Markovian Projection onto a Displaced Diffusion: Generic Formulas with Applications

A. Antonov* and T. Misirpashaev†

October 9, 2006

Abstract

We develop a systematic approach to Markovian projection onto an effective displaced diffusion, and work out a set of computationally efficient formulas valid for a large class of non-Markovian underlying processes. The generic derivation is followed by applications, including the calculation of FX options in cross-currency models and swaption pricing in LIBOR Market Models, where we are able to recover in an unambiguous way many known analytical approximations and derive several new ones.

1 Introduction

'Markovian projection' is a term introduced to the vocabulary of quantitative finance by Piterbarg (2006a and references therein) in a series of works solving various appearances of the problem of efficient analytical approximation to options with European exercise. The term refers to a technique that is based on a theorem, proven by Gyöngy (1986) and used in a narrower context by Dupire (1997), which explains how a complicated non-Markovian process can be replaced by a Markovian process with the same marginal distributions. Unfortunately, the equivalent Markovian process is usually still too complicated to enable analytical progress and needs to be approximated by a displaced diffusion which is a linear function of state, possibly with time-dependent coefficients. The step of reduction to a displaced diffusion is an approximation that often invokes heuristic considerations and has not been developed into an unambiguous technical procedure. The purpose of this paper is to state an analytical criterion for an optimal displaced diffusion and derive a formulaic solution that can be uniformly applied to a large class of problems.

*NumeriX Software Ltd. 41 Eastcheap Street, 2nd Floor London EC3M 1DT UK; antonov@numerix.com
†NumeriX LLC 4320 Winfield Rd, Suite 200, Warrenville, IL 60555 USA; missir@numerix.com

Generic result. We begin with a statement of the key result. Our task is to develop an efficient approximation for European options, $E[(S(T)-K)^+]$. It is usually possible to derive an SDE for the underlying process $S(t)$ in its equivalent martingale measure,

$$dS(t) = S_0\Lambda(t)\cdot dZ(t), \qquad S(0) = S_0, \tag{1}$$

where $Z(t)$ is the standard F-dimensional Brownian motion and $\Lambda(t)$ is an F-dimensional stochastic process adapted to a suitable augmentation of the filtration generated by $Z(t)$. The notation $f\cdot g$ is used for a scalar product in the space of Brownian motion factors. The process $\Lambda(t)$ has a meaning of normalized volatility and is often available in a closed form. This form however can be very complicated, generally resulting in a non-Markovian process $S(t)$.

We are looking for the parameter functions $\sigma(t)$ and $\beta(t)$ of an effective displaced diffusion process $S^*(t)$ that obeys the SDE

$$dS^*(t) = (S^*(t)\beta(t) + (1-\beta(t))S^*(0))\,\sigma(t)\cdot dZ(t), \qquad S^*(0) = S_0, \tag{2}$$

and optimally approximates the marginals of $S(t)$. Assuming that an expansion in powers of volatility is possible, we get the following leading behavior[1],

$$|\sigma(t)|^2 = E\left[|\Lambda(t)|^2\right], \tag{3}$$

$$\beta(t) = \frac{S_0 E\left[|\Lambda(t)|^2 (S(t)-S_0)\right]}{2E\left[|\Lambda(t)|^2\right] E\left[(S(t)-S_0)^2\right]}. \tag{4}$$

Having found the effective displaced diffusion, we can invoke Piterbarg's (2005a) formula for skew averaging

$$\bar\beta_T = \frac{\int_0^T \beta(t)|\sigma(t)|^2 \int_0^t |\sigma(\tau)|^2 d\tau dt}{\int_0^T |\sigma(t)|^2 \int_0^t |\sigma(\tau)|^2 d\tau dt}, \tag{5}$$

and finally reduce the pricing of a European option with maturity T to a variant of the Black-Scholes formula,

$$E[(S(T)-K)^+] = \frac{S_0}{\bar\beta_T}\mathcal{N}(d_+) - \left(K + \frac{S_0(1-\bar\beta_T)}{\bar\beta_T}\right)\mathcal{N}(d_-), \tag{6}$$

$$d_\pm = \frac{\ln\left(S_0/(K\bar\beta_T + S_0(1-\bar\beta_T)\right) \pm V/2}{\sqrt{V}}, \quad V = \bar\beta_T^2 \int_0^T |\sigma(t)|^2 dt. \tag{7}$$

[1] This is deduced from a more general result valid to all orders in volatility, which however is difficult to use in practice since the calculation of expected values occurring in it gets very cumbersome beyond the leading order.

Separable volatility process. To give an example where all calculations can be completed in a closed form, assume that we succeeded in casting the martingale measure SDE for the underlying process $S(t)$ in the form

$$dS = S_0 \sum_n X_n(t) a_n(t) \cdot dZ, \tag{8}$$

where $a_n(t)$ are deterministic vector functions with scalar components $a_{n,\nu}(t)$ ($\nu = 1, \ldots, F$), and $X_n(t)$ are stochastic processes obeying SDEs of the form

$$dX_n = \mu_n(t, \{X_k(t)\})dt + \sigma_n(t, \{X_k(t)\}) \cdot dZ(t). \tag{9}$$

Each local volatility function $\sigma_n(t, \{X_k\})$ here has F scalar components. The volatility components $\sigma_{n,\nu}(t, \{X_k\})$ and the drifts can, in general, depend on $X_n(t)$, and on the values achieved by the processes $X_k(t)$ with $k \neq n$. Therefore, the processes $X_n(t)$ considered in isolation are not necessarily Markovian, but taken together constitute a multi-dimensional (possibly infinitely dimensional) Markov process. Under these conditions, we call the process $S(t)$ a separable volatility process.

We assume furthermore that the drift terms μ_n are small in the sense that they are of the second or higher order in volatilities[2], $\mu_n = O(\sigma_+^2)$, where $\sigma_+ = \max\{\sigma_{n,\nu}(t, \{X_k(t)\}), a_{n,\nu}(t)\}$. We will show that, for a separable volatility process with small drifts (in the sense just described), the expectations in Eqs. (3) and (4) can be computed in the leading order, resulting in the following explicit form suitable for immediate applications:

$$\sigma(t) = \sum_n X_n(0) a_n(t) + O(\sigma_+^3), \tag{10}$$

$$\beta(t) = \frac{\sum_n (a_n(t) \cdot \sigma(t)) \int_0^t (\sigma_n(\tau, \{X_k(0)\}) \cdot \sigma(\tau))\, d\tau}{|\sigma(t)|^2 \int_0^t |\sigma(\tau)|^2 d\tau} + O(\sigma_+^2). \tag{11}$$

The rest of the paper is organized as follows. A derivation of the generic Eqs. (3) and (4), and more specific Eqs. (10) and (11) in section 2 is followed by applications that show the universality and robustness of the method. As a first example, in section 3, we consider the projection onto a displaced diffusion of a generic process with local volatility.

Section 4 is devoted to applications for the pricing of FX options in cross-currency models. In section 4.1, we reproduce Piterbarg's (2005b) results for the model with Gaussian interest rate components and a CEV process for the FX rate. We move on to the case of LIBOR Market Models for interest rates in section 4.2, and obtain approximations for FX options which—to our knowledge—were not derived before.

[2] A motivating case where this restriction for drifts is satisfied is offered by LIBOR Market Models where the basis processes X_n are LIBOR rates, and the drifts picked up as a result of measure changes are quadratic in volatilities.

In section 5, we apply the method to the important problem of single-currency swaption valuation in LIBOR Market Models and obtain a new formula which captures the effect of the skew in a systematic way and offers an alternative to a more heuristic swap rate approximation of Piterbarg (2005a).[3]

2 Derivation of the key formulas

Generic process. We proceed to the proof of Eqs. (3) and (4) for the parameters of an effective displaced diffusion. The technique is based on a theorem by Gyöngy (1986) according to which the Markovian process

$$dS^{**}(t) = \Xi(t, S^{**}(t))dZ(t), \qquad S^{**}(0) = S_0, \tag{12}$$

has exactly the same marginal distributions as $S(t)$ provided the local volatility $\Xi(t,x)$ satisfies the equation

$$|\Xi(t,x)|^2 = E\left[S_0^2|\Lambda(t)|^2 \,\middle|\, S(t) = x\right]. \tag{13}$$

For every fixed t, the conditional expectation in the right hand side can be characterized as a function of state $|\Xi(\cdot,x)|^2$ which minimizes the L_2 distance from the true variance,

$$\chi^2 = E\left[\left(S_0^2|\Lambda(t)|^2 - |\Xi(t,S(t))|^2\right)^2\right] \to \min. \tag{14}$$

To obtain a tractable model, instead of minimizing the functional in Eq. (14) over the space of all local volatility functions, we minimize it only over a subspace of affine linear functions of state, expressible in the form

$$\Xi(t,S(t)) = (\beta(t)S(t) + (1-\beta(t))S_0))\,\sigma(t). \tag{15}$$

This corresponds to a minimization over $\beta(t)$ and $\sigma(t)$ of the function

$$\chi^2(\sigma(t),\beta(t)) = E\left[\left(S_0^2|\Lambda(t)|^2 - (S_0 + \beta(t)\Delta S(t))^2\,|\sigma(t)|^2\right)^2\right] \tag{16}$$

for every fixed t. Here $\Delta S(t) = S(t) - S_0$. Equating to zero the variations of $\chi^2(\sigma(t),\beta(t))$ over $\sigma(t)$ and $\beta(t)$ gives a pair of equations,

$$S_0^2 E\left[(|\Lambda(t)|^2(S_0 + \beta(t)\Delta S(t))^2\right] = |\sigma(t)|^2 E\left[(S_0 + \beta(t)\Delta S(t))^4\right], \tag{17}$$

$$S_0^2 E\left[|\Lambda(t)|^2(S_0 + \beta(t)\Delta S(t))\Delta S(t)\right] = |\sigma(t)|^2 E\left[(S_0 + \beta(t)\Delta S(t))^3\Delta S(t)\right]. \tag{18}$$

The averages of the form

$$E_{i,j}(t) = E\left[|\Lambda(t)|^i \Delta S(t)^j\right] \tag{19}$$

[3]More recent, unpublished result of Piterbarg (2006b) agrees with ours in the case of one-factor LMM.

encountered in Eqs. (17) and (18) are unconditional and can be efficiently computed using a perturbation theory as we will show in a moment. Assuming that all necessary expectations $E_{i,j}(t)$ are found with sufficient accuracy, one could solve Eqs. (17) and (18) exactly. Dividing one of the equations by another eliminates the unknown $|\sigma(t)|^2$ and leads to an algebraic equation of 4th order for $\beta(t)$,

$$\begin{aligned}\beta^4(E_{2,2}E_{0,3} - E_{2,1}E_{0,4}) + \beta^3 S_0(3E_{0,2}E_{2,2} - 2E_{2,1}E_{0,3} - E_{2,0}E_{0,4})\\ -3\beta^2 S_0^2 E_{2,0}E_{0,3} + \beta S_0^3(E_{2,2} - 3E_{2,0}E_{0,2}) + S_0^4 E_{2,1} = 0.\end{aligned} \quad (20)$$

(We used the martingale property $E_{0,1}(t) = E[S(t) - S_0] \equiv 0$ but kept all non-vanishing terms.) Knowing $\beta(t)$, we can find $|\sigma(t)|^2$ from any of the two equations or their linear combination. One possible representation is

$$|\sigma|^2 = \frac{S_0^3 E_{2,0} + \beta S_0^2 E_{2,1}}{S_0^3 + 3\beta^2 S_0 E_{0,2} + \beta^3 E_{0,3}}. \quad (21)$$

It is more practical, however, to use an expansion in volatilities. Keeping only the leading terms constitutes a second technical approximation—the first and principal approximation being the restriction to the subspace (15) of displaced diffusions. For that we need to examine the order of the various averages $E_{i,j}(t)$.

The technique for the estimation of the leading order of the moments is developed in the second part of this section for the case of a separable volatility process; the generic case requires minor adaptations. The averages $E_{2,0}(t) = E\left[|\Lambda(t)|^2\right]$ and $E_{0,2}(t) = E\left[(S(t) - S_0)^2\right]$ have the highest quadratic order $O(|\Lambda|^2)$. Next are the averages $E_{2,1}(t)$, $E_{2,2}(t)$, $E_{0,3}(t)$, and $E_{0,4}(t)$, each of which is of the order $O(|\Lambda|^4)$. Keeping only the leading order terms, Eq. (20) reduces to a linear equation with the solution,

$$\beta(t) = \frac{S_0 E_{2,1}(t)}{3E_{2,0}(t)E_{0,2}(t) - E_{2,2}(t)}. \quad (22)$$

Noting that the difference between $E_{2,2}(t)$ and the product $E_{2,0}(t)E_{0,2}(t)$ is of an order higher than $O(|\Lambda|^4)$, we can further simplify this expression to the result (4) for the effective skew $\beta(t)$ stated in the introduction. The leading order behavior (3) for the effective volatility $\sigma(t)$ is a simple consequence of Eq. (21).

Separable volatility process. We turn to the specific case when the volatility $\Lambda(t)$ has a form consistent with Eq. (8),

$$\Lambda(t) = \sum_n a_n(t) X_n(t), \quad (23)$$

where X_n's satisfy the SDE (9),

$$dX_n = \mu_n(t, \{X_k(t)\})dt + \sigma_n(t, \{X_k(t)\}) \cdot dZ(t),$$

with the accompanying condition on the smallness of the drifts, $\mu_n(t) = O(\sigma_+^2)$, discussed in the introduction. (Note that $a_n(t) = O(\sigma_+)$, by definition of the characteristic scale σ_+.) Applying Ito's lemma to the product process $X_n(t)X_m(t)$,

$$d(X_n X_m) = (X_n \mu_m + X_m \mu_n + \sigma_n \cdot \sigma_m) dt + (X_n \sigma_m + X_m \sigma_n) \cdot dZ, \tag{24}$$

we observe that the drift of $X_n(t)X_m(t)$ is also of the order $O(\sigma_+^2)$; therefore

$$E\left[X_n(t)X_m(t)\right] = X_n(0)X_m(0) + O(\sigma_+^2). \tag{25}$$

The leading order of the average $E_{2,0}(t) = E[|\Lambda(t)|^2]$ follows,

$$E_{2,0}(t) = \sum_{n,m} (a_n(t) \cdot a_m(t)) E\left[X_n(t)X_m(t)\right] = \Big|\sum_n a_n(t) X_n(0)\Big|^2 + O(\sigma_+^4). \tag{26}$$

This leads to Eq. (10) as one possible solution for the effective volatility $\sigma(t)$. All other solutions are related by orthogonal rotations in the factor space and lead to the same final answers for option prices.

We introduce the notation

$$\bar{\Lambda}(t) = \sum_n a_n(t) X_n(0) \tag{27}$$

for the volatility $\Lambda(t)$ computed with frozen values of the processes $X_n(t)$. Then

$$E_{2,0}(t) = |\bar{\Lambda}(t)|^2 + O(\sigma_+^4). \tag{28}$$

The leading order of the other averages cannot be directly found with a replacement of the processes by their initial values because the result would vanish. A general strategy is to use Ito's lemma to take consecutive differentials of the processes appearing under the expectation sign until the replacement by the initial values becomes possible. For example, for $E_{0,2}(t) = E[(S(t) - S_0)^2]$ we have

$$d(\Delta S(t)^2) = 2 S_0 \Delta S(t) \Lambda(t) \cdot dZ(t) + S_0^2 |\Lambda(t)|^2 dt. \tag{29}$$

Taking the expectations, we obtain an ODE

$$dE[\Delta S(t)^2] = S_0^2 E[|\Lambda(t)|^2] dt, \tag{30}$$

which can be expressed as

$$dE_{0,2}(t) = S_0^2 E_{2,0}(t) dt, \tag{31}$$

and easily solved in the leading order,

$$E[\Delta S(t)^2] = S_0^2 \int_0^t |\bar{\Lambda}(\tau)|^2 d\tau + O(\sigma_+^4). \tag{32}$$

Similarly, applying Ito's lemma to $\Delta S(t)^k$, we get a generalization of Eq. (31),

$$dE_{0,k}(t) = \frac{1}{2}k(k-1)S_0^2 E_{2,k-2}(t)dt, \tag{33}$$

which shows that $E_{0,3}(t)$ is of the same order as $E_{2,1}(t)$ while $E_{0,4}(t)$ is of the same order as $E_{2,2}(t)$.

We now turn to the calculation of the leading behavior of the average $E_{2,1}(t)$,

$$E_{2,1}(t) = E[|\Lambda(s)|^2 \Delta S(t)] = \sum_{n,m} (a_n(t) \cdot a_m(t)) E\left[\Delta S(t) X_n(t) X_m(t)\right]. \tag{34}$$

We want to compute the average $E[\Delta S(t) X_n(t) X_m(t)]$ with the accuracy up to quadratic terms. Applying Ito's lemma to the product $\Delta S(t) X_n(t) X_m(t)$, we get

$$d(\Delta S\, X_n X_m) = S_0 X_n X_m \Lambda \cdot dZ + \Delta S\, d(X_n X_m) + \langle d\Delta S, d(X_n X_m)\rangle. \tag{35}$$

Next, we substitute Eq. (24) and take the average to obtain

$$dE\left[\Delta S X_n X_m\right] = E\left[\Delta S(X_n \mu_m + X_m \mu_n + \sigma_n \cdot \sigma_m)\right] dt + S_0 E\left[(\Lambda(t) \cdot (X_n \sigma_m + X_m \sigma_n)\right] dt. \tag{36}$$

The first term in the right hand side gives a contribution of the order $O(\sigma_+^4)$ and the second gives the leading quadratic contribution which can be computed with frozen processes $X_n(t) \to X_n(0)$, so that

$$dE\left[\Delta S(t) X_n(t) X_m(t)\right] = S_0 \bar{\Lambda}(t) \cdot \left(X_n(0) \sigma_m(t, \{X_k(0)\}) + X_m(0) \sigma_n(t, \{X_k(0)\})\right) dt + O(\sigma_+^4). \tag{37}$$

Solving this ODE and substituting into Eq. (34), we finally get

$$E_{2,1}(t) = 2S_0 \sum_n (\bar{\Lambda}(t) \cdot a_n(t)) \int_0^t (\bar{\Lambda}(\tau) \cdot \sigma_n(\tau, \{X_k(0)\})) d\tau + O(\sigma_+^6). \tag{38}$$

Similar manipulations can be used to prove that

$$E_{2,2}(t) = E_{2,0}(t) E_{0,2}(t) + O(\sigma_+^6) = |\bar{\Lambda}(t)|^2 \int_0^t |\bar{\Lambda}(\tau)|^2 d\tau + O(\sigma_+^6). \tag{39}$$

Eqs. (38), (39), and (22) lead to Eq. (11), which completes the derivation of the effective displaced diffusion for the separable volatility process.

3 Projection for a Markovian process

As a warm-up before proceeding to non-trivial applications, it is instructive to see how the generic formulas for Markovian projection onto a displaced diffusion work in the case of a Markovian underlying process,

$$dS(t) = S_0 \Lambda(t, S(t)) \cdot dZ(t). \tag{40}$$

Here, the local volatility function is apparently not in a separable form (23), but it can be represented in such a form by expanding around S_0,

$$\Lambda(t, S(t)) = \sum_{n=0}^{\infty} a_n(t) X_n(t), \tag{41}$$

with

$$a_n(t) = \frac{1}{n!} \frac{\partial^n \Lambda(t, S)}{\partial S^n}\Big|_{S=S_0}, \qquad X_n(t) = (S(t) - S_0)^n. \tag{42}$$

By Ito's lemma,

$$dX_n(t) = nS_0 X_{n-1}(t)\Lambda(t, S(t))\cdot dZ(t) + \frac{1}{2} n(n-1) S_0^2 X_{n-2}(t) |\Lambda(t, S(t))|^2 dt. \tag{43}$$

We identify the local volatilities,

$$\sigma_n(t, \{X_k\}) = nX_{n-1}(t)\Lambda(t, S_0 + X_1(t)), \tag{44}$$

and observe that the drift term is quadratic in the volatilities. Therefore, the conditions for the validity of the formulas (10) and (11) are satisfied. With the initial values $X_0(0) = 1$, $X_n(0) = 0$ for $n \neq 0$, all frozen volatilities $\sigma_n(t, \{X_k(0)\})$ with $n \neq 1$ vanish, and

$$\sigma_1(t, \{X_k(0)\}) = \Lambda(t, S_0). \tag{45}$$

The parameters of the effective displaced diffusion are

$$\sigma(t) = a_0(t) X_0(0) = \Lambda(t, S_0), \tag{46}$$

$$\beta(t) = \frac{(a_1(t)\cdot\Lambda(t, S_0)) \int_0^t |\Lambda(\tau, S_0)|^2 d\tau}{|\Lambda(t, S_0)|^2 \int_0^t |\Lambda(\tau, S_0)|^2 d\tau} = \frac{\left(\frac{\partial \Lambda(t,S)}{\partial S}\Big|_{S=S_0} \cdot \Lambda(t, S_0) \right)}{|\Lambda(t, S_0)|^2}, \tag{47}$$

which is the result we should expect from the heuristic considerations of linearized local volatility, as argued by Piterbarg (2005a). According to Gyöngy's theorem, an equivalent local volatility (in the sense of one-dimensional marginals) always exists, but finding it is a difficult computational task. Our results (3), (4), (10), and (11) generalize the notion of linearized local volatility and show how to obtain its position and slope starting directly from stochastic volatility.

4 FX option in cross-currency models

In this section, we apply the techniques developed in section 2 to the problem of efficient analytical approximations for European FX options in cross-currency models. We first consider a setup with Gaussian interest rate single-currency models (such as Hull-White models or 1-factor HJM), and then a setup with LIBOR Market Models for single-currency interest rates. In both cases, a skew process (which can be either CEV or displaced diffusion) is assumed for the spot FX rate $X(t)$. We work in T-forward measure, having as a numeraire the domestic bond $P(t,T)$. The foreign bond is denoted $\tilde{P}(t,T)$. (We systematically use the tilde for foreign-currency quantities.) In T-forward measure, the forward FX rate process

$$F(t,T) = \frac{X(t)\tilde{P}(t,T)}{P(t,T)} \tag{48}$$

is a martingale, and the price reduces to a discounted expectation,

$$P(0,T)E_T\left[(F(T,T)-K)^+\right]. \tag{49}$$

4.1 Gaussian interest rate models with FX skew

Piterbarg (2005b) sets up the following cross-currency model with FX skew in the domestic risk-neutral measure,

$$dP(t,T)/P(t,T) = r(t)dt - \sigma(t,T)\cdot dZ(t) \tag{50}$$

$$d\tilde{P}(t,T)/\tilde{P}(t,T) = (\tilde{r}(t) + \tilde{\sigma}(t,T)\cdot\gamma(t,X(t))dt - \tilde{\sigma}(t,T)\cdot dZ(t) \tag{51}$$

$$dX(t)/X(t) = (r(t)-\tilde{r}(t))dt + \gamma(t,X(t))\cdot dZ(t). \tag{52}$$

All the processes here are driven by the same standard F-dimensional Brownian motion, and we continue to assume that the volatilities are also F-dimensional. Thus, in our notations, rate correlations do not appear explicitly, but are contained in the scalar products of volatilities, such as $\tilde{\sigma}(t,T)\cdot\gamma(t,X(t))$. The volatilities $\sigma(t,T)$ and $\tilde{\sigma}(t,T)$ are deterministic[4], and for the $\gamma(t,X)$ the following form is assumed,

$$\gamma(t,x) = \nu(t)\left(\frac{x}{L(t)}\right)^{\alpha(t)-1}, \tag{53}$$

where $\nu(t)$, $L(t)$ and $\alpha(t)$ are deterministic functions. The function $\nu(t)$ has F components.

The SDE for the forward FX rate $F(t,T)$ in T-forward measure reads

$$dF(t,T) = F(t,T)\left(\sigma(t,T) - \tilde{\sigma}(t,T) + \gamma(t,X(t))\right)\cdot dZ(t). \tag{54}$$

To apply the formulas (10) and (11), we need to express the volatility as a linear combination of processes whose drift is of second order or higher in volatilities. The FX rate process $X(t)$ itself

[4]We included a negative sign in the definition of the volatilities to make the final answers consistent with those of section 4.2 where the model is defined in terms of LIBOR rates as opposed to bond prices.

is not a good candidate to be a part of such an expansion because it has a large drift expected value which is related to the forward curve $F(0,t)$. We isolate this dominant drift in the FX rate by introducing a process

$$R(t) = X(t)/F(0,t). \tag{55}$$

The process $R(t)$ satisfies the following SDE,

$$dR(t) = \mu_R(t)dt + R(t)\gamma(t, R(t)F(0,t))\cdot dZ(t), \qquad R(0) = 1, \tag{56}$$

where $\mu_R(t)$ is quadratic in volatilities. We also normalize the forward rate process by defining

$$S(t) = F(t,T)/F(0,T). \tag{57}$$

Using the homogeneity of $\gamma(t,x)$ in x, we can write the SDE for $S(t)$ in the following form,

$$dS(t) = \Big(S(t)(\sigma(t,T) - \tilde{\sigma}(t,T)) + S(t)R(t)^{\alpha(t)-1}\gamma(t, F(0,t))\Big)\cdot dZ(t). \tag{58}$$

We now use an expansion

$$R(t)^{\alpha(t)-1} = e^{(\alpha(t)-1)\ln R(t)} = \sum_{n=0}^{\infty} \frac{(\ln R(t))^n(\alpha(t)-1)^n}{n!}, \tag{59}$$

and finally arrive at a desired representation of the volatility of the process $S(t)$ in the form $\sum_{n=0}^{\infty} a_n(t)X_n(t)$ with

$$X_n(t) = S(t)(\ln R(t))^n, \tag{60}$$

$$a_0(t) = \sigma(t,T) - \tilde{\sigma}(t,T) + \gamma(t, F(0,t)), \tag{61}$$

$$a_n(t) = \frac{(\alpha(t)-1)^n}{n!}\gamma(t, F(0,t)), \quad n \geq 1. \tag{62}$$

Similar to the example of the previous section, in a straightforward application of Ito's lemma, almost all frozen volatilities $\sigma_n(t, \{X_k(0\})$ vanish,

$$\sigma_0(t, \{X_k(0\}) = \sigma(t,T) - \tilde{\sigma}(t,T) + \gamma(t, F(0,T)), \tag{63}$$

$$\sigma_1(t, \{X_k(0\}) = \gamma(t, F(0,T)), \tag{64}$$

$$\bar{\sigma}_n(t, \{X_k(0\}) = 0, \quad n \geq 2. \tag{65}$$

We are now in a position to use formulas (10) and (11) for the volatility and skew of the effective displaced diffusion,

$$\sigma(t) = \sigma(t,T) - \tilde{\sigma}(t,T) + \gamma(t, F(0,t)), \tag{66}$$

$$\beta(t) = 1 - \frac{(1-\alpha(t))(\gamma(t, F_0(t))\cdot\sigma(t)) \int_0^t (\gamma(\tau, F(0,\tau))\cdot\sigma(\tau))d\tau}{|\sigma(t)|^2 \int_0^t |\sigma(\tau)|^2 d\tau}. \tag{67}$$

This agrees with the result of Piterbarg (2005b).

4.2 Cross-currency LMM model

Following a standard setup of LIBOR Market Models, we assume a common set of maturities $0 = T_0 < T_1 < \ldots < T_N$ for both single-currency rate LMM components and define domestic and foreign LIBOR rates, $L_n(t)$ and $\tilde{L}_n(t)$, as forward rates starting at T_n and ending at T_{n+1}. In terms of T-maturity zero-coupon bonds, $P(t,T)$ and $\tilde{P}(t,T)$, the LIBOR rates are given by

$$L_n(t) = \frac{1}{\delta_n}\left(\frac{P(t,T_n)}{P(t,T_{n+1})} - 1\right), \qquad \tilde{L}_n(t) = \frac{1}{\delta_n}\left(\frac{\tilde{P}(t,T_n)}{\tilde{P}(t,T_{n+1})} - 1\right), \tag{68}$$

where δ_n is the daycount fraction from T_n to T_{n+1}.

An important part in the analytics will be played by bond ratios,

$$\begin{aligned} R_n(t) &= \frac{P(t,T_n)}{P(t,T_{n+1})} = 1 + \delta_n\, L_n(t), \\ \tilde{R}_n(t) &= \frac{\tilde{P}(t,T_n)}{\tilde{P}(t,T_{n+1})} = 1 + \delta_n\, \tilde{L}_n(t). \end{aligned} \tag{69}$$

We specify the dynamics directly in terms of the bond ratios. In the martingale measure of each LIBOR,

$$dR_n(t) = (\beta_n(t)R_n(t) + (1-\beta_n(t))R_n(0))\sigma_n(t)\cdot dZ_n(t), \tag{70}$$

$$d\tilde{R}_n(t) = (\tilde{\beta}_n(t)\tilde{R}_n(t) + (1-\tilde{\beta}_n(t))\tilde{R}_n(0))\tilde{\sigma}_n(t)\cdot d\tilde{Z}_n(t). \tag{71}$$

We take for the numeraire in the domestic currency the rolling spot account, corresponding to Jamshidian's (1997) spot-LIBOR measure,

$$N(t) = \frac{1}{P(T_0,T_1)} \cdots \frac{1}{P(T_{n-1},T_n)}\, \frac{P(t,T_{n+1})}{P(T_n,T_{n+1})} \qquad \text{for } T_n < t \le T_{n+1}, \tag{72}$$

and set the volatility of the short bond $P(t,T_{n+1})$ equal to 0. As was shown by Schlögl (2002), this numeraire is equivalent to the continuously compounded savings account in the sense that they lead to the same measure. The foreign-currency numeraire $\tilde{N}(t)$ is chosen similarly.

The FX dynamics linking the single-currency models into a cross-currency model are defined by postulating a process of the numeraire-discounted asset $Y(t)$ related to the foreign exchange rate $X(t)$ by

$$Y(t) = \frac{X(t)\,\tilde{N}(t)}{N(t)}, \tag{73}$$

with $Y(0) = X(0) = Y_0$. We set

$$dY(t) = Y(t)\gamma(t, Y(t))\cdot dZ(t) \tag{74}$$

in the spot-LIBOR measure with F-component local volatility $\gamma(t,x)$. In accordance with section 4.1, we assume the following functional form for the volatility,

$$\gamma(t,Y(t)) = \sigma_Y(t)\left(\frac{Y(t)}{Y_0}\right)^{\alpha(t)-1}, \tag{75}$$

with deterministic functions $\sigma_Y(t)$ and $\alpha(t)$—where again $\sigma_Y(t)$ has F components. We chose here a slightly different approach, specifying the dynamics for the natural martingale $Y(t)$ instead of the FX rate $X(t)$, so that we do not need to eliminate the drift.

The forward FX rate process can be expressed in terms of the bond ratios, $R_n(t)$ and $\tilde{R}_n(t)$, and the process $Y(t)$,

$$F(t,T_M) = \prod_{k=0}^{\eta(t)-1} \frac{\tilde{P}(T_{k-1},T_k)}{P(T_{k-1},T_k)}\, Y(t) \prod_{n=\eta(t)}^{M-1} R_n(t)\tilde{R}_n^{-1}(t), \tag{76}$$

where $\eta(t) = n+1$ for $T_n < t \le T_{n+1}$. The SDE for the martingale $S(t) = F(t,T)/F(0,T)$ reads

$$dS(t) = \Lambda(t)\cdot dZ(t), \qquad S(0) = 1, \tag{77}$$

with

$$\begin{aligned}\Lambda(t) &= S(t)\Big(\gamma(t,Y(t)) \\ &+ \sum_{n=\eta(t)}^{M-1} \frac{\beta_n(t)\,R_n(t) + (1-\beta_n(t))\,R_n(0)}{R_n(t)}\sigma_n(t) \\ &- \sum_{n=\eta(t)}^{M-1} \frac{\tilde{\beta}_n(t)\,\tilde{R}_n(t) + (1-\tilde{\beta}_n(t))\,\tilde{R}_n(0)}{\tilde{R}_n(t)}\tilde{\sigma}_n(t)\Big).\end{aligned} \tag{78}$$

We can deal with the pure local volatility term $S(t)\gamma(t,Y(t))$ as we did in section 4.1, representing it as a series,

$$\left(\frac{Y(t)}{Y_0}\right)^{\alpha(t)-1} = \sum_{n=0}^{\infty} \frac{(\ln Y(t)/Y_0)^n\,(\alpha(t)-1)^n}{n!}, \tag{79}$$

or, alternatively, we could invoke the result of section 3. Applying formulas (10) and (11), we finally get

$$\sigma(t) = \sum_{n=\eta(t)}^{M-1} (\sigma_n(t) - \tilde{\sigma}_n(t)) + \sigma_Y(t), \tag{80}$$

$$\begin{aligned}
\beta(t) &= 1 - \frac{(1-\alpha(t))(\sigma_Y(t)\cdot\sigma(t))\int_0^t(\sigma_Y(\tau)\cdot\sigma(\tau))\,d\tau}{|\sigma(t)|^2\int_0^t|\sigma(\tau)|^2 d\tau} \\
&- \frac{\sum_{n=\eta(t)}^{M-1}(1-\beta_n(t))(\sigma_n(t)\cdot\sigma(t))\int_0^t(\sigma_n(\tau)\cdot\sigma(\tau))d\tau}{|\sigma(t)|^2\int_0^t|\sigma(\tau)|^2 d\tau} \\
&+ \frac{\sum_{n=\eta(t)}^{M-1}(1-\tilde{\beta}_n(t))(\tilde{\sigma}_n(t)\cdot\sigma(t))\int_0^t(\tilde{\sigma}_n(\tau)\cdot\sigma(\tau))d\tau}{|\sigma(t)|^2\int_0^t|\sigma(\tau)|^2 d\tau},
\end{aligned} \tag{81}$$

where we substituted $\sigma_Y(t)$ for $\gamma(t, Y_0)$.

A particular case of this result with $\alpha(t) \equiv 1$ was derived, and numerically checked, by Antonov and Misirpashaev (2006). It is worth mentioning that the same parameters for effective displaced diffusion of the forward FX rate would be obtained if we started with an effective displaced diffusion for the process $Y(t)$,

$$dY(t) = (Y(t)\alpha(t) + Y_0(1-\alpha(t)))\sigma_Y(t)\cdot dZ(t). \tag{82}$$

5 Single-currency swaption valuation in LMM

Now we apply the method to the problem of single-currency swaption valuation in LIBOR Market Models. We want to obtain a new formula which captures the effect of the skew in a systematic way. We begin with a discussion of generic functionals of displaced diffusions.

5.1 Functionals of displaced diffusions

Let L_n be a set of processes such that

$$dL_n = \mu_n(t)dt + (\beta_n(t)L_n + (1-\beta_n(t))l_n)\sigma_n(t)\cdot dZ(t), \quad L_n(0) = l_, \tag{83}$$

where $\beta_n(t)$ are scalar deterministic functions, $\sigma_n(t)$ are F-component deterministic functions, and the drifts are of second or higher order in volatilities, $\mu_n(t) = O(\sigma_+^2)$ with $\sigma_+ = \max(\{\sigma_n(t)\})$. Let $S(L_1, \ldots, L_N)$ be a function that does not depend on time explicitly. Assume that the measure is such that $S(t)$ considered as a process is a martingale and denote $S(l_1, \ldots, l_N) = S_0$. We apply the technique of Markovian projection to find the effective displaced diffusion that optimally approximates univariate marginals of the process $S(t)$. An immediate application of this result to European swaption pricing is considered in section 5.2. (We will take L_n to be LIBOR rates and S the swap rate.)

We have a driftless SDE for the process $S(t)$ in its martingale measure,

$$dS(t) = S_0\Lambda(t)\cdot dZ(t), \tag{84}$$

with

$$\Lambda(t) = \frac{1}{S_0}\sum_n \beta_n(t)\sigma_n(t)(L_n - l_n)\frac{\partial S}{\partial L_n} + \sigma_n(t) l_n \frac{\partial S}{\partial L_n}. \tag{85}$$

In the right hand side, we recognize the desired sum of the form $\sum_n a_n(t)X_n(t)$ with two types of terms, which we distinguish using odd and even indexes,

$$a_{2n+1}(t) = \beta_n(t)\sigma_n(t), \quad X_{2n+1}(t) = \frac{L_n - l_n}{S_0}\frac{\partial S}{\partial L_n}, \tag{86}$$

$$a_{2n}(t) = \sigma_n(t), \quad X_{2n}(t) = \frac{l_n}{S_0}\frac{\partial S}{\partial L_n}. \tag{87}$$

Diffusion terms frozen at the starting values of the processes $X_n(t)$ are easily found,

$$dX_{2n+1}(t) = \frac{1}{S_0}\frac{\partial S_0}{\partial l_n} l_n \sigma_n(t)\cdot dZ(t) + O(\sigma_+^2), \tag{88}$$

$$dX_{2n}(t) = \frac{l_n}{S_0}\sum_m \frac{\partial^2 S_0}{\partial l_n \partial l_m} l_m \sigma_m(t)\cdot dZ(t) + O(\sigma_+^2). \tag{89}$$

We are ready to apply formulas (10) and (11). The result can be represented in the form

$$\sigma(t) = \sum_n \frac{l_n}{S_0}\frac{\partial S_0}{\partial l_n}\sigma_n(t) = \sum_n \frac{\partial(\ln S_0)}{\partial(\ln l_n)}\sigma_n(t), \tag{90}$$

$$\beta(t) = \frac{\sum_n \left(\frac{1}{2}\frac{\partial|\sigma(t)|^2}{\partial(\ln l_n)} + \frac{\partial(\ln S_0)}{\partial(\ln l_n)}\left(|\sigma(t)|^2 - (1-\beta_n(t))(\sigma(t)\cdot\sigma_n(t))\right)\right)\int_0^t (\sigma_n(\tau)\cdot\sigma(\tau))d\tau}{|\sigma(t)|^2\int_0^t |\sigma(\tau)|^2 d\tau}. \tag{91}$$

We also write down the final result for the maturity-dependent effective skew β_T averaged using Piterbarg's formula (5) for the case of time-independent shifts $\beta_n(t) \equiv \beta_n$ in the processes $L_n(t)$,

$$\bar{\beta}_T = 1 + \frac{\sum_{n,m}\frac{\partial^2(\ln S_0)}{\partial(\ln l_n)\partial(\ln l_m)} I_n I_m - \sum_n \frac{\partial(\ln S_0)}{\partial(\ln l_n)}(1-\beta_n) I_n^2}{V^2}, \tag{92}$$

where

$$I_n = \int_0^T (\sigma_n(\tau)\cdot\sigma(\tau))d\tau, \quad V = \int_0^T |\sigma(\tau)|^2 d\tau. \tag{93}$$

We now examine a few simple special cases.

Log-normal processes. The case of log-normal processes L_n corresponds to $\beta_n \equiv 1$. The formula for the effective skew simplifies to

$$\beta(t) = \frac{\sum_n \left(\frac{1}{2}\frac{\partial|\sigma(t)|^2}{\partial \ln l_n} + \frac{\partial \ln S_0}{\partial \ln l_n}|\sigma(t)|^2\right)\int_0^t (\sigma(\tau)\cdot\sigma_n(\tau))d\tau}{|\sigma(t)|^2\int_0^t |\sigma(\tau)|^2 d\tau}. \tag{94}$$

Product processes. The effective skew for the product $S = \prod_n L_n$ of processes which are not necessarily log-normal is given by

$$\beta(t) = 1 - \frac{\sum_n (1-\beta_n(t))(\sigma(t)\cdot\sigma_n(t)) \int_0^t (\sigma(\tau)\cdot\sigma_n(\tau))d\tau}{|\sigma(t)|^2 \int_0^t |\sigma(\tau)|^2 d\tau}. \tag{95}$$

The averaged skew in the case of time-independent shift $\beta_n(t) \equiv \beta_n$ is given by the formula

$$\beta_T = 1 - \frac{\sum_n (1-\beta_n) I_n^2}{V^2}. \tag{96}$$

Linear combination of processes. In the case of a linear combination process $S = \sum_n \alpha_n L_n$, the result simplifies to

$$\beta(t) = \frac{\sum_n \alpha_n \beta_n(t) l_n (\sigma(t)\cdot\sigma_n(t)) \int_0^t (\sigma(\tau)\cdot\sigma_n(\tau))d\tau}{|\sigma(t)|^2 \int_0^t |\sigma(\tau)|^2 d\tau}. \tag{97}$$

If all the processes L_n are normal so that all shifts $\beta_n \equiv 0$, the linear combination is obviously also a normal process.

5.2 Swap rate projected onto a displaced diffusion

We use the LMM setup of section 4.2, restricting our attention to the domestic interest rate model. Swap rate $S(t)$ with first fixing at T_B and last payment at T_E can be written as

$$S(t) = \frac{P(t,T_B) - P(t,T_E)}{\sum\limits_{i=B+1}^{E} \delta_{i-1} P(t,T_i)}. \tag{98}$$

A payer swaption price with a fixed rate K and exercise $T < T_B$ is expressed via an expectation in the martingale measure for the process $S(t)$,

$$\text{Swaption}(T,K) = E[(S(T)-K)^+] \sum_{i=B+1}^{E} \delta_{i-1} P_i, \tag{99}$$

where we introduced a notation for bond values at the origin $P_j = P(0,T_j)$.

To apply formulas (90) and (91) to the swaption pricing, it is sufficient to calculate the derivatives $\partial(\ln S_0)/\partial(\ln l_n)$ and $\partial^2(\ln S_0)/\partial(\ln l_n)\partial(\ln l_m)$, taking for S_0 the swap rate (98) expressed as a function of the initial LIBOR values $l_n = L_n(0)$,

$$S_0 = \frac{\prod\limits_{i=B}^{E-1} (1+\delta_i l_i) - 1}{\sum\limits_{i=B+1}^{E} \delta_{i-1} \prod\limits_{j=i}^{E-1} (1+\delta_j l_j)}. \tag{100}$$

The effective volatility (90) first derived by Andersen and Andreasen (2000) is now standard and is cited in many sources, most of which restrict the calculation of swaptions to a purely log-normal approximation. The formula (91) for the effective skew is new. Piterbarg (2005a) introduced a swap rate approximation with a skew

$$\beta(t) = \frac{\sum_n \beta_n(t)(\sigma_n(t)\cdot\sigma(t))}{(E-B)|\sigma(t)|^2}, \tag{101}$$

which is different from our result because it does not include the convexity of the swap rate contained in the second derivatives over the LIBOR rates. Note, however, that a more recent, unpublished result of Piterbarg (2006b) agrees with ours in the case of one-factor LMM.

A new swaption formula. We conclude this section by listing explicit formulas for the derivatives of the swap rate and the final pricing formula for the European payer swaption in the case of a log-normal LIBOR Model. Andersen and Andreasen (2000) gave a compact expression for the first derivative in terms of the bonds at the origin,

$$\frac{\partial(\ln S_0)}{\partial(\ln l_n)} = \frac{P_n - P_{n+1}}{P_n}\left(\frac{P_E}{P_B - P_E} + \frac{\sum\limits_{i=n+1}^{E}\delta_{i-1}P_i}{\sum\limits_{i=B+1}^{E}\delta_{i-1}P_i}\right), \tag{102}$$

which allows for an efficient computation of the effective volatility

$$\sigma(t) = \sum_{n=B}^{E-1}\frac{\partial(\ln S_0)}{\partial(\ln l_n)}\sigma_n(t). \tag{103}$$

The second derivative is only slightly more cumbersome,

$$\begin{aligned}\frac{\partial^2(\ln S_0)}{\partial(\ln l_n)\partial(\ln l_m)} &= \mathbf{1}_{n=m}\frac{P_{n+1}}{P_n}\frac{\partial(\ln S_0)}{\partial(\ln l_n)} - \frac{(P_n - P_{n+1})(P_m - P_{m+1})}{P_n P_m} \\ &\times\left(\frac{P_B P_E}{(P_B - P_E)^2} + \frac{\sum\limits_{i=S+1}^{\min(m,n)}\delta_{i-1}P_i\sum\limits_{j=\max(m,n)+1}^{E}\delta_{j-1}P_j}{\left(\sum\limits_{i=B+1}^{e}\delta_{i-1}P_i\right)^2}\right).\end{aligned} \tag{104}$$

Setting $\beta_n = 1$ in Eq. (92), we find the effective maturity-dependent skew of the swap rate in the case of a purely log-normal LIBOR market model,

$$\bar{\beta}_T = 1 + \frac{\sum_{n,m=B}^{E-1}\frac{\partial^2(\ln S_0)}{\partial(\ln l_n)\partial(\ln l_m)}\int_0^T(\sigma_n(\tau)\cdot\sigma(\tau))d\tau\int_0^T(\sigma_m(\tau)\cdot\sigma(\tau))d\tau}{\left(\int_0^T|\sigma(\tau)|^2d\tau\right)^2}. \tag{105}$$

Using Eqs. (99) and (6), we obtain a new formula for the price of a payer European swaption with maturity T and strike K,

$$\text{Swaption}(T,K) = \left(\frac{S_0}{\bar{\beta}_T} \mathcal{N}(d_+) - \left(K + \frac{S_0(1-\bar{\beta}_T)}{\bar{\beta}_T} \right) \mathcal{N}(d_-) \right) \sum_{i=B+1}^{E} \delta_{i-1} P_i, \tag{106}$$

where $d_\pm$ and V are defined by Eq. (7). Numerical comparison of Eq. (106) against other alternatives will be presented elsewhere.

6 Conclusions and future directions

We developed a regular technique for projecting a general stochastic volatility process onto a displaced diffusion and worked out the applications to FX rates in cross-currency models and swap rates in LIBOR Market Models. Such a projection profoundly changes dynamic properties of the process; the univariate marginals, however, can sometimes be approximated quite accurately. Having correct univariate marginals is sufficient for the valuation of European options on the underlying process, hence the primary realm of the applications of the technique is to the problems of model calibration to European options on various financial rates. As is known, displaced diffusions are effective in capturing implied volatility skew but not smile, which is the principal limitation of the method. According to a general theorem by Gyöngy, it is always possible (although not computationally easy) to find an equivalent local volatility process with unchanged marginals. Reducing the space of local volatility functions to the class of displaced diffusions introduces an error which can be reduced only by extending the target space of local volatility functions. Thus, the direction of future research is to identify wider classes of tractable volatility processes and develop efficient projection methods. The approach of this work is based on L_2-distance minimization and is perfectly suited for projections onto spaces other than displaced diffusions.

In preparing this work, we greatly benefited from the insights of Vladimir Piterbarg. We are grateful to our colleagues at NumeriX and especially to Gregory Whitten for providing a stimulating research environment.

References

[1] L. Andersen and J. Andreasen (2000) "Volatility skews and extensions of the LIBOR market model," *Appl. Math. Finance*, 7(1), 1–32.

[2] A. Antonov and T. Misirpashaev (2006) "Cross-Currency LMM Model with Efficient Calibration to FX Skew by Markovian Projection," working paper, NumeriX.

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

[4] I. Gyöngy (1986) "Mimicking the One-Dimensional Marginal Distributions of Processes Having an Ito Differential" *Probability Theory and Related Fields*, 71, 501–516.

[5] F. Jamshidian (1997) "Libor and Swap Market Models and Measures," *Finance and Stochastics*, 1, 293–330.

[6] V. Piterbarg (2005a) "Stochastic volatility model with time-dependent skew," *Applied Mathematical Finance*, 12(2),147–185.

[7] V. Piterbarg (2005b) "A Multi-currency Model with FX Volatility Skew," available at SSRN: http://ssrn.com/abstract=685084.

[8] V. Piterbarg (2006a) "Markovian Projection Method for Volatility Calibration," available at SSRN: http://ssrn.com/abstract=906473.

[9] V. Piterbarg (2006b) "Modern Approaches to Stochastic Volatility Calibration," Proceedings of the WBS 3rd Fixed Income Conference, Amsterdam.

[10] E. Schlögl (2002) "A Multicurrency Extension of the Lognormal Interest Rate Market Models," *Finance and Stochastics*, Springer, 6(2), pages 173–196.