Mind the cap

Peter Jäckel*

First version: 2003
Last update: 10th October 2004

Abstract

Market models of interest rates are based on the decomposition of the yield curve into a set of discrete forward rates. In this article, I analyse the implications for non-canonical caplets in the framework of a Libor market model in the presence of deterministic funding spreads against the stochastically evolved market rates which are subject to a user-controlled skew of implied volatilities generated by a displaced diffusion process.

1 Introduction

Few of the readers will have missed the recent proliferation of articles on various aspects of the increasingly popular market models of interest rates. The reasons for this trend are easy to see: market models allow traders to design the risk-neutral volatility functions and correlations for their exotic pricing models as close as they wish to the real-world structure of uncertainty they can see in the market-observables. Despite the fact that risk-neutral model parameters are, from a purely theoretical point of view, not really required to be similar to the real world behaviour, it is intuitively clear that the capturing of fundamental features of real-world dynamics in any given model process will lead to more realistic and thus stable hedge ratios. A poignant example for this is the characteristically unimodal evoution of both realised[1] and implied volatility of any given caplet: having undergone a long and slow ris, just before the caplet's expiry those volatility figures tend to decrease noticeably. Not surprisingly, any trader hedging some exposure to caplet volatilities using the underlying futures contract would like his quants to design the modelling framework to take this into account. Another reason for the plethora of work lately published on market models is the progress made both by computer hardware manufacturers and by practitioners' Monte Carlo techniques. The framework for distributed calculations of simulations using additional variance reduction techniques is more and more readily implemented in all the major derivatives houses, and specifically for the Libor market model, fast drift approximations that obviate the need for short-stepped Euler schemes are available [HJJ01, PPvR02]. What's more, with the recent developments of algorithms that allow for the approximate pricing of products that depend on the exercise strategy of the investor such as Bermudan swaptions [LS98, And00, Jäc02], market models have now become the method of choice for the pricing of complex interest rate derivatives.

All of these developments created an ever more urgent need for fast calibration procedures for the Libor market model that are viable in a production environment. At the heart of any fast calibration procedure is an analytical or semi-analytical pricing formula for the given calibration instruments. Since the Libor market model reprices the canonical caplets by construction, it is only natural that so far most of the attention for analytical approximations of other market instruments has been on swaptions, and

*ABN AMRO, 250 Bishopsgate, London EC2M 4AA.

[1] measured over a suitable window in time

some very impressive formulae have been found [Reb99, JR00, Jäc02, HW00, Sch00]. However, in practice it is also important to be able to calculate (semi-)analytical prices for all the possible caplets, not just those that coincide in expiry and payment date as well as accrual period with the abstract discretisation of the yield curve used within the model's own discetised framework. For instance, a given Libor market model implementation may be based on 1-monthly discrete forward rates, but we may wish to calibrate to caplet prices for contracts on 3-month Libor rates. All of the existing swaption approximations work well whenever there is a significant averaging effect due to the swap rate being effectively a weighted sum of all of the discrete forward rates. In contrast, when a non-canonical rate depends only on a small number of the discrete forward rates, or when the payment frequency of the fixed side of a swap does not match the floating side exactly, the known low-order approximations start to break down, and higher order corrections are required. A 3-month caplet that is to be composed from 1-monthly forward rates is such an example. Another, and probably more important case is the value of a caplet on a 6-month forward Libor rate in a model framework of 3-monthly discrete forward rates, or even an option on a 12-month rate. These situations require some kind of basket approximation, and, ideally, the method should allow for some sort of implied volatility skew to be embedded in the model, and it should take into account the possible funding spread difference between 3-monthly and 6-monthly (or 12-monthly, respectively) Libor rates a firm may be subject to. The obvious application of the latter two features is the Japanese interest rate market where not only the caplet skew is far too pronounced to be ignored, but where in addition to all other complications the fact that most US and European investment houses fund at a rate that is *lower* than Yen-Libor leads to significant pricing implications.

2 A simple Libor market model with a skew

There are many methods to incorporate a skew in a Libor market model. Examples include the constant elasticity of variance model [AA00], quadratic volatility specifications [Züh02b], and jump-diffusion processes [GM02]. For reasons of simplicity, I choose the *displaced diffusion* setup [Rub83] which is also known as *affine volatility* [Züh02a]. In this framework, the discrete forward rates evolve according to the stochastic differential equation

$$\frac{\mathrm{d}(f_i + s_i)}{f_i + s_i} = \mu_i(\boldsymbol{f}, \boldsymbol{s}, t)\mathrm{d}t + \sigma_i(t)\mathrm{d}W_i \tag{1}$$

with some constant shift s_i associated with the forward rate f_i. Equation (1) describes the stochastic evolution of geometric Brownian motion for the quantity $(f_i + s_i)$ with instantaneous deterministic volatility $\sigma_i(t)$ and instantaneous indirectly stochastic drift $\mu_i(\boldsymbol{f}(t), \boldsymbol{s}, t)$. This feature will be important later when we approximate the drift as a constant and thus render the expression $(f_i + s_i)$ as a lognormal variate.

2.1 The skew parametrisation

Any given forward rate is drift-free in its own natural measure, i.e.

$$\mathrm{d}(f + s) = \mathrm{d}f = \sigma(t)(f + s)\mathrm{d}W \ . \tag{2}$$

Since it has become more and more common practice to express the volatility-rate dependence as some equivalent constant elasticity-of-variance parameter [Kar02, HKL02], it is desirable to find a scaling for the skew that allows us to specify the proximity to the lognormal or normal volatility setting directly in a similar fashion, albeit in a somewhat approximate way. One such possible parametrisation is to replace the term $(f+s)$ on the right hand side of equation (2) by $(f \cdot q + f(0)(1-q))$ for some constant q, i.e.

$$\mathrm{d}f = \sigma_q \cdot [\, q \cdot f + (1-q) \cdot f(0)\,] \cdot \mathrm{d}W \ . \tag{3}$$

This approach allows the continuous transition from the lognormal framework for $q = 1$ to the normal model first introduced by Bachelier [Bac00] at $q = 0$. However, this kind of parametrisation has two practical drawbacks. Firstly, the end-users of any model tend to explore the available parameter scales in a rather indiscriminate fashion in order to achieve the skew they desire to model. For $q < 0$, the above parametrisation unfortunately results in a shifted lognormal distribution *with inverted asymmetry* stretching from $-\infty$ to $f_{\text{max}} = f(0) \cdot \left(\frac{q-1}{q}\right)$. In other words, it predicts that the forward rate will at expiry not exceed a certain positive threshold $f_{\text{max}} > f(0)$, but may take on any negative value with potentially quite considerable probability. One of the shortcomings of the extended Vasicek model, in comparison, that the Libor market model is frequently used to remedy for, is that it allows for negative forward rates with an approximately normal distribution. It therefore seems natural to impose a skew limitation at the point where the q parametrisation meets the Bachelier model, or possibly even before. This leads to an alternative parametrisation in a new skew parameter Q that gently approaches the normal model as $Q \to 0$ but requires $Q \in (0, 2)$ by virtue of the following definition:

$$Q := 2^{-\frac{s}{f(0)}} \tag{4}$$

or, equivalently

$$s := -f(0) \log_2 Q \tag{5}$$

which leads to

$$\mathrm{d}f = \sigma_Q \cdot [\, f - \log_2 Q \cdot f(0)\,] \cdot \mathrm{d}W \ . \tag{6}$$

The transformation from the (q, σ_q) to the (Q, σ_Q) parametrisation is given by

$$\begin{array}{rcl} Q &=& 2^{\frac{q-1}{q}} \\ \sigma_Q &=& q \cdot \sigma_q \end{array} \ . \tag{7}$$

It is easy to see that this parametrisation is equivalent to the former at the three most important points: the lognormal model is in both cases given by $q = Q = 1$, the approximation for the square root model is given at $q = Q = {}^1\!/_2$, and the normal (Bachelier) model is given by $q = 0$ and approximated in the limit of $Q \to 0$ but never quite reached. The feature of the Q parametrisation only being able to approach the Bachelier model in the limit of $Q \to$ in a very explicit fashion is, rather subtly, shared by but somewhat disguised in the q parametrisation. In fact, for the q paramatrisation and the Q encoding of the skew alike, analytical formulæas well as Monte Carlo schemes based on approximations to the transfer densities need a distinctive switch to the Bachelier framework as q, or Q, for that matter, vanish, since the transition from displaced lognormal to normal distributions undergoes a singular change at $q = Q = 0$ in a similar fashion as $\int x^{q-1}\,\mathrm{d}x$ switches structurally from $\frac{x^q}{q}$ to $\ln x$ at $q \equiv 0$.

It should be pointed out that the discussion in the following sections equally holds regardless of whether one prefers the q or the Q parametrisations outlined above since both of them result in volatility specifications of *affine* nature and are thus equivalent to the displaced diffusion equation (1). The choice of parametrisation does, in practice, though make a difference to the user-friendliness of a given model, and, in my experience, the limitation of a control parameter such as the skew coefficient Q, or q, respectively, to a *finite* interval tends to be more intuitive. The constraints of the skew control coefficient are directly related to the range of the skew that we want to allow for, and this is elaborated in the next section.

2.2 The skew range

In order to establish whether the restriction $Q > 0$ poses in practice any noticeable limitation, let us define the skew χ as the change in implied volatility incurred at the money as the strike is varied by one $^1\!/_{10}$-th of the forward, i.e.

$$\chi \;:=\; \left.\frac{\mathrm{d}\hat{\sigma}(K)}{\mathrm{d}K}\right|_{K=f} \cdot \frac{f}{10}\;. \tag{8}$$

Since the implied volatility $\hat{\sigma}$ relates to the price given in the limit of $Q \to 0$ via the Black and the Bachelier pricing formulæ, respectively, we have

$$V_{\text{Black}}(f,K,\hat{\sigma},T) \;=\; V_{\text{Bachelier}}(f,K,\hat{\sigma}_{\text{Bachelier}},T) \tag{9}$$

and thus

$$\left.\frac{\partial V_{\text{Black}}(f,K,\hat{\sigma},T)}{\partial K}\right|_{K=f} + \left.\frac{\partial V_{\text{Black}}(f,K,\hat{\sigma},T)}{\partial \hat{\sigma}}\right|_{K=f} \cdot \left.\frac{\mathrm{d}\hat{\sigma}(K)}{\mathrm{d}K}\right|_{K=f} \;=\; \left.\frac{\partial V_{\text{Bachelier}}(f,K,\hat{\sigma}_{\text{Bachelier}},T)}{\partial K}\right|_{K=f}\;. \tag{10}$$

Since the right hand side of equation (10) is exactly given by $-^1\!/_2$ times the discount factor to the payment date, this leads to

$$\left.\frac{\mathrm{d}\hat{\sigma}(K)}{\mathrm{d}K}\right|_{K=f} \;=\; \frac{\mathrm{N}\left(\frac{-\hat{\sigma}\sqrt{T}}{2}\right) - \frac{1}{2}}{f\sqrt{T}\,\varphi\left(\frac{\hat{\sigma}\sqrt{T}}{2}\right)} \tag{11}$$

with $\varphi(x) = \mathrm{d}N(x)/\mathrm{d}x = \mathrm{e}^{-\frac{1}{2}x^2}\Big/\sqrt{2\pi}$. A first-order expansion of the numerator and denominator in $\hat{\sigma}\sqrt{T}$ gives us the approximate rule

$$\chi_{\text{Bachelier}} \;\simeq\; -\frac{\hat{\sigma}}{20}\;. \tag{12}$$

In other words, for implied volatilities around 20%, the maximum (negative) attainable skew of the displaced diffusion model is approximately 1% which is usually more than sufficient for caplets. What's more, the markets that require a stronger skew calibration such as Japan tend to have significantly higher volatilities and this means that the displaced diffusion approach can be calibrated to the skew prevailing there, too.

Of course, it is arguable whether one should allow for negative interest rates at all in any Libor market model. In this context it is helpful to note that the at-the-money-forward skew required in most major interest rate markets for most maturities is considerably less strong than predicted by the normal, or equivalently, extended Vasicek model [Vas77, HW90]. The probabilities of negative interest rates are thus even smaller than in the Hull-White or extended Vasicek model, and should therefore in practice be of no concern. The negativity of rates would be completely suppressed in a CEV modelling framework as suggested in [AA00]. However, the CEV framework suffers from one major drawback: for most market-calibrated parameters especially long-dated forward rates incur a rather too large probability of absorption at zero. This may seem innocuous in comparison to the possibility of stochastic paths to spend some time in the negative domain. However, the absorption feature makes the whole concept of pricing in a risk-neutral measure rather questionable since it jeopardises the existence of an equivalent martingale measure [Pla02]. On the other hand, some people might argue that effective Libor rates should actually be allowed to become temporarily slightly negative, although this line of reasoning almost inevitably leads to a debate based on economic grounds that is of no particular relevance here.

Nevertheless, should one desire to adjust the displaced diffusion framework not to allow for negative rates at the expense of an absorbing boundary at zero, it is indeed possible to include such a boundary condition for the affine volatility specification of the displaced diffusion model, and still obtain a very simple closed form solution for options on canonical caplets [Züh02b]. Also, the incorporation of an absorbing boundary at zero poses no problem to any Monte Carlo simulation whatsoever. However, as shown in [Züh02b], unless we are concerned with calibration to extremely far out-of-the-money floorlets, the distinction between the displaced diffusion setup with and without absorbing barrier at zero makes no practical difference for implied volatilities whence I neglect this issue in the following.

One of the attractive features of the stochastic differential equation (1) is that it not only allows for a negative skew, but also for a positive dependence of implied volatilities on the strike. In particular for calibration at high strikes, the positive skew observed in the market poses frequently a severe problem for HJM models based on a quasi-Gaussian evolution of the forward rates such as the extended Vasicek or Hull-White model. Whilst it is usually still possible to calibrate those models at any given strike, the implied risk-neutral distribution as given by a quasi-Gaussian forward rate evolution differs significantly from the distribution as implied by the market's smile [BL78] which can give rise to substantial pricing differences if, for instance, an exotic contract is valued that contains any form of forward rate related digital features.

In analogy to the analysis and expansion that led to the expression for the skew in the Bachelier model given by equation (12), we arrive at the following approximation for the Q skew:

$$\chi_Q \quad \simeq \quad \frac{\hat{\sigma}}{20} \cdot \frac{\log_2 Q}{1 - \log_2 Q} \quad \simeq \quad \frac{\hat{\sigma}}{20} \cdot (q-1) \quad . \tag{13}$$

This formula requires that the parameter Q must be in the interval $(0, 2)$. In fact, for $Q \to 2$, the skew expression diverges. This effect can be understood better if we have a look at the risk-neutral densities shown in figure 1. As we can see, for $Q \gtrsim 3/2$, the density becomes more and more peaked. In fact, in the limit of $Q \to 2$, the density approaches a Dirac distribution. This feature of the displaced diffusion equations for $Q > 1$ bears consequences for any Monte Carlo simulation: when the density is strongly peaked and has a very long but thin tail, the simulation converges rather poorly since most variates are drawn in the area of the peak, and only very few fall in the tail. In this case, it may be advisable to employ importance sampling or *sampler density* [Jäc02] techniques that lay much more emphasis on the long tail and thus improve convergence considerably. In practice, I would recommend not to use values of Q greater than $3/2$.

2.3 The drift conditions in the displaced diffusion framework

Following the convention that the canonical discrete forward rate f_i with associated accrual factor τ_i fixes at time t_i, and that the chosen numéraire is given by a zero coupon maturing at t_N, the drift

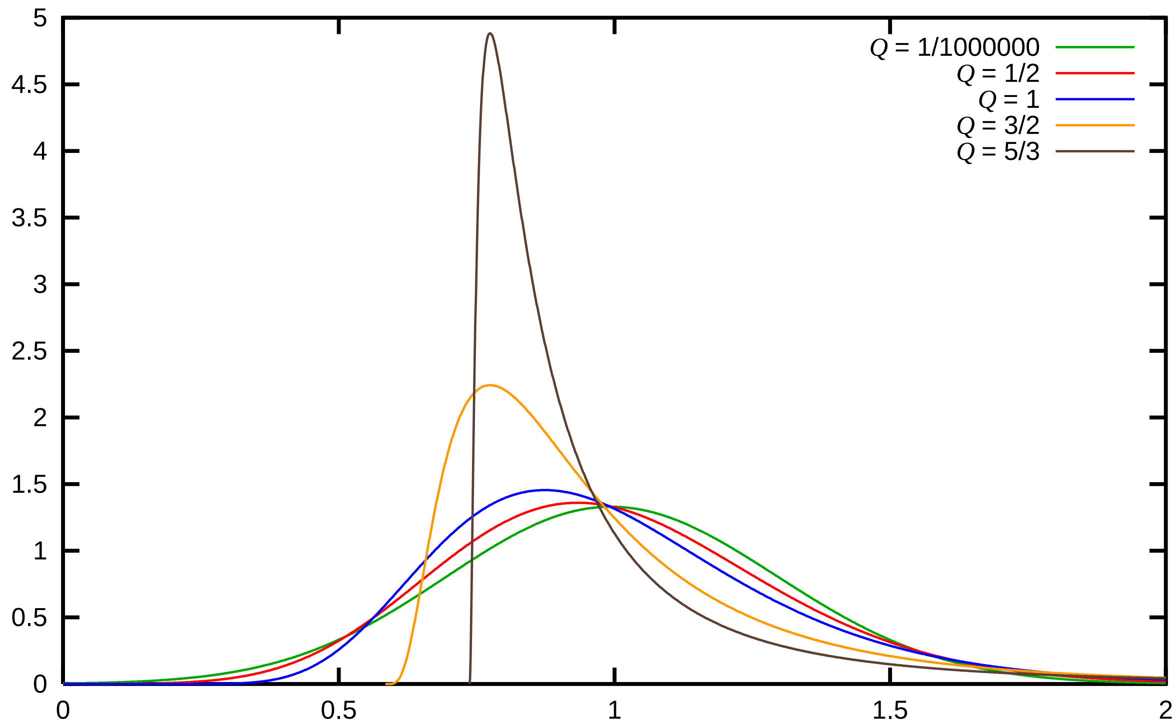

Figure 1: The forward rate density for different levels of the skew coefficient Q with $T = 1$, $\sigma_Q = 30\%/(1 - \log_2 Q)$, and $f_0 = 1$.

conditions for the forward rates subject to the stochastic differential equation (1) are

$$\mu_i\left(\boldsymbol{f}(t), t\right) = \underbrace{-\sigma_i \sum_{k=i+1}^{N-1} \frac{(f_k(t) + s_k)\tau_k}{1 + f_k(t)\tau_k} \sigma_k \rho_{ik}}_{\text{non-zero for } i<N-1} + \underbrace{\sigma_i \sum_{k=N}^{i} \frac{(f_k(t) + s_k)\tau_k}{1 + f_k(t)\tau_k} \sigma_k \rho_{ik}}_{\text{non-zero for } i\geq N} \quad . \tag{14}$$

2.4 Interpolating Libors from canonical discrete forward rates

It is common to highlight the fundamental features of Libor market models using the example of interest rate products that depend only on cashflows occurring precisely on dates coinciding with the model's yield curve discretisation. In practice, however, a Libor market model implementation has to cope with many intermediate cashflows, with settlement delays, fixing conventions, and many other idiosyncracies of the fixed income market. This means that it may be necessary to compute discount factors that span several canonical periods, potentially with a stub discount factor covering only part of the associated discrete forward rate's accrual period. An example for this is given in figure 2. It is difficult to construct non-canonical discount factors from a given set of discrete forward rates in a completely arbitrage-free manner. However, in practice, it is usually sufficient to choose an approximate interpolation rule such that the residual error is well below the levels where arbitrage could be enforced. It is also important to remember that the numerical evaluation of any complex deal with a Libor market model is ultimately still subject to inevitable errors resulting from the calculation scheme: Monte Carlo simulations, non-recombining trees, or recombining trees with their own drift approximation problems. In this context it may not be surprising that the following discount factor interpolation approach is highly accurate for practical purposes.

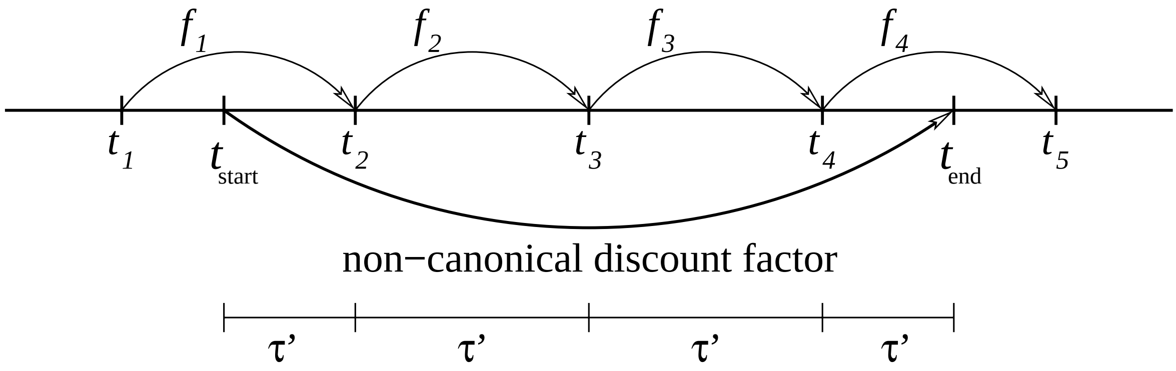

Figure 2: A non-canonical discount factor and its decomposition into canonical forward rates.

Given any forward discount factor $P(t; t_{\text{start}}, t_{\text{end}})$ at time $t \leq t_{\text{start}} < t_{\text{end}}$ that represents the forward funding cost of borrowing one currency unit at time t_{start} and paying back $1/P(t; t_{\text{start}}, t_{\text{end}})$ at time t_{end}, we compute $P(t; t_{\text{start}}, t_{\text{end}})$ from the discrete forward rates according to

$$P(t; t_{\text{start}}, t_{\text{end}}) \;=\; \prod_i \left(1 + f_i(t)\tau_i'\right)^{-1} \;. \tag{15}$$

The product on the right hand side of equation (15) is hereby over all the forward rates that are partly or completely spanned by the discount factor period $(t_{\text{start}}, t_{\text{end}})$. The modified accrual factors τ_i' reflect the potentially partial coverage at either end of the period as depicted in figure 2 where both τ_1' and τ_4' are smaller than τ_1 and τ_4, respectively. When a firm's funding cost happens to be given directly by (forward) Libor rates for a given period τ as they are observed in the market, the relationship between the Libor rate $L(t; t_{\text{start}}, t_{\text{start}} + \tau)$ and the discount factor over the associated accrual period is

$$L(t; t_{\text{start}}, t_{\text{start}} + \tau) \;=\; \left(\frac{1}{P(0; t_{\text{start}}, t_{\text{start}} + \tau)} - 1 \right) \Big/ \, \tau \;. \tag{16}$$

In other words, using the decomposition (15) into canonical forward rates, we have

$$1 + L_\tau \cdot \tau \;=\; \prod_i \left(1 + f_i(t)\tau_i'\right) \tag{17}$$

where I have dropped the explicit mentioning of the dependence on t and t_{start}. In the following, I shall assume that the yield curve is sufficiently smooth in between canonical forward rate dates to justify the simple accrual factor adjustment of the stub periods at either end of the Libor rate accrual interval akin to discrete rate interpolations customary in the short dated money markets. However, it is straightforward to add an additional Libor rate correction factor γ_τ by setting

$$L_\tau \cdot \tau \;=\; \gamma_\tau \cdot \left[\prod_i \left(1 + f_i(t)\tau_i'\right) \;-\; 1 \right] \tag{18}$$

with

$$\gamma_\tau \;=\; \frac{L_\tau(0) \cdot \tau}{\prod_i \left(1 + f_i(0)\tau_i'\right) \;-\; 1} \tag{19}$$

which would correct the Libor rate exactly in the limit of vanishing volatilities. Equation (17) will form the basis of the analytical valuation of non-canonical caplets. First, however, let us have a look at yet another interesting feature of the fixed income market: the spread between funding and interbank offered rates.

2.5 Spread differentials

Most of the major investment houses fund their cash requirements in the Euro, Dollar, and Sterling markets at rates that are very close to the official interbank offered rates. After all, it is precisely this interbank borrowing and lending for funding purposes that originally gave rise to the introduction of the Interbank Offered Rates (IBOR) quotation averages. Some financial institutions, however, have the privilege of higher-than-average credit ratings, and fund themselves accordingly somewhat more cheaply, and others can only borrow at less advantageous rates. In the Yen market, for instance, this phenomenon is particularly pronounced where most of the Western investment banks fund significantly more cheaply than TIBOR. There are several ways to incorporate such a spread between funding and IBOR rates into a market model. In the following, I present a simple procedure based on adjustment factors that are structurally similar to discount factors.

Let us assume that we are building a Libor market model that is based on a 3-monthly canonical forward rate discretisation of the yield curve. In this framework, it may be desirable to be able to price options on forward rate agreements that happen to fall precisely on the canonical dates by a straightforward application of Black's formula and a multiplication by a funding discount factor. In other words, for all forward rates' displacement coefficients Q_i being exactly unity, we may wish to see no skew for such canonical caplet prices struck at different levels. In order to accomplish a setup that allows for spreads, and indeed for spread differentials since the spread between funding rates and 3 month Libor rates may be different than the spread between funding and 6 month rates, I define the (forward) *Libor equivalent discount factor*

$$\tilde{P}(t; t_{\text{start}}, t_{\text{start}} + \tau) \;\; = \;\; \frac{1}{1 + L_\tau(t; t_{\text{start}}) \cdot \tau} \; . \tag{20}$$

Funding discount factors P and Libor equivalent discount factors $\tilde{P}$ are related by virtue of a deterministic *spread factor,* i.e.

$$\tilde{P}(t; t_{\text{start}}, t_{\text{start}} + \tau) \;\; = \;\; P(t; t_{\text{start}}, t_{\text{start}} + \tau) \cdot \zeta_\tau(t_{\text{start}}, t_{\text{start}} + \tau) \; . \tag{21}$$

The spread factor $\zeta_\tau(t_{\text{start}}, t_{\text{start}} + \tau)$ is less than unity whenever funding can be done at a more favourable rate than Libor. Since the spread factor is effectively a credit spread discount factor that represents a simplified amalgamation of default hazard rates into a single number, it is decreasing in the accrual period τ. The decomposition of (forward) funding discount factors now becomes

$$\zeta_{\tau^*}(t_{\text{start}}, t_{\text{end}}) \cdot P(t; t_{\text{start}}, t_{\text{end}}) \;\; = \;\; \prod_i (1 + f_i(t)\tau_i')^{-1} \tag{22}$$

where τ^* stands for the model's canonical discretisation period. All Libor rates that are not for a period that is equal to τ^* can then be computed indirectly via the funding discount factors. This yields

$$L_\tau \cdot \tau \;\; = \;\; \frac{\zeta_{\tau^*}}{\zeta_\tau} \cdot \prod_i (1 + f_i(t)\tau_i') \;\; - \;\; 1 \tag{23}$$

wherein both ζ_{τ^*} and ζ_τ are of course to be taken over the accrual period of the Libor rate L_τ. For $\tau \neq \tau^*$, i.e. when we are interested in a Libor rate that is based on an accrual period different from the model's intrinsic discretisation period, in the presence of a spread differential of the spread between funding and τ-Libor versus the spread between funding and τ^*-Libor, the multiplicative spread ratio term $\frac{\zeta_{\tau^*}}{\zeta_\tau}$ on the right hand side of equation (23) gives rise to a *spread differential induced skew* as we will see in the following.

3 Analytical caplet valuation

The analytical valuation of a caplet[2] is based on the evaluation of the expectation

$$\mathsf{E}\left[(L\cdot\tau - K\cdot\tau)_+\right] \ . \tag{24}$$

3.1 First order approximation ignoring the drift

For τ^* not too large, and for moderate interest rates, a Taylor expansion of the product on the right hand side of equation (23) is an obvious approach:

$$L\cdot\tau \;=\; (\delta-1) \,+\, \delta\cdot\sum_i f_i\tau_i' \,+\, \mathcal{O}\left((f_i\tau_i')^2\right) \quad \text{with} \quad \delta := \frac{\zeta_{\tau^*}}{\zeta_\tau} \ . \tag{25}$$

The right hand side of expansion (25) is $(\delta-1)$ plus a sum of displaced lognormals. In other words, we have a constant term plus a sum of correlated lognormal variates. Now, taking into account the displacements s_i, let us define:-

$$\gamma \;:=\; \frac{L(0)\cdot\tau+1-\delta}{\delta\cdot\sum_i f_i(0)\cdot\tau_i'} \tag{26}$$

$$x_i \;:=\; \gamma\cdot\delta\cdot(f_i+s_i)\cdot\tau_i' \tag{27}$$

$$\kappa \;:=\; K\cdot\tau+1-\delta+\gamma\cdot\delta\cdot\sum_i s_i\cdot\tau_i' \tag{28}$$

This enables us to write the first order approximation for (24) as

$$\mathsf{E}\left[\left(\sum_i x_i - \kappa\right)_+\right] \ . \tag{29}$$

Note that the scaling factor γ was introduced to ensure that the (undiscounted) forward contract $\mathsf{E}[(\sum_i x_i - \kappa)]$ is priced exactly.

Ignoring the fact that most of the involved forward rates are not drift-free in the terminal payment measure of the caplet, we can evaluate (29) as a basket option on a linear combination of lognormal variates x_i with individual expectations $x_i(0)$ struck at κ. This means we have now reduced the first order caplet approximation to the calculation of the expectation (29) where the x_i are lognormal variates with expectations

$$\mathsf{E}[x_i] \;=\; x_i(0) \;=\; \gamma\cdot\delta\cdot(f_i(0)+s_i)\cdot\tau_i' \tag{30}$$

and log-covariances

$$\mathsf{E}[\ln x_i\cdot\ln x_j] - \mathsf{E}[\ln x_i]\cdot\mathsf{E}[\ln x_j] \;=\; c_{ij} \;=\; \int_0^{T_{\text{expiry}}} \sigma_i(t)\sigma_j(t)\rho_{ij}(t)\,\mathrm{d}t \ . \tag{31}$$

[2] I restrict the discussion to caplets. The translation to floorlets is, naturally, straightforward, and should not pose a problem to the reader if I succeed in my attempt to make the exposition of the case of a caplet sufficiently clear.

There are many methods for the approximation of basket options such as Mike Curran's excellent geometric conditioning approach [Cur94], the matching of two moments to a lognormal distribution [Lev92], the matching of three moments to a Johnson distribution (which, incidentally, is the distribution resulting from a displaced diffusion), the method by Turnbull and Wakeman [TW91], or Taylor expansion approaches [Ju01, RDK01]. For the specific case here, however, the particularly fast *rank reduction* method lends itself readily since we can take advantage of the fact that all of the involved forward rates are typically very strongly positively correlated. This method is based on an analysis of the pricing of options on baskets of perfectly correlated lognormally distributed coupons that arises in a single factor extended Vasicek modelling environment [Jam89] and is detailed in appendix A. The rank reduction method works extremely well when correlations are moderate to high, volatilities are at similar levels, and the expectations of the constituents of the basket are also of comparative magnitude. All of these criteria are satisfied by the basket option problem at hand in equation (29). In addition, the rank reduction method is very fast indeed and particularly easy to implement, and all of this is why it is the designated method of choice for the caplet approximation.

3.2 Second order approximation with drift estimate

Let us denote the number of forward rates that contribute to the value of the caplet based on the non-canonical Libor rate L as m. Extending the expansion of the Libor decomposition (23) to second order, we obtain

$$L \cdot \tau \;=\; (\delta - 1) \;+\; \delta \cdot \sum_{i=1}^{m} f_i \tau'_i \;+\; \delta \cdot \sum_{i=1}^{m} \sum_{j=1}^{i-1} f_i \tau'_i f_j \tau'_j \;+\; \mathcal{O}\left((f_i \tau'_i)^3 \right) . \tag{32}$$

This time, it is not immediately obvious how we can substitute the expansion (32) into the caplet pricing formula (24) and treat the resulting expectation as a basket option on a sum of correlated lognormal variates. However, rewriting the second order expansion (32) as

$$\begin{aligned} L \cdot \tau \;\simeq\; & (\delta - 1) + \delta \cdot \sum_{i=1}^{m} \eta (f_i + s_i) \tau'_i - \delta \cdot \sum_{i=1}^{m} \eta s_i \tau'_i + \delta \cdot \sum_{i=1}^{m} \sum_{j=1}^{i-1} \eta (f_i + s_i) \tau'_i \eta (f_j + s_j) \tau'_j \\ & - \delta \cdot \sum_{i=1}^{m} \sum_{j=1}^{i-1} \eta f_i \tau'_i \eta s_j \tau'_j - \delta \cdot \sum_{i=1}^{m} \sum_{j=1}^{i-1} \eta s_i \tau'_i \eta f_j \tau'_j - \delta \cdot \sum_{i=1}^{m} \sum_{j=1}^{i-1} \eta s_i \tau'_i \eta s_j \tau'_j \end{aligned} \tag{33}$$

$$\begin{aligned} =\; & (\delta - 1) - \delta \cdot \sum_{i=1}^{m} \eta s_i \tau'_i + \delta \cdot \sum_{i=1}^{m} \sum_{j=1}^{i-1} \eta s_i \tau'_i \eta s_j \tau'_j \\ & + \delta \cdot \sum_{i=1}^{m} \left(1 + \eta s_i \tau'_i - \sum_{j=1}^{m} \eta s_j \tau'_j \right) \eta (f_i + s_i) \tau'_i + \delta \cdot \sum_{i=1}^{m} \sum_{j=1}^{i-1} \eta (f_i + s_i) \tau'_i \eta (f_j + s_j) \tau'_j \end{aligned} \tag{34}$$

with some constant scaling coefficient η (that is to be determined later) provides some insight. The terms on the right hand side of equation (34) form three groups. The first group consists of all the constant terms on the right hand side of the first line. If we approximate the drift conditions (14) for the forward rates by a constant expression, we can treat the second group as a sum of lognormal variates as it comprises only terms of the form *constant* $\cdot\, (f_i + s_i)$. The last group is then a sum of bilinear combinations of lognormal variates, and this is where we can take advantage of a feature of the lognormal distribution: products of lognormals are again lognormally distributed, and we can compute their expectations and covariances with the original set of lognormals analytically!

Before we proceed to the calculation of the covariances of all the linear and bilinear terms, though, we ought to remember that particularly for caplets on accrual periods that are significantly longer than the model's intrinsic discretisation period, the risk-neutral drift of the involved discrete forward rates is no longer entirely negligible. Chosing the numéraire given by a zero coupon bond that pays one currency unit at the end of the (potentially truncated) accrual period of the last involved discrete forward rate, i.e. at $t_m + \tau'_m$ in our previous notation, we arrive at the following constant drift approximation

$$\mathsf{E}_{T_{\text{expiry}}}[(f_i + s_i)] \quad \approx \quad (f_i(0) + s_i) \cdot \prod_{j=i+1}^{m} \mathrm{e}^{-\frac{\left(f_j(0)+s_j\right)\tau'_j}{1+f_j(0)\tau'_j} c_{ij}} \tag{35}$$

There are, of course, a whole series of rather ad-hoc assumptions in equation (35). As we know, the drift of the discrete forward rates is neither constant nor deterministic[3] due to its instantaneous dependence on the forward rates that bridge the gap between the payment time of any one forward rate and the numéraire asset. This means wherever we have used the initial values for the forward rates in equation (35) we are both using the wrong value to represent the path-average for the evolution of the forward rates (since we are using the initial value), and we are ignoring the indirect stochasticity of the drift since we are using a constant value for each and every forward rate. In my experience, the suppresion of the variance of the drift term due to the stochasticity of the forward rates is typically the dominant error in the constant drift expression. As the drift term is in the exponent, it is Jensen's inequality that is raising its head here. Ignoring the variability of the forward rates in the expression $\frac{(f_j(0)+s_j)\tau'_j}{1+f_j(0)\tau'_j}$ leads to a much bigger discrepancy than the fact that we are ignoring the drift or path-average for f_j when we replace it by a constant value. This phenomenon is reasonably well understood and has led to the development of highly accurate stepwise drift approximations that enable us to construct Monte Carlo schemes that do not need short time steps as we would with the Euler method [HJJ01, PPvR02]. For our caplet calculations, however, this effect is fortunately quite small. Still, we can try to correct for it to some extent by the approximation that each of the terms $(f_j + s_j)$ is almost lognormally distributed, i.e.

$$(f_j + s_j) \quad \approx \quad (f_j(0) + s_j)\,\mathrm{e}^{-\frac{1}{2}c_{jj}+\sqrt{c_{jj}}\cdot z_j} \qquad \text{with} \qquad z_j \sim \mathcal{N}(0,1)\ . \tag{36}$$

In this way, we can expand each of the terms in the product of the right hand side of equation (35) individually in c_{jj} and integrate over an independent normal standard normal distribution for z_j, i.e.

$$\begin{aligned} \mathrm{e}^{\frac{-(f_j+s_j)\mathrm{e}^{-\frac{1}{2}c_{jj}+\sqrt{c_{jj}}\cdot z_j}}{1+(f_j+s_j)\mathrm{e}^{-\frac{1}{2}c_{jj}+\sqrt{c_{jj}}\cdot z_j}-s_j}c_{ij}} \quad &\approx \\ \mathrm{e}^{\frac{-(f_j+s_j)}{1+f_j}c_{ij}} \cdot &\left(1 + \frac{(f_j+s_j)^2\,(1-s_j)\,((1-s_j)\,c_{ij} + 2\,(1+f_j))\,c_{ij}}{2\,(1+f_j)^4}c_{jj}\right) \;+\; \mathcal{O}\left(c_{jj}^2\right) \end{aligned} \tag{37}$$

where I have suppressed the modified accrual factors and dropped all initial value $\cdot(0)$ notation for clarity. Let us now define the approximate expectation for the displaced forward rate using the above expansions as

$$\mathsf{E}_{T_{\text{expiry}}}[(f_i + s_i)] \quad \approx \quad e_i \tag{38}$$

[3] The only exception is, of course, the one forward rate that pays at the same time as the numéraire

with

$$e_i \;:=\; (f_i(0)+s_i)\cdot\prod_{j=i+1}^{m}\mathrm{e}^{-\frac{\left(f_j(0)+s_j\right)\tau_j'}{1+f_j(0)\tau_j'}c_{ij}}\cdot\left(1+\frac{\left(f_j\tau_j'+s_j\tau_j'\right)^2\left(1-s_j\tau_j'\right)\left(\left(1-s_j\tau_j'\right)c_{ij}+2\left(1+f_j\tau_j'\right)\right)c_{ij}}{2\left(1+f_j\tau_j'\right)^4}c_{jj}\right)\;. \tag{39}$$

I now turn the attention to the earlier introduced scaling coefficient η. In analogy to the scaling coefficient γ that we used in the lower order approximation, η is supposed to ensure that our analytical approximation will return the correct expectation of forward rate agreements exactly. To compute η, we need the expectation of all the terms on the right hand side of equation (34). Rearranging the resulting terms as coefficients of a quadratic expression in η, we obtain

$$\mathsf{E}[L\cdot\tau] \;=\; (\delta-1)+\alpha_1\cdot\eta+\alpha_2\cdot\eta^2 \tag{40}$$

with

$$\alpha_1 \;=\; \delta\cdot\sum_{i=1}^{m}\left(e_i\tau_i'-s_i\tau_i'\right) \tag{41}$$

$$\alpha_2 \;=\; \delta\cdot\sum_{i=1}^{m}\sum_{j=1}^{i-1}\left(s_i\tau_i's_j\tau_j'+e_i\tau_i'e_j\tau_j'\mathrm{e}^{c_{ij}}\right)+\delta\cdot\sum_{i=1}^{m}e_i\tau_i'\left(s_i\tau_i'-\sum_{j=1}^{m}s_j\tau_j'\right)\;. \tag{42}$$

Naturally, the solution for η that will ensure the correct value for forward rate agrements within our analytical approximations is

$$\eta \;=\; \begin{cases} \frac{\alpha_1}{2\alpha_2}\left(\sqrt{1+\frac{4\alpha_2}{\alpha_1^2}\left[L(0)\cdot\tau+1-\delta\right]}-1\right) & \text{for}\quad \alpha_2\neq 0 \\ \frac{1}{\alpha_1}\left[L(0)\cdot\tau+1-\delta\right] & \text{for}\quad \alpha_2=0\;. \end{cases} \tag{43}$$

We now have almost all the components that we need to put together an approximate caplet valuation formula based on the rank reduction method applied to an option on the basket of lognormal variates. Since we have a second order expansion of equation (23), the vector of lognormal variates with expectation $\boldsymbol{\xi}$ will in total have

$$N := \frac{m(m+1)}{2} \tag{44}$$

elements of which the first m account for the first order terms, and the remaining $\frac{m(m-1)}{2}$ result from the bilinear combinations. The individual expectations are given by:-

$$\xi_k \;=\; \begin{cases} \delta\eta e_k\tau_k'\left(1-\sum\limits_{j=1,j\neq k}^{m}\eta s_j\tau_j'\right) & & \text{for}\quad k\leq m \\ \delta\eta^2 e_i\tau_i'e_j\tau_j'\cdot\mathrm{e}^{c_{ij}} & \text{with}\quad k=m+\frac{(i-1)(i-2)}{2}+j,\ i=2..m,\ j=1..(i-1) & \text{for}\quad k>m \end{cases} \tag{45}$$

The extended log-covariance matrix C' has N^2 entries. Its elements c_{kl}' can be expressed as sums of elements of the original matrix $C\in\mathbb{R}^{m\times m}$. They are:-

$$c_{kl}' \;=\; \begin{cases} c_{kl} & & \text{for}\quad k\leq m\ \text{and}\ l\leq m \\ c_{il}+c_{jl} & \text{with}\quad k=m+\frac{(i-1)(i-2)}{2}+j,\ i=2..m,\ j=1..(i-1) & \text{for}\quad k>m\ \text{and}\ l\leq m \\ c_{kp}+c_{kq} & \text{with}\quad l=m+\frac{(p-1)(p-2)}{2}+q,\ p=2..m,\ q=1..(p-1) & \text{for}\quad k\leq m\ \text{and}\ l>m \\ c_{ip}+c_{iq}+c_{jp}+c_{jq} & \text{with}\quad \left\{\begin{array}{c} k=m+\frac{(i-1)(i-2)}{2}+j,\ i=2..m,\ j=1..(i-1) \\ \text{and} \\ l=m+\frac{(p-1)(p-2)}{2}+q,\ p=2..m,\ q=1..(p-1) \end{array}\right\} & \text{for}\quad k>m\ \text{and}\ l>m \end{cases} \tag{46}$$

Finally, we need to know the effective strike that is to be used in the basket formula. It is given by

$$\lambda \;:=\; K\cdot\tau + 1 - \delta + \sum_{i=1}^{m} s_i\tau'_i - \sum_{i=1}^{m}\sum_{j=1}^{i-1} s_i\tau'_i s_j\tau'_j \tag{47}$$

Using all of the above definitions, the non-canonical caplet approximation is finally given by the expectation

$$\mathsf{E}\left[\left(\sum_{k=1}^{N} x_k - \lambda\right)_+\right] \tag{48}$$

for lognormal variates x_k with expectations

$$\mathsf{E}[x_k] \;=\; \xi_k \tag{49}$$

and log-covariances

$$\mathsf{E}[\ln x_k \cdot \ln x_l] - \mathsf{E}[\ln x_k]\cdot\mathsf{E}[\ln x_l] \;=\; c'_{kl} \tag{50}$$

which can be computed with any basket approximation such as the rank reduction method given in appendix A.

4 Analysis of the skew resulting from the approximation formulæ

There are various effects that contribute to the skew that we can observe in the implied volatilies of caplets as given by the prices we obtain from Monte Carlo simulations with a Libor market model. First of all there is, of course, the skew that was deliberately put into the model by virtue of, for instance, a displaced diffusion evolution of the canonical forward rates. In addition to that, though, non-canonical caplets incur other effects leading to a skew just by themselves, even if the underlying canonical forward rates were designed to be as lognormally distributed as possible (for instance, by setting $q = Q = 1$).

4.1 The basket effect

The first effect, albeit that it is the smaller out of the two addressed in this article, is due to the fact that a non-canonical caplet bears some similarity to an option on a basket. To analyse this feature, I shall assume that a caplet can indeed be priced very accurately using an expansion of the Libor calculation formula (17) in conjunction with the rank reduction method. To simplify matters, I will also assume that an expansion as presented in sections 3.1 and 3.2 is sufficiently precise not to taint the results significantly. Let the basket pricing formula given by the rank reduction method be denoted by

$$v(\boldsymbol{x}, K, R, C) \tag{51}$$

where $\boldsymbol{x}$ stands for a vector of expectations of displaced forward rates (or products thereof), K is the strike, R is a strike displacement, and C is the effective log-covariance matrix of lognormally distributed variates whose sum comprises the basket. All mentioning of the modified accrual factors τ'_i has been suppressed since they can be absorbed into the the entries of the vector $\boldsymbol{x}$, the strike K, the strike displacement R, and the Libor rate L, respectively. The skew as defined in equation (8) is then implicitly given by the equation

$$\left.\frac{\partial V_{\text{Black}}(L,K,\hat{\sigma},T)\cdot\tau}{\partial K}\right|_{K=L}+\left.\frac{\partial V_{\text{Black}}(L,K,\hat{\sigma},T)\cdot\tau}{\partial\hat{\sigma}}\right|_{K=L}\cdot\left.\frac{\mathrm{d}\hat{\sigma}(K)}{\mathrm{d}K}\right|_{K=L}=\left.\frac{\partial v(\boldsymbol{x},K,R,C)}{\partial K}\right|_{K=L}. \tag{52}$$

The rank reduction approximation involves a modification of the covariance matrix such that its rank is reduced to one, and the calculation of the expectation

$$v(\boldsymbol{x},K,R,C)=\mathsf{E}\left[\left(\sum_{i=1}^{n}x_i\mathrm{e}^{-\frac{1}{2}\tilde{\sigma}_i^2T+\tilde{\sigma}_i\sqrt{T}\,y}-(K+R)\right)_+\right] \tag{53}$$

where the $\tilde{\sigma}_i$ stand for the square root of the diagonal elements of the modified and rank reduced time-to-expiry-averaged covariance matrix, i.e. $\tilde{\sigma}_i=a_i/\sqrt{T}$ with a_i defined in equations (83) and (84) in appendix A. By virtue of the condition $Q>0$, all of the elements of the vector $\boldsymbol{x}$ are positive, and since we assume positive correlation between all forward rates, the expectation in equation (53) can be expressed as

$$v(\boldsymbol{x},K,R,C)=\sum_{i=1}^{n}x_i\cdot\mathrm{N}(\tilde{\sigma}_i\sqrt{T}-y^*)-(K+R)\cdot\mathrm{N}(-y^*) \tag{54}$$

with $y^*=y^*(K)$ being the solution of

$$\sum_{i=1}^{n}x_i\mathrm{e}^{-\frac{1}{2}\tilde{\sigma}_i^2T+\tilde{\sigma}_i\sqrt{T}\,y^*}=(K+R)\;, \tag{55}$$

as shown in appendix A. Equations (54) and (55) can be used to compute the unknown quantity on the right hand side of equation (52). This yields

$$\frac{\partial v}{\partial K}\;=\;\left[(K+R)\cdot\varphi\left(y^*(K)\right)-\sum_{i=1}^{n}x_i\cdot\varphi\left(\tilde{\sigma}_i\sqrt{T}-y^*(K)\right)\right]\cdot\frac{\partial y^*(K)}{\partial K}\;-\;\mathrm{N}\left(-y^*(K)\right)\;. \tag{56}$$

Thanks to the fact that equation (55) can be rewritten as

$$(K+R)\;=\;\sum_{i=1}^{n}x_i\frac{\varphi\left(\tilde{\sigma}\sqrt{T}-y^*(K)\right)}{\varphi\left(y^*(K)\right)}\;, \tag{57}$$

equation (56) can be simplified to

$$\frac{\partial v}{\partial K}\;=\;-\mathrm{N}\left(-y^*(K)\right)\;. \tag{58}$$

As a consequence, for $K=L$, the skew is governed by

$$\left.\frac{\mathrm{d}\hat{\sigma}(K)}{\mathrm{d}K}\right|_{K=L}\;=\;\frac{\mathrm{N}(-\frac{1}{2}\hat{\sigma}\sqrt{T})-\mathrm{N}(-y^*(L))}{\varphi(\frac{1}{2}\hat{\sigma}\sqrt{T})L\sqrt{T}}\;, \tag{59}$$

where $\hat{\sigma}$ stands for the implied Black volatility consistent with the caplet price.

At this point, in order to make some more progress on our understanding of the skew resulting from the basket effect on the skew, I resort to Taylor expansions. First, let us remember that for small ε, we have

$$\mathrm{N}(\varepsilon)\simeq\frac{1}{2}+\frac{\varepsilon}{\sqrt{2\pi}}-\frac{\varepsilon^3}{\sqrt{2\pi}}+\mathcal{O}\left(\varepsilon^5\right)\;. \tag{60}$$

Also, let us recall that *at-the-money* means that the unconditional expectation of the basket is equal to the displaced Libor rate:

$$(L+R) \;=\; \sum_{i=1}^{n} x_i \;. \tag{61}$$

Combining equations (61) and (55), and expanding the exponentials in equation (55) to first order, we can approximate y^* as

$$y^* \;\simeq\; \frac{1}{2}\cdot\frac{\sum_i x_i\tilde{\sigma}_i^2 T}{\sum_i x_i\tilde{\sigma}_i\sqrt{T}} \;. \tag{62}$$

Equally, expanding the at-the-money Black formula

$$V_{\text{Black}}(L,L,\hat{\sigma},T) \;=\; L\cdot\Big[\mathrm{N}(\tfrac{1}{2}\hat{\sigma}\sqrt{T}) - \mathrm{N}(-\tfrac{1}{2}\hat{\sigma}\sqrt{T})\Big] \tag{63}$$

and the rank reduction basket pricing formula (54) at the money for small T using (60), we arrive at

$$\frac{L\hat{\sigma}\sqrt{T}}{\sqrt{2\pi}} \;\simeq\; \frac{\sum_i x_i\tilde{\sigma}_i\sqrt{T}}{\sqrt{2\pi}} \qquad\qquad \text{i.e.} \qquad\qquad L\hat{\sigma} \;\simeq\; \sum_i x_i\tilde{\sigma}_i \;. \tag{64}$$

Now, substituting (64) and (62) into (59), expanding according to (60), and using (61), we obtain

$$\begin{aligned} \left.\frac{\mathrm{d}\hat{\sigma}(K)}{\mathrm{d}K}\right|_{K=L} \;&=\; \frac{1}{2}\mathrm{e}^{\frac{1}{2}\hat{\sigma}^2 T}\frac{1}{L}\left[\frac{\sum_i x_i\tilde{\sigma}_i^2}{\sum_i x_i\tilde{\sigma}_i} - \hat{\sigma}\right] \\ &=\; \frac{1}{2}\mathrm{e}^{\frac{1}{2}\hat{\sigma}^2 T}\frac{1}{L\hat{\sigma}}\left[\sum_i \nu_i\tilde{\sigma}_i^2 - \left(\sum_i \nu_i\tilde{\sigma}_i\right)^2\right] \;, \end{aligned} \tag{65}$$

where I have used the defition

$$\nu_i \;:=\; \frac{x_i}{\sum_i x_i - R} \;\simeq\; \frac{x_i}{L} \quad . \tag{66}$$

A closer look at the terms in the square brackets on the right hand side of equation (65) reveals that, within the scope of the used approximations, the equation can be rewritten as

$$\left.\frac{\mathrm{d}\hat{\sigma}(K)}{\mathrm{d}K}\right|_{K=L} \;=\; \frac{1}{2}\mathrm{e}^{\frac{1}{2}\hat{\sigma}^2 T}\frac{1}{L\hat{\sigma}}\left[\sum_i \nu_i\,(\tilde{\sigma}_i-\hat{\sigma})^2 - \frac{R}{L}\hat{\sigma}^2\right] \;. \tag{67}$$

Obviously, equation (67) implies that the skew is positive for $R \leq 0$, i.e. for $Q \geq 1$. What is interesting about this equation is that it predicts that even for $R = 0$, i.e. for instance for $Q = 1$ (which means that all the canonical forward rates are lognormally distributed in their own natural measure) and in the absence of any spread differential, a non-canonical caplet would display a very small, but *positive* skew, unless all the involved foward rates have identical modified average volatility $\tilde{\sigma}_i$. This means, even when we keep the effective at-the-money volatility of a non-canonical caplet fixed, and even when we keep the effective implied volatility of all canonical caplets fixed or virtually unchanged, it is possible to increase the skew of the given non-canonical caplet ever so slightly by a simple change to the term structure of the instantaneous volatility of the canonical forward rates. This is because, out of all the discrete forward rates that contribute to the value of the non-canonical caplet, at most one of them expires naturally on the same date as the caplet. The values of all the remaining canonical discrete forward rates that eventually contribute to the fixing value of the non-canonical rate that determines the payoff of the caplet are taken as a snapshot *before* their natural expiry. This means that the root-mean-square volatility they realise until the fixing date of the non-canonical caplet is not given by their canonical implied volatility, but misses out on the last part of instantaneous volatility between expiry of the non-canonical caplet and the natural fixing date of the individual contributing discrete forward rates. Since we are free to tailor term structures of instantaneous volatilities of canonical forward rates at will in the Libor market model framework, we can change the shape of the volatility curve, and thus the value of the partially averaged root-mean-square volatility to expiry of the non-canonical caplet, whilst keeping the implied volatility of each canonical caplet unchanged.

Fortunately, the basket effect for non-canonical caplets is very small as long as the non-canonical accrual period doesn't span too many canonical periods and thus proves to be of no practical importance. It is, however, from a theoretical point of view astounding to observe a noticeable effect of the shape of term structure of the canonical forward rates on the skew of non-canonical caplets. It remains to be seen if this kind of effect is also observable in other financial modelling environments, and to what extent it can be detected in the skew of the implied volatilities associated with European swaptions.

As a side note, it may be worth mentioning that the positive sign of the skew effect resulting from the summation of lognormally distributed assets is fairly well known for Asian and basket options when they are approximated by a Johnson distribution. The Johnson distribution is identical to a displaced lognormal distribution. For Asian and basket options, it is fairly straightforward to write down the equations for the matching of the first three moments, and to show that the displacement is negative, thus giving rise to a positive skew when all the underlying constituents are strongly positively correlated.

4.2 The spread differential induced skew

The spread induced displacement is negative if the spread incurred by any one Libor rate is larger than the spread of the Libor that determines the dynamics of the model. For example, if we build the model from a 3m Libor rate with a spread of 10bp (i.e. our funding is 10bp cheaper than 3m Libor), and we have a 20bp Libor spread, then we will end up with a negative spread induced displacement resulting in a positive skew for options on the 6m-Libor rate.

Since I am at this point at serious risk of stretching the readers' patience beyond redemption, I shall only outline the analysis of the spread differential induced skew. As a starting point, we can approximate the non-canonical Libor rate as a single lognormal variate with relative volatility σ subject to a spread differential induced skew as given in equation (25). We assume $\delta \gtrsim 1$ since we place ourselves in the position of a financial institution that funds approximately at the 3m Libor cost but writes a caplet on a longer accrual period for which the equivalent Libor rates are higher than the simple compounding effect for the longer period could justify. The equivalent Black volatility at the money is implicitly (approximately) specified by the leading terms in equation (25), i.e.

$$V_{\text{Black}}(L, K, \hat{\sigma}, T)\big|_{K=L} \;=\; V_{\text{Black}}((1+h\tau)\cdot L^*, K-h, \sigma, T)\big|_{K=L} \ , \tag{68}$$

where I have used the abbreviation $h := (\delta-1)/\,\tau$ and assumed that we can model the basket of canonical forward rates as a single lognormally distributed L^* with volatility $\hat{\sigma}^*$. Straightforward expansions of equation (68) lead to

$$\hat{\sigma} \;=\; \hat{\sigma}^* \cdot \left(1 - \frac{h}{L} + \frac{1}{2}h\tau\right) + \mathcal{O}\left(h^2\right) \ . \tag{69}$$

The next step is then to differentiate (68) with respect to K, and carry out some further Taylor expansions and simplifications. We finally arrive at a dependence of the spread differential induced skew as defined in equation (8) on the at-the-money implied volatility of the non-canonical Libor rate, the spread differential h, and the non-canonical forward rate L itself given by

$$\chi \;\approx\; \frac{\hat{\sigma}^*}{20} \cdot \frac{h}{L} + \mathcal{O}\left(h^2\right) \;. \tag{70}$$

The interesting fact here is that the spread differential induced skew diverges as Libor rates approach zero, and that it grows linearly with h (as long as implied volatilities or times to maturity are small since I used first order expansions in $\hat{\sigma}\sqrt{T}$ and $\hat{\sigma}^*\sqrt{T}$.). For small values of the actual spread and the assumption that spread discount factors are given by

$$\zeta_\tau(t_{\text{start}}, t_{\text{start}} + \tau) \;=\; \mathrm{e}^{-\varepsilon_\tau \cdot \tau} \tag{71}$$

with ε_τ representing the cumulatively compounded spread rate for τ-period Libor rates, we obtain

$$h \cdot \tau \;=\; (\varepsilon_\tau - \varepsilon_{\tau^*}) \cdot \tau \;+\; \mathcal{O}\left(((\varepsilon_\tau - \varepsilon_{\tau^*}) \cdot \tau)^2\right) \;, \tag{72}$$

which means that the skew is approximately linear in the spread differential $h = (\varepsilon_\tau - \varepsilon_{\tau^*})$. If we recall that spread differentials are currently noticeably pronounced in Japan, where rates are low and volatilities high, we may expect the spread differential induced skew to be of non-negligible size in that market.

5 Numerical examples

The first example I give to demonstrate the accuracy of the presented approximations is an option on a 12m Libor rate, expiring in 12 months from inception. All the discrete forward rates that contribute to this caplet are initially set to values near 4%, and are assumed to be perfectly lognormal in their natural measure, i.e. $Q = 1$. I used the same instantaneous term structure of volatility for all of the canonical forward rates given by

$$\sigma_i(t) \;=\; [a + b(T_i - t)] \cdot \mathrm{e}^{-c\cdot(T_i - t)} + d \tag{73}$$

with $a = 0.1$, $b = 1$, $c = 2$, $d = 0.1$, and a time-constant correlation structure given by

$$\rho_{ij} \;=\; \mathrm{e}^{-\beta\cdot(T_i - T_j)} \tag{74}$$

with $\beta = 0.1$. In figure 3, I show the results from numerical simulations using 2^{20} Sobol' vector draws and analytical expansions for the given term structure (labelled as “peaked volatility”) in comparison with the numbers we would obtain if we had set the volatility of all canonical forward rates to 26.05% (denoted as “flat volatility”). The description “first order” refers to the expansion outlined in section 3.1, whereas “second order” is the implied volatility curve resulting from the method explained in section 3.2. The skew as defined in equation (8) associated with the curves is given in table 1. As we can see, the agreement of the first order expansion with the numerical results is for practical purposes just about at the edge of being useful, whereas the agreement of the second order expansion with the numerical data is rather excellent indeed.

The figure and table highlight several features that we had already identified in the analytical discussion in section 4. Firstly, there is clear evidence of the small but positive skew as a consequence of

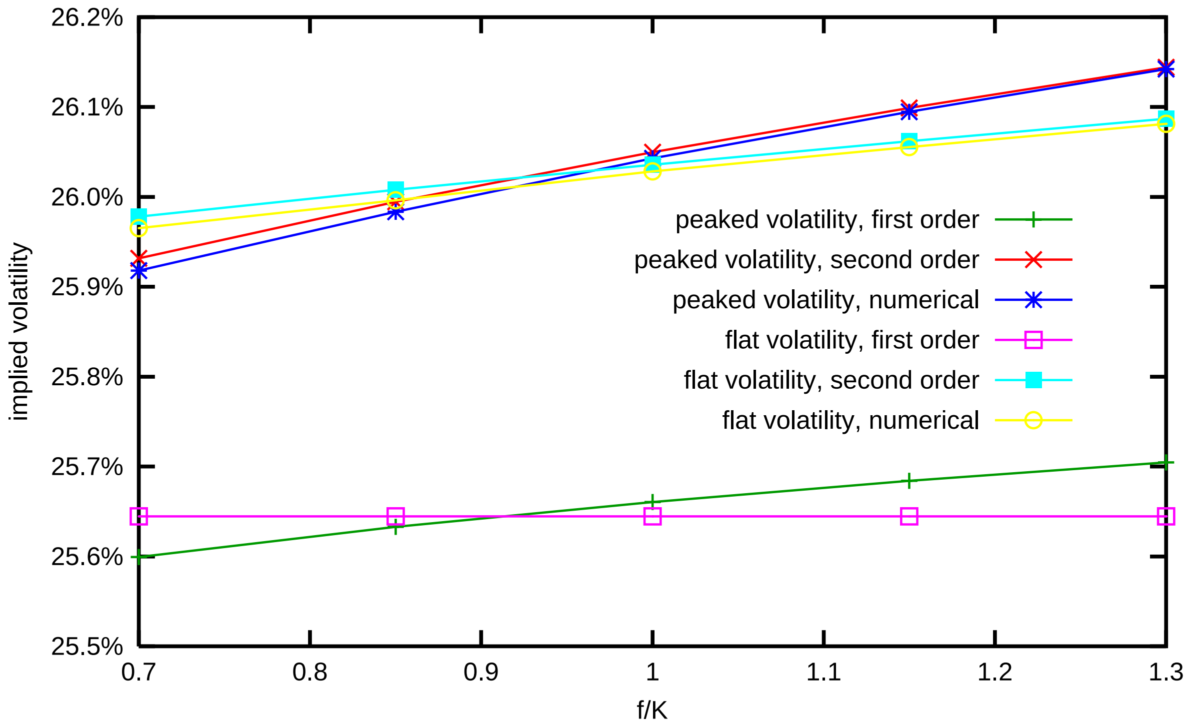

Figure 3: The implied volatilities of a 12m caplet on a 12m Libor rate.

volatility type	first order expansion χ	second order expansion χ	numerically χ
peaked	0.017%	0.035%	0.037%
flat	0	0.018%	0.020%

Table 1: The skew numbers associated with the curves in figure 3.

the basket effect explained in section 4.1. In accordance with the analysis given in that section, the skew increases as we switch from equal and flat volatility of the forward rates to a peaked term structure of volatility. The term structure of instantaneous volatility gives rise to the effective variances of all the contributing discrete forward rates to differ, and as we can tell from equation (67), this in turn causes the skew to increase. The fact that there is still some residual skew even for flat volatilities can also be explained if we compare the first and second order expansion results. Since the second order expansion takes into account the effect of the (nearly) lognormally distributed products of forward rates which have a larger variance than the first order terms, it effectively values a basket of differing constituents, and that in turn causes a slight skew, as discussed at great length by now.

The next example I give is to show the effect of the spread differential for a caplet on a 6m Libor rate with 12 months to expiry as analysed in section 4.2, using lower levels of interest rates and somewhat higher volatilities than before, albeit not quite as extreme as those prevailing in the Japanese market. Again, I set $Q = 1$ for all the canonical forward rates, and I choose the parametrisation $a = 0.2$, $b = 1$, $c = 1$, $d = 0.2$, and $\beta = 0.1$ for the correlation coefficient. A total of 12 curves are displayed in figure 4, representing the implied volatilities computed numerically and analytically (using the second order expansion) from different rate and spread differential settings. In the legend of the figure, the level of the canonical forward rates is indicated by either L=60bp or L=30bp, which is to mean that

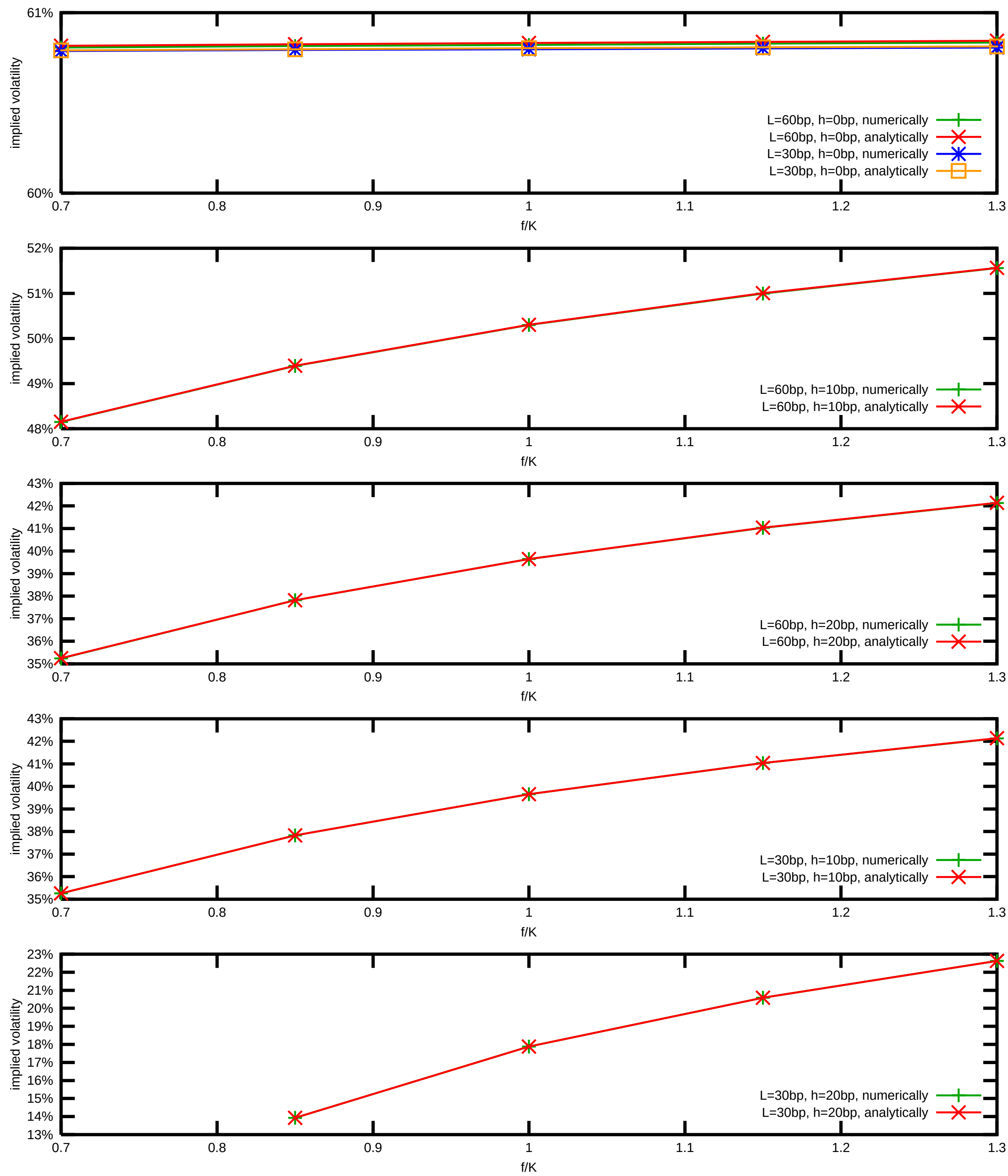

Figure 4: The implied volatility ($\hat{\sigma}$) of a caplet on a 6m Libor rate expiring in 12 months for different levels of the non-canonical forward rate L and different spread differentials h.

the Libor rates are just slightly lower than the given numbers. The spread differential between the 3m canonical rates and the 6m Libor rate is given by h=0bp, h=10bp or h=20bp. In all four cases, the analytical approximation matches the numerically computed results extremely well. This good agreement between numerical and analytical figures for a 12m$\times$6m caplet is not that surprising if we consider that the 6m rate in question is composed of two canonical 3m Libor rates which in turn means that there are no third order terms in equation (3.2) that would be neglected by the approximation given in section 3.2. What's more, just as one would expect from the relationship (70), the implied volatilities for $L \approx 60$bp and the spread differential at 20bp coincide with the values for $L \approx 30$bp and around

10bp spread differential. I should also explain why the point at $f/K = 0.7$ is missing for $L \approx 30$bp and around 20bp spread differential. The reason is that this is where the effective negative displacement of the Libor rate results in a floorlet struck at $0.7 \cdot L$ being perfectly worthless, which is why no equivalent Black volatility can be implied.

To summarise the results on the skew, I give in table 2 the skew figures that were computed from

L	h	χ (numerically)	χ (from analytical prices)	χ from approximation (70)
60bp	0bp	0.0044%	0.0047%	0
30bp	0bp	0.0029%	0.0034%	0
60bp	10bp	0.5357%	0.5361%	0.51%
60bp	20bp	1.0709%	1.0714%	1.01%
30bp	10bp	1.0690%	1.0693%	1.01%
30bp	20bp	2.2158%	2.2160%	2.03%

Table 2: The skew numbers associated with the curves in figure 3.

the results shown in figure 4. Clearly, the significant magnitude of the skew that is induced by spread differentials emphasises how important it is that the forward rates that are evolved in a Libor market model are directly linked to interbank offered rates, and not immediately to funding rates, since this would cause an unintended skew to be built into the model. This is to say that even when we correct the volatility levels such that the effective implied volatilities at the money are calibrated to the market, we still have to bear in mind that there may be a significant skew for non-canonical caplets when spread differentials are present.

Finally, I present an example of the accuracy of the approximations for a user-controlled skew. In order to show how strong the given higher order approximations are, I have chosen the scenario of a non-canonical 3m caplet with 49 months and 2 weeks to expiry in a 3m Libor market model. This means, the non-canonical rate is almost exactly split between two canonical discrete forward rates which makes it a particularly hard test. The volatility parameters are $a = 0.1$, $b = 1$, $c = 2$, $d = 0.1$, and this time I use a term structure of instantaneous volatility given by

$$\rho_{ij}(t) \; = = \mathrm{e}^{-\beta \left| \, |T_i - t|^{\varkappa} - |T_j - t|^{\varkappa} \, \right|} \tag{75}$$

with $\beta = 0.8$ and $\varkappa = 0.2$. This term structure of instantaneous volatility and correlation allows for quite a considerable decorrelation of the forward rates. In addition to that, I used forward rates near 9%. As you can see, the numerical and analytical results agree very well for different levels of the skew, even for options on the Libor rate that are considerably far away from the money.

In summary, I would like to say that I was surprised how complicated it turned out to find a sufficiently accurate caplet approximation in the framework of a Libor market model with a simple user-controlled skew such as given by the stochastic differential equation (1). After all, we are talking here about an interest rate model *that is designed to meet the market features of options on Libor rates by design*, and the pricing of caplets is rarely what the model is originally implemented for. However, since the trading of exotic derivatives valued with a Libor market model requires the model to be reasonably calibrated to market instruments (which sometimes includes options on 6m Libor rates where they are sufficiently liquid, and always includes many non-canonical caplets), and since the handling of many different instruments in a consistent framework requires not only the ability to value all exotics using Monte Carlo simulations, but also the much larger numbers of simpler derivatives such as caps and floors (that are typically in any interest rate option book) in a timely fashion, analytical approximations for caplets and floorlets for the given model may be a very desirable thing to have.

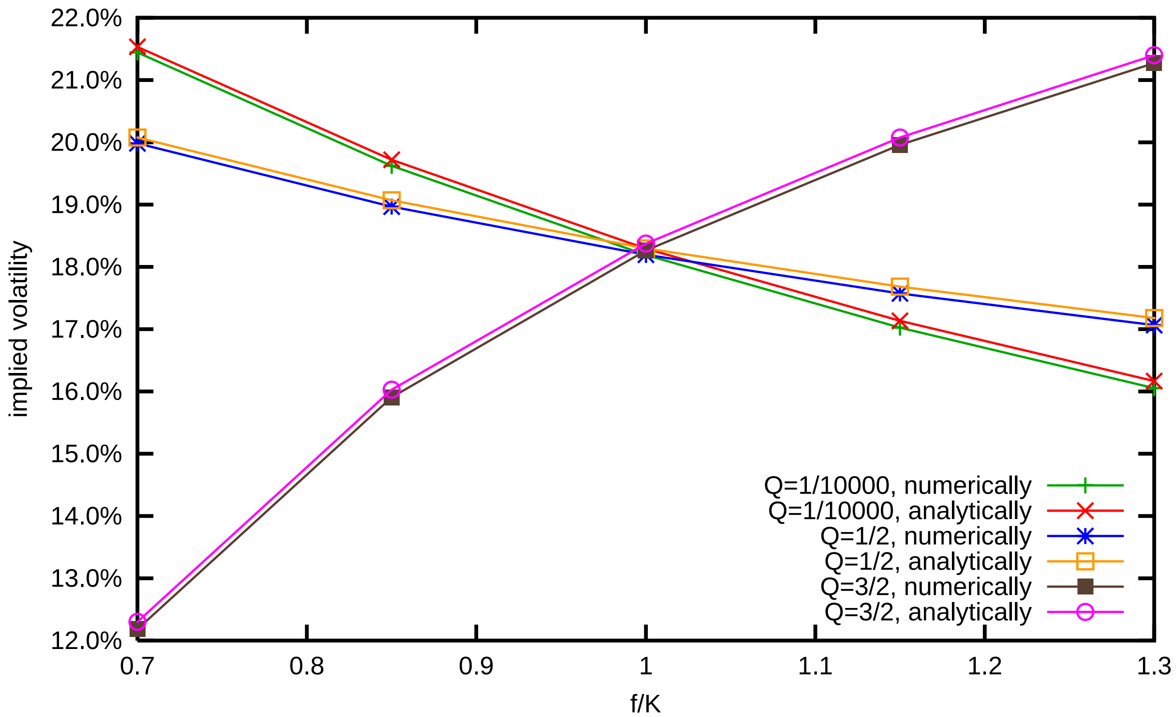

Figure 5: The implied volatility ($\hat{\sigma}$) of a non-canonical caplet on a 3m Libor rate expiring in 49 months and two weeks for different values of the skew parameter Q.

A The rank reduction method for options on baskets of positively correlated lognormals

The problem at hand is the pricing of a call or put option on a weighted average of correlated lognormal variates with expectation f_i. In general, there is no requirement for the fixing of the associated correlated assets to occur simultaneously, which means we could also allow for the pricing of Asian and Asian basket options. All we need for the pricing of the basket option is the covariance matrix C of the logarithmic returns and the weights w_i. When the fixing of all of the involved assets is to be simultaneous at time T, we would have

$$c_{ii} = \sigma_i^2 T \tag{76}$$

and

$$c_{ij} = \sigma_i \sigma_j \rho_{ij} T \quad \text{for} \quad i \neq j \tag{77}$$

using the usual notation for implied volatility and correlation. The basket, or weighted average, of the involved n lognormal variates is given by

$$B = \sum_{i=1}^{n} \omega_i \mathrm{e}^{-\frac{1}{2}c_{ii}+z_i} \tag{78}$$

with the modified weights

$$\omega_i = w_i f_i \tag{79}$$

and the normal variates z_i satisfying the covariance conditions

$$\mathsf{E}[z_i] = 0 \quad \text{and} \quad \mathsf{E}[z_i z_j] = c_{ij} \ . \tag{80}$$

The pricing of an option on the geometric average of lognormal variates can be done without any difficulty since the geometric average is itself lognormally distributed. However, for an arithmetic average, this can only be done if the covariance matrix is of rank 1, subject to an additional criterion that is elaborated in the following.

The key idea of the rank reduction method is to substitute the original covariance matrix C with a matrix C' of rank one such that the log-variance of a geometric basket with the same modified weighting coefficients as B is preserved. In other words, we need to find a covariance matrix C' such that

$$\sum_{i,j=1}^{n} \omega_i \omega_j c_{ij} = \sum_{i,j=1}^{n} \omega_i \omega_j c'_{ij} \tag{81}$$

Any symmetric positive semi-definite matrix C' of rank one can be written as the dyadic product of a vector $\boldsymbol{a}$ with itself:

$$C' = \boldsymbol{a} \cdot \boldsymbol{a}^\top \tag{82}$$

In order to retain the ratios of the standard deviations of all of the constituents, we set

$$a_i := s \cdot \sqrt{c_{ii}} \tag{83}$$

with some common scaling factor s. This factor can be determined from the geometric basket log-variance preserving condition (81) :

$$s := \sqrt{\frac{\sum\limits_{i,j=1}^{n} \omega_i \omega_j c_{ij}}{\sum\limits_{i,j=1}^{n} \omega_i \omega_j \sqrt{c_{ii} c_{jj}}}} \tag{84}$$

Once we have computed the coefficients a_i, the approximate (undiscounted) price of a call option on the arithmetically weighted basket struck at K is given by

$$\mathsf{E}\left[\left(\sum_{i=1}^{n} \omega_i \mathrm{e}^{-\frac{1}{2}a_i^2 + a_i y} - K\right)_+\right] \tag{85}$$

where y is a standard normal variate. As long as the function

$$g(y) = \sum_{i=1}^{n} \omega_i \mathrm{e}^{-\frac{1}{2}a_i^2 + a_i y} \tag{86}$$

is monotonic in y, we can compute expectation (85) comparatively easily. A sufficient condition for the monotonicity of the function $g(y)$ is given if all of the weighting coefficients ω_i are positive. For general basket options such as a the option on a bond that not only pays coupons but also demands repayments (which would involve negative weights), this requirement may be too strict. Even when there are some slightly negative weighting coefficients, the function $g(y)$ may still remain monotonic in y. However, for simplicity, we demand at this point that

$$\omega_i \cdot a_i \geq 0 \; . \tag{87}$$

In practice, this restriction rarely poses a problem. Given (87), we can price the call option on the basket by first identifying the critical value y^* where

$$g(y^*) - K = 0 \; . \tag{88}$$

The value y^* can be found by the use of the standard Newton method, and converges very rapidly due to the smoothness of the function g. A good initial guess is usually given by the second order expansion of $g(y)$ in y around zero. Given the definitions

$$\begin{aligned} b &:= \sum_{i=1}^{n} \tfrac{1}{2} a_i^2 \omega_i \mathrm{e}^{-\frac{1}{2}a_i^2} \\ c &:= \sum_{i=1}^{n} a_i \omega_i \mathrm{e}^{-\frac{1}{2}a_i^2} \\ d &:= \sum_{i=1}^{n} \omega_i \mathrm{e}^{-\frac{1}{2}a_i^2} - K \; , \end{aligned}$$

calculate the discriminant $\delta := c^2 - 4bd$. Then, if the discrimant δ is positive, use

$$y_{\text{initial guess from second order expansion}} := \frac{\sqrt{\delta} - c}{2b} \tag{89}$$

as your initial guess, else use

$$y_{\text{initial guess from first order expansion}} := -\frac{d}{c} \; . \tag{90}$$

The second order expansion is usually already within a relative accuracy of 10^{-5} and may thus be a sufficiently precise approximation for y^* for certain applications. Nonetheless, due to the availability of an extremely good initial guess, any subsequent Newton iterations typically converge to sufficient precision within a single step. Having established the critical value y^*, the approximate value of the call option is given by

$$\mathsf{E}\left[\left(\sum_{i=1}^{n} \omega_i \mathrm{e}^{-\frac{1}{2}c_{ii} + z_i} - K\right)_+\right] \;\simeq\; \left(\sum_{i=1}^{n} \omega_i \mathrm{N}(-y^* + a_i) - K\mathrm{N}(-y^*)\right) \tag{91}$$

wherein $\mathrm{N}(\cdot)$ is the cumulative normal distribution function. Equally, the approximation for the value of a put option can be computed as

$$\mathsf{E}\left[\left(K - \sum_{i=1}^{n} \omega_i \mathrm{e}^{-\frac{1}{2}c_{ii} + z_i}\right)_+\right] \;\simeq\; \left(K\mathrm{N}(y^*) - \sum_{i=1}^{n} \omega_i \mathrm{N}(y^* - a_i)\right) \; . \tag{92}$$

References

[AA00] L. Andersen and J. Andreasen. Volatility Skews and Extensions of the Libor Market Model. *Applied Mathematical Finance*, 7(1), 2000.

[And00] L. Andersen. A simple approach to the pricing of Bermudan swaptions in the multifactor LIBOR market model. *The Journal of Computational Finance*, 3(2):5–32, Winter 1999/2000.

[Bac00] Bachelier. *Théorie de la Spéculation*. PhD thesis, Université de Paris, 1900.

[BL78] D. T. Breeden and R. H. Litzenberger. Prices of state-contingent claims implicit in option prices. *Journal of Business*, 51(4):621–651, 1978.

[BS96] A. N. Borodin and P. Salminen. *Handbook of Brownian Motion — Facts and Formulae*. Birkhäuser, 1996. ISBN 3-7643-5463-1.

[Cur94] Mike Curran. Valuing Asian and portfolio options by conditioning on the geometric mean price. *Management science*, 40:1705–1711, 1994.

[GM02] P. Glasserman and N. Merener. Cap and swaption approximations in libor market models with jumps. Working paper, Columbia University, 2002. `www-1.gsb.columbia.edu/faculty/pglasserman/Other/swaption.pdf`.

[HJJ01] C. Hunter, P. Jäckel, and M. Joshi. Getting the Drift. *Risk*, July 2001. `www.rebonato.com/MarketModelPredictorCorrector.pdf`.

[HKL02] P. Hagan, D. Kumar, and A. S. Lesniewski. Managing smile risk. *Wilmott*, pages 84–108, September 2002.

[HW90] J. Hull and A. White. Pricing interest rate derivative securities. *Review of Financial Studies*, 3(4):573–592, 1990.

[HW00] J. Hull and A. White. Forward rate volatilities, swap rate volatilities, and the implementation of the LIBOR market model. *Journal of Fixed Income*, 10(2):46–62, September 2000.

[Jäc02] P. Jäckel. *Monte Carlo methods in finance*. John Wiley and Sons, February 2002.

[Jam89] F. Jamshidian. An exact bond option pricing formula. *Journal of Finance*, 44(9):205–209, 1989.

[JR00] P. Jäckel and R. Rebonato. The link between caplet and swaption volatilities in a Brace-Gatarek-Musiela/Jamshidian framework: approximate solutions and empirical evidence. *The Journal of Computational Finance*, 6(4):41–59, 2003 (submitted in 2000). `www.jaeckel.org/LinkingCapletAndSwaptionVolatilities.pdf`.

[Ju01] Nengjiu Ju. Pricing Asian and Basket Options via Taylor expansion of the underlying volatility. Working paper, The Robert H. Smith School of Business, University of Maryland, College Park, MD 20742, +1 301 405 2934, nju@rhsmith.umd.edu, March 2 2001.

[Kar02] P. Karasinski. Do we need a new market volatility standard? Exploring rate-level dependence of volatility. presented at the Quantitative Finance London 2002 conference, Fixed Income Research Citigroup Schroder Salomon Smith Barney, November 25 2002.

[Lev92] E. Levy. Pricing European average rate currency options. *Journal of International Money and Finance*, 11:474–491, 1992.

[LS98] F. A. Longstaff and E. S. Schwartz. Valuing American options by simulation: A simple least squares approach. Working paper, The Anderson school, UCLA, 1998.

[Pla02] E. Platen. Consistent pricing and hedging for a modified constant elasticity of variance model. presented at the Quantitative Finance London 2002 conference, University of Technology Sydney, November 25 2002.

[PPvR02] R. Pietersz, A. Pelsser, and M. van Regenmortel. Fast drift approximated pricing in the BGM model. Working paper, Erasmus University Rotterdam, 2002. www.few.eur.nl/few/people/pelsser/papers/Pietersz_Pelsser_vanRegenmorte\%l-FastDriftApproxBGM-02.pdf.

[RDK01] E. Reiner, D. Davidov, and R. Kumanduri. A rapidly convergent expansion method for Asian and Basket options. In *Risk Europe*, Paris, April 11th 2001.

[Reb99] R. Rebonato. Calibration of the BGM model. *Risk*, March:74–79, 1999.

[Rub83] M. Rubinstein. Displaced diffusion option pricing. *Journal of Finance*, 38:213–217, March 1983.

[Sch00] J. G. M. Schoenmakers. Calibration of LIBOR models to caps and swaptions: a way around intrinsic instabilities via parsimonious structures and a collateral market criterion. Working paper, Weierstrass Institute, Berlin, 2000. www.wias-berlin.de/publications/preprints/740.

[TW91] S. Turnbull and L. Wakeman. A quick algorithm for pricing european average options. *Journal of Financial and Quantitiative Analysis*, 26(3):377–390, 1991.

[Vas77] O. A. Vasicek. An equilibrium characterisation of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.

[Züh02a] C. Zühlsdorff. Extended libor market models with affine and quadratic volatility. Working paper, University of Bonn, 2002. ftp.wiwi.uni-bonn.de/papers/bgse/2002/bgse6_2002.pdf.

[Züh02b] C. Zühlsdorff. The pricing of derivatives on assets with quadratic volatility. Working paper, University of Bonn, 2002. ftp.wiwi.uni-bonn.de/papers/bgse/2002/bgse5_2002.pdf.