, \\
\eta(t) &= \bar{\eta}\sqrt{z(t)}, \\
B(t) &= \bar{B}(t),
\end{aligned}$$

where γ, $\bar{\eta}$, and $\bar{B}(\cdot)$ are given in the statement of the proposition. The proposition follows.
∎

Test results show that the approximation is quite accurate, although it does deteriorate for longer expiries. The accuracy can be improved by calculating various terms more accurately. For example, let us demonstrate how this can be done for the terms $\sqrt{\frac{z_2(t)}{z_1(t)}}$, $\sqrt{\frac{z_1(t)}{z_2(t)}}$, $\sqrt{z_i(t) z_j(t)}$. The results of Proposition 6.3 were obtained by replacing them with 1. A more accurate approximation is obtained by calculating their expected values in a log-normal approximation.

Recall that

$$\mathsf{E} z_i(t) = 1, \quad \mathsf{Var}(z_i(t)) = \frac{\eta_i^2 \rho_{2+i,2+i}}{2\theta}\left(1 - e^{-2\theta t}\right), \quad i = 1, 2. \tag{37}$$

The covariance between $z_1(t)$ and $z_2(t)$ can be approximated by[2]

$$\mathsf{covar}(z_1(t), z_2(t)) = \frac{\eta_1\eta_2\rho_{34}}{2\theta}\left(1 - e^{-2\theta t}\right). \tag{38}$$

Let us approximate

$$z_i(t) \approx \exp\left(\xi_i - \frac{1}{2}\mathsf{Var}(\xi_i(t))\right) \quad i = 1, 2,$$

where the Gaussian vector (ξ_1, ξ_2) has zero mean and the variance-covariance matrix such that the second-order moments of $(z_1(t), z_2(t))$ given by (37), (38) are recovered. the latter can be achieved by inverting the moments relationships

$$\begin{aligned}\mathsf{Var}(z_i(t)) &= \exp(\mathsf{Var}(\xi_i)) - 1, \quad i = 1, 2, \\ \mathsf{covar}(z_1(t), z_2(t)) &= \exp(\mathsf{covar}(\xi_1, \xi_2)) - 1.\end{aligned}$$

Then,

$$\begin{aligned}\mathsf{E}\sqrt{z_1(t) z_2(t)} &= \exp\left(-\frac{1}{8}\mathsf{Var}(\xi_1 - \xi_2)\right) \\ \mathsf{E}\sqrt{z_1(t)/z_2(t)} &= \exp\left(-\frac{1}{8}(\mathsf{Var}(\xi_1) + 2\mathsf{covar}(\xi_1, \xi_2) - 3\mathsf{Var}(\xi_2))\right),\end{aligned}$$

and so on.

6.3. **Example 3: Fitting the market with a stochastic volatility model.** While the focus of this paper is on approximating more complicated models with simpler ones for calibration purposes, direct calibration of stochastic volatility models to the market is another possible application of the techniques we use.

Let $S(t)$ be the value of a given market variable (for example, an equity). Let S_0 be the spot value. Suppose market-implied Black volatilities $\sigma_{\text{mkt}}(T, K)$ are known for all expiries T and strikes K. These can be easily converted (see (5)) into the market-implied Dupire's local volatility $b_{\text{mkt}}(t, x)$, ie a local volatility function such that the market European option prices are consistent with the model

$$dY(t) = b_{\text{mkt}}(t, Y(t))\, dW(t), \quad Y(0) = S_0.$$

Suppose we would like to construct a stochastic volatility model

$$dS(t) = b(t, S(t))\sqrt{z(t)}\, dW(t), \quad S(0) = S_0, \tag{39}$$

consistent with market European option prices. As follows from Theorem 2.1, the local volatility function $b(t, x)$ is then given by

$$b^2(t, x) = \frac{b^2_{\text{mkt}}(t, x)}{\mathsf{E}(z(t) | S(t) = x)}. \tag{40}$$

As mentioned before, the challenge of computing $\mathsf{E}(z(t)|S(t) = x)$ makes the method difficult to implement in practice. Numerical methods are often applied to compute the conditional expected value in a forward Kolmogorov PDE for $(S(t), z(t))$. Instead, we would like to discuss a number of approaches to obtain (analytic) approximations.

[2] Usiang Gaussian approximation to the dynamics of z_1, z_2

A direct Gaussian-approximation-based approach (see Section 4) is possible, but is likely to be inaccurate in this case if $b_{\mathrm{mkt}}(t,x)$ has a high degree of convexity. What we would like to explore instead are methods based on comparing (39) with another *stochastic volatility* model, a model in which European option prices can be cheaply computed[3]. Suppose we have such a "proxy" model, as defined by the (known) local volatility function $\tilde{b}(t,x)$, for the underlying $X(t)$ and (the same) stochastic variance process $z(t)$,

$$dX(t) = \tilde{b}(t,X(t))\sqrt{z(t)}\,dW(t), \quad X(0) = S_0. \tag{41}$$

Typically $\tilde{b}(t,x)$ is a simple function, most common being $\tilde{b}(t,x) = \beta$ ("normal" case), $\tilde{b}(t,x) = \nu x$ ("log-normal" case), or $\tilde{b}(t,x) = \nu x + \beta$ ("shifted log-normal" case). Let $\sigma_{\mathrm{proxy}}(T,K)$ be the implied Black volatilities of European options from this model. By construction, $\sigma_{\mathrm{proxy}}(T,K)$ are cheap to compute. Then, they can be turned into the "proxy" Dupire's local volatility $b_{\mathrm{proxy}}(t,x)$ (by the means of (5)). Then, rewriting (40),

$$\mathsf{E}(z(t)|X(t)=x) = \frac{b^2_{\mathrm{proxy}}(t,x)}{\tilde{b}^2(t,x)}. \tag{42}$$

Hence, having a stochastic volatility model which cheaply-computable European option prices allows us to compute the conditional expected values $\mathsf{E}(z(t)|X(t)=x)$ easily. One way to take advantage of this observation is to combine (40) and (42),

Proposition 6.4. *The local volatility function $b(t,x)$ that makes the model (39) consistent with the market is given by*

$$b(t,x) = \tilde{b}(t,x) \times \frac{b_{\mathrm{mkt}}(t,x)}{b_{\mathrm{proxy}}(t,x)} \times \left(\frac{\mathsf{E}(z(t)|X(t)=x)}{\mathsf{E}(z(t)|S(t)=x)}\right)^{/12},$$

where $X(t)$ follows the "proxy" model (41) with a known local volatility function $\tilde{b}(t,x)$.

Remark 6.4. *The local volatility function $b(t,x)$ that we are trying to find is given by the "proxy" local volatility $\tilde{b}(t,x)$ times two corrections, the ratio of Dupire's local volatilities $\frac{b_{\mathrm{mkt}}(t,x)}{b_{\mathrm{proxy}}(t,x)}$, and the ratio of conditional expected values $\frac{\mathsf{E}(z(t)|X(t)=x)}{\mathsf{E}(z(t)|S(t)=x)}$.*

The ratio $\frac{b_{\mathrm{mkt}}(t,x)}{b_{\mathrm{proxy}}(t,x)}$ is, as discussed, cheap to compute. Approximating

$$\frac{\mathsf{E}(z(t)|X(t)=x)}{\mathsf{E}(z(t)|S(t)=x)} \approx 1,$$

we obtain the following useful corollary.

Corollary 6.1. *Approximately,*

$$b(t,x) \approx \tilde{b}(t,x) \times \frac{b_{\mathrm{mkt}}(t,x)}{b_{\mathrm{proxy}}(t,x)}.$$

[3]This approach has originally been suggested in Forde (2006).

To obtain a more sophisticated approximation, we can look for an (approximate) functional relationship between $X(t)$ and $S(t)$. Denote

$$h(t,x) = \int_{x_0}^{x} \frac{dy}{b(t,y)}, \tag{43}$$
$$\tilde{h}(t,x) = \int_{x_0}^{x} \frac{dy}{\tilde{b}(t,y)},$$
$$H(t,x) = \tilde{h}^{-1}(t,h(t,x)),$$

where $\tilde{h}^{-1}(t,x)$ is the inverse of $\tilde{h}(t,x)$ in the second (ie, x) argument. Furthermore, denote

$$\tilde{X}(t) = H(t,S(t)).$$

Then

$$\begin{aligned} d\tilde{X}(t) &= \frac{\partial}{\partial x} H(t,x)\bigg|_{x=S(t)} dS(t) + \ldots dt \\ &= \tilde{b}\left(t,\tilde{X}(t)\right)\sqrt{z(t)}\,dW(t) + \ldots dt. \end{aligned}$$

We see that $\tilde{X}(t)$ and $X(t)$ have the same diffusion coefficient. This inspires an approximation

$$X(t) \approx H(t,S(t)), \tag{44}$$

that leads to the following result.

Proposition 6.5. *The local volatility function $b(t,x)$ that makes the model (39) approximately consistent with the market is given by*

$$b(t,x) = \tilde{b}(t,H(t,x))\frac{b_{\text{mkt}}(t,x)}{b_{\text{proxy}}(t,H(t,x))}, \tag{45}$$

with $H(\cdot,\cdot)$ given by (43).

Proof. By (44),

$$\begin{aligned} \mathsf{E}(z(t)|S(t)=x) &= \mathsf{E}(z(t)|H(t,S(t))=H(t,x)) \\ &\approx \mathsf{E}(z(t)|X(t)=H(t,x)). \end{aligned}$$

From (42),

$$\mathsf{E}(z(t)|S(t)=x) = \frac{b^2_{\text{proxy}}(t,H(t,x))}{\tilde{b}^2(t,H(t,x))},$$

and the result follows from (40).∎

Remark 6.5. *Note that $H(\cdot,\cdot)$ depends on the (unknown) function $b(t,x)$, hence (45) should be treated as a functional equation for $b(t,x)$.*

In the small-time limit, the relationship between a local and implied volatilities are much simpler than implied by (5). The following corollary is well-known (see for example Andersen and Brotherton-Ratcliffe (2001)).

Corollary 6.2. *Let $X(\cdot)$ follow the time-independent local volatility model*

$$dX(t) = \varphi(X(t))\,dW(t), \quad X(0) = X_0.$$

Then the implied Black volatility is given to the first order in t by

$$\sigma(K) = \log(K/X_0)\left(\int_{X_0}^{K} \frac{du}{\varphi(u)}\right)^{-1}.$$

By using this simple (harmonic average) relationship between the two, Propositions 6.4 and 6.5 can be restated in terms of the implied, rather than local, volatilities. We denote $b(x) \triangleq \lim_{t\to 0+} b(t,x)$, etc. Proposition 6.5 in the short-time limit was first proved in Forde (2006),

Corollary 6.3. *In the short-time limit,*

$$b(x) = \tilde{b}(H(x))\left(\frac{\log(x/S_0)}{\sigma_{\mathrm{proxy}}(H(x))}\right)' \Bigg/ \left(\frac{\log(x/S_0)}{\sigma_{\mathrm{mkt}}(x)}\right)'.$$

One question remains unanswered, and that is the choice of the "proxy" model. If $z(\cdot)$ follows the standard mean-reverting square root process (as in the Heston model), then the following models admit closed-form expressions for European options,

$$\tilde{b}(t,x) = \nu x$$

(the original Heston model), and

$$\tilde{b}(t,x) = \nu x + \beta,$$

the shifted Heston model, as previously explained. Moreover, using the parameter averaging techniques, this could be extended to a shifted-lognormal Heston model with time-dependent coefficients,

$$\tilde{b}(t,x) = \nu(t)\,x + \beta(t).$$

The last model, is, probably, the best choice for the proxy model. The time-dependent coefficients can be chosen to make the "proxy" model as similar as possible to the model for $X(\cdot)$, thus improving the quality of various approximations made. Ideally we should use

$$\beta(t) = b(t,S_0),$$
$$\nu(t) = \frac{\partial}{\partial x} b(t,S_0),$$

the first-order approximation to the local volatility $b(t,x)$ "along the forward". The value and the derivative of the local volatility $b(t,x)$ are of course unknown a-priori; one can envision various approximations or, conceivably, an iterative procedure where the local volatility approximation obtained on step n is used to define the proxy local volatility for step $n+1$.

7. Conclusions

We have developed a generic method for obtaining closed-form approximations to European options and demonstrated its usage in a number of simple examples. The accuracy of the approximations critically depends on how accurately the conditional expected values can be estimated. In particular, any improvements over Gaussian-based methods will have an immediate effect on the European option approximations.

References

Andersen, L. B. and Andreasen, J. (2002). Volatile volatilities. *Risk*, 15(12): 163–168.

Andersen, L. B. and Brotherton-Ratcliffe, R. (2001). Extended Libor market models with stochastic volatility. Working paper.

Andreasen, J. (2005). Back to the future. *Risk Magazine*, 18(9): 104–109.

Avellaneda, M., Boyer-Olson, D., Busca, J., and Friz, P. (2002). Reconstructing volatility. *Risk*, 15(10).

Dupire, B. (1994). Pricing with a smile. *Risk*, 7(1).

Dupire, B. (1997). A unified theory of volatility. Banque Paribas working paper.

Forde, M. (2006). Calibrating local-stochastic volatility models. University of Bristol working paper, http://www.maths.bris.ac.uk/ mamsf/.

Gyöngy, I. (1986). Mimicking the one-dimensional distributions of processes having an Ito differential. *Prob. Th. Rel. Fields*, 71: 501 – 516.

Karatzas, I. and Shreve, S. E. (1997). *Brownian Motion and Stochastic Calculus*. Springer.

Piterbarg, V. V. (2004). A stochastic volatility forward Libor model with a term structure of volatility smiles. SSRN Working paper.

Piterbarg, V. V. (2005a). A multi-currency model with FX volatility skew. SSRN working paper.

Piterbarg, V. V. (2005b). Stochastic volatility model with time-dependent skew. *Applied Mathematical Finance*, 12(2): 147–185.

Piterbarg, V. V. (2005c). Time to smile. *Risk Magazine*, 18(5): 71–75.

Piterbarg, V. V. (2006a). Smiling hybrids. *Risk Magazine*, 19(5): 66–71.

Piterbarg, V. V. (2006b). Towards a multi-stochastic volatility model for CMS spread exotics. Global Derivatives Trading and Risk Management Conference, Paris.

Savine, A. (2000). A theory of volatility. Reech working paper, http://www.reech.com/company/research.papers/.

Barclays Capital, 5 The North Colonnade, E14 4BB, London
E-mail address: `vladimir.piterbarg@barclayscapital.com`