Interest Rate and Credit Models

8. LIBOR Market Model

Andrew Lesniewski

Baruch College
New York

Spring 2020

Outline

LIBOR market model

- The real challenge in modeling interest rates is the existence of a term structure of interest rates embodied in the shape of the forward curve.
- Fixed income instruments typically depend on a segment of the forward curve rather than a single point.
- Pricing such instruments requires thus a model describing a stochastic time evolution of the entire forward curve.
- There exists a large number of term structure models based on different choices of state variables parameterizing the curve, number of dynamic factors, volatility smile characteristics, etc.

LIBOR market model

- The industry standard for interest rates modeling that has emerged over the past few years is the LIBOR market model (LMM).
- Unlike the older approaches (short rate models which we discussed in the previous lecture), where the underlying state variable is an unobservable instantaneous rate, LMM captures the dynamics of the entire curve of interest rates by using the (market observable) LIBOR forwards as its state variables.
- The time evolution of the forwards is given by a set of intuitive stochastic differential equations in a way which guarantees arbitrage freeness of the process.

LIBOR market model

- The model is intrinsically multi-factor, meaning that it captures accurately various aspects of the curve dynamics: parallel shifts, steepenings / flattenings, butterflies, etc.
- In this lecture we discuss two versions of the LMM methodology:
  (i) the classic LMM with a local volatility specification, and
  (ii) its stochastic volatility (SABR style) extension.
- One of the main difficulties experienced by the pre-LMM term structure models is the fact that they tend to produce unrealistic volatility structures of forward rates.
- The persistent “hump” occurring in the short end of the volatility curve leads to overvaluation of instruments depending on forward volatility.

LIBOR market model

- The LMM model offers a solution to this problem by allowing one to impose an approximately stationary volatility and correlation structure of LIBOR forwards.
- This reflects the view that the volatility structure of interest rates retains its shape over time, without distorting the valuation of instruments sensitive to forward volatility.
- On the downside, LMM is far less tractable than, for example, the Hull-White model.
- In addition, it is not Markovian in the sense short rate models are Markovian.
- As a consequence, all valuations based on LMM have to be done by means of Monte Carlo simulations.

LIBOR market model

- We shall consider a sequence of approximately equally spaced dates $0 = T_0 < T_1 < \ldots < T_N$ which will be termed the *standard tenors*.
- A standard LIBOR forward rate L_j, $j = 0, 1, \ldots, N-1$ is associated with a FRA which starts on T_j and matures on T_{j+1}.
- Usually, it is assumed that $N = 120$ and the L_j's are 3 month LIBOR forward rates.
- Note that these dates refer to the actual start and end dates of the contracts rather than the LIBOR "fixing dates", i.e. the dates on which the LIBOR rates settle.
- To simplify the notation, we shall disregard the difference between the contract's start date and the corresponding forward rate's fixing date. Proper implementation, however, must take this distinction into account.

LIBOR market model

- Each LIBOR forward L_j is modeled as a continuous time stochastic process $L_j(t)$. This process gets killed at $t = T_j$, as the LIBOR rate fixes.
- The dynamics of the forward process is driven by an N-dimensional, correlated Wiener process $W_1(t), \ldots, W_{N-1}(t)$.
- We let ρ_{jk} denote the correlation coefficient between $W_j(t)$ and $W_k(t)$:

$$\mathrm{E}\left[dW_j(t)\, dW_k(t)\right] = \rho_{jk}\, dt\,,$$

where E denotes expected value.

No arbitrage condition

- Let us first consider the world in which there is no volatility of interest rates. The shape of the forward curve would be set once and for all by a higher authority, and each LIBOR forward would have a constant value $L_j(t) = L_{j0}$.
- In other words,
$$dL_j(t) = 0,$$
for all j's.
- The fact that the rates are stochastic forces us to replace this simple dynamical system with a system of stochastic differential equations of the form:
$$dL_j(t) = \Delta_j(t)\,dt + C_j(t)\,dW_j(t). \tag{1}$$
- Here
$$\Delta_j(t) = \Delta_j(t, L(t)),$$
$$C_j(t) = C_j(t, L(t)),$$
are the drift and instantaneous volatility, respectively.

No arbitrage condition

- As discussed in Lecture Notes #4, the *no arbitrage* requirement of asset pricing forces a relationship between the drift term and the diffusion term: the form of the drift term depends thus on the choice of numeraire.
- Recall that L_k is a martingale under the T_{k+1}-forward measure Q_k, and so its dynamics reads:
$$dL_k(t) = C_k(t)\, dW_k(t),$$
where $C_k(t)$ is an instantaneous volatility function which will be defined later.
- For $j \neq k$,
$$dL_j(t) = \Delta_j(t)\, dt + C_j(t)\, dW_j(t).$$
Since the j-th LIBOR forward settles at T_j, the process for $L_j(t)$ is killed at $t = T_j$.
- We shall determine the drifts $\Delta_j(t)$ by requiring lack of arbitrage.

No arbitrage condition

- Let us first assume that $j < k$. The numeraires for the measures Q_j and Q_k are the prices $P\left(t, T_{j+1}\right)$ and $P\left(t, T_{k+1}\right)$ of the zero coupon bonds expiring at T_{j+1} and T_{k+1}, respectively.
- Explicitly,

$$P(t, T_{j+1}) = P(t, T_{\gamma(t)}) \prod\nolimits_{\gamma(t) \leq i \leq j} \frac{1}{1 + \delta_i F_i\left(t\right)}, \tag{2}$$

where F_i denotes the OIS forward[1] spanning the accrual period $[T_i, T_{i+1})$, and where $\gamma : [0, \ T_N] \to \mathbb{Z}$ is defined by

$$\gamma\left(t\right) = m + 1, \quad \text{if } t \in [T_m, \ T_{m+1}) \ .$$

- Notice that $P(t, T_{\gamma(t)})$ is the "stub" discount factor over the incomplete accrual period $[t, T_{\gamma(t)}]$.

[1] Recall that all discounting is done on OIS.

No arbitrage condition

- Since the drift of $L_j(t)$ under $\mathbb{Q}_j$ is zero, formula (48) (or (49)) of Lecture #4 yields:

$$\begin{aligned}\Delta_j(t) &= \frac{d}{dt}\left[L_j,\ \log\frac{P(\cdot,T_{k+1})}{P(\cdot,T_{j+1})}\right](t)\\ &= -\frac{d}{dt}\left[L_j,\ \log\prod\nolimits_{j+1\le i\le k}(1+\delta_i F_i)\right](t)\\ &= -\frac{d}{dt}\int_0^t\sum\nolimits_{j+1\le i\le k} dL_j(s)\ \frac{\delta_i dF_i(s)}{1+\delta_i F_i(s)}\\ &= -C_j(t)\sum\nolimits_{j+1\le i\le k}\frac{\rho_{ji}\delta_i C_i(t)}{1+\delta_i F_i(t)}\ ,\end{aligned}$$

  where, in the third line, we have used the fact that the spread between L_j and F_j is deterministic, and thus its contribution to the quadratic variation is zero.
- Similarly, for $j > k$, we find that

$$\Delta_j(t) = C_j(t)\sum\nolimits_{k+1\le i\le j}\frac{\rho_{ji}\delta_i C_i(t)}{1+\delta_i F_i(t)}\ .$$

No arbitrage condition

- We can thus summarize the above discussion as follows. We let $dW_j(t) = dW_j^{Q_k}(t)$ denote the Wiener process under the measure Q_k.
- Then the dynamics of the LMM model is given by the following system of stochastic differential equations. For $t < \min(T_k, T_j)$,

$$dL_j(t) = C_j(t) \times \begin{cases} -\sum_{j+1 \le i \le k} \frac{\rho_{ji}\delta_i C_i(t)}{1+\delta_i F_i(t)}\, dt + dW_j(t), & \text{if } j < k, \\ dW_j(t), & \text{if } j = k, \\ \sum_{k+1 \le i \le j} \frac{\rho_{ji}\delta_i C_i(t)}{1+\delta_i F_i(t)}\, dt + dW_j(t), & \text{if } j > k. \end{cases} \tag{3}$$

- These equations have to be supplied with initial values for the LIBOR forwards:

$$L_j(0) = L_{j0}, \tag{4}$$

where L_{j0} is the current value of the forward.

No arbitrage condition

- In addition to the forward measures discussed above, it is convenient to use the spot measure.
- It is expressed in terms of the numeraire:

$$B(t) = \frac{P(t, T_{\gamma(t)})}{\prod_{1 \leq i \leq \gamma(t)} P(T_{i-1}, T_i)} . \tag{5}$$

- Under the spot measure, the LMM dynamics reads:

$$dL_j(t) = C_j(t) \left(\sum_{\gamma(t) \leq i \leq j} \frac{\rho_{ji} \delta_i C_i(t)}{1 + \delta_i F_i(t)} dt + dW_j(t) \right). \tag{6}$$

Structure of the instantaneous volatility

- So far we have been working with a general instantaneous volatility $C_j(t)$ for the forward $L_j(t)$.
- In practice, we assume $C_j(t)$ to be one of the following standard volatility specifications discussed in Lecture Notes #5:

$$C_j(t) = \begin{cases} \sigma_j(t) & \text{(normal model)}, \\ \sigma_j(t) L_j(t)^{\beta_j} & \text{(CEV model)}, \\ \sigma_j(t) L_j(t) & \text{(lognormal model)}, \\ \sigma_j(t) L_j(t) + \vartheta_j(t) & \text{(shifted lognormal model)}, \end{cases} \tag{7}$$

where the functions $\sigma_j(t)$ and $\vartheta_j(t)$ are deterministic, and where $\beta_j \leq 1$.

Structure of the instantaneous volatility

- In the following, we will be assuming the CEV model specification, and thus the dynamics of the LIBOR forwards is given by the system:

$$dL_j(t) = \sigma_j(t) L_j(t)^{\beta_j} \times \begin{cases} -\sum_{j+1 \le i \le k} \frac{\rho_{ji}\delta_i\sigma_i(t) L_i(t)^{\beta_i}}{1+\delta_i F_i(t)} dt + dW_j(t), & \text{if } j < k, \\ dW_j(t), & \text{if } j = k, \\ \sum_{k+1 \le i \le j} \frac{\rho_{ji}\delta_i\sigma_i(t) L_i(t)^{\beta_i}}{1+\delta_i F_i(t)} dt + dW_j(t), & \text{if } j > k, \end{cases} \tag{8}$$

under Q_k.

Structure of the instantaneous volatility

- Under the spot measure:

$$dL_j(t) = \sigma_j(t) L_j(t)^{\beta_j} \left(\sum_{\gamma(t) \le i \le j} \frac{\rho_{ji} \delta_i \sigma_i(t) L_i(t)^{\beta_i}}{1 + \delta_i F_i(t)} dt + dW_j(t) \right). \quad (9)$$

- We recall from the discussion in Lecture Notes #5 that the CEV model needs care at zero forward.
- Experience shows that the Dirichlet (absorbing) boundary condition at zero works better than the Neumann (reflecting) condition, and we will assume that the Dirichlet condition is imposed.
- What it means is that if a path realizing the process for L_j hits zero, it gets killed and stays zero forever.

Dimension reduction

- In a market where the forward curve spans 30 years, there are 120 quarterly LIBOR forwards and thus 120 stochastic factors.
- So far we have not imposed any restrictions on the number of these factors, and thus the number of Brownian motions driving the LIBOR forward dynamics is equal to the number of forwards.
- Having a large number of factors poses severe problems with the model's implementation. On the numerical side, the "curse of dimensionality" kicks in, leading to unacceptably slow performance.
- On the financial side, the parameters of the model are severely underdetermined and the calibration of the model becomes unstable.
- We are thus led to the idea that only a small number d of independent Brownian motions $Z_a(t)$, $a = 1, \ldots, d$, with

$$\mathrm{E}\left[dZ_a(t)\, dZ_b(t)\right] = \delta_{ab} dt\,, \tag{10}$$

should drive the process. Typically, $d = 1, 2, 3$, or 4.

Dimension reduction

- We set

$$dW_j(t) = \sum_{1 \le a \le d} U_{ja}\, dZ_a(t), \tag{11}$$

where U is an $N \times d$ matrix with the property that UU' is close to the correlation matrix.

- Of course, it is in general impossible to have $UU' = \rho$. We can easily rewrite the dynamics of the model in terms of the independent Brownian motions:

$$dL_j(t) = \Delta_j(t)\, dt + \sum_{1 \le a \le d} B_{ja}(t)\, dZ_a(t), \tag{12}$$

where

$$B_{ja}(t) = U_{ja}\, C_j(t). \tag{13}$$

- We shall call this system the *factor reduced* LMM dynamics. It is the factor reduced form of LMM that is used in practice.

Calibration of the LMM model

- Calibration (to a selected collection of benchmark instruments) is a choice of the model parameters so that the model reprices the benchmark instruments to a desired accuracy.
- The choice of the calibrating instruments is dictated by the characteristics of the portfolio to be managed by the model.
- An special feature of LMM is that it leads to pricing formulas for caps and floors which are consistent with the market practice of quoting the prices of these products in terms of Black's model. This makes the calibration of LMM to caps and floors straightforward.
- On the other hand, from the point of view of the LMM model, swaptions are exotic structures whose fast pricing poses serious challenges. We describe a strategy of dealing with these issues.
- A key ingredient of any efficient calibration methodology for LMM is rapid and accurate swaption valuation.

Approximate valuation of swaptions

- A swap rate is a non-linear function of the underlying LIBOR forward rates.
- The stochastic differential equation for the swap rate implied by the LMM model cannot be solved in closed form, and thus pricing swaptions within LMM requires Monte Carlo simulations.
- This poses a serious issue for efficient model calibration, as such simulations are very time consuming.
- We describe a closed form approximation which can be used to calibrate the model. Consider a standard forward starting swap, whose start and end dates are denoted by T_m and T_n, respectively.

Approximate valuation of swaptions

- Recall from Lecture Notes #1 that the level function of the swap is defined by:

$$A_{mn}(t) = \sum_{m \leq j \leq n-1} \alpha_{j+1} P\left(t, T_{j+1}\right), \tag{14}$$

  where α_j are the day count fractions for fixed rate payments, and where $P\left(t, T_j\right)$ is the time t value of \$1 paid at time T_j.
- Typically, the payment frequency on the fixed leg is not the same as that on the floating leg[2] (which we continue to denote by δ_j).
- This fact causes a bit of a notational nuisance but needs to be taken properly into account for accurate pricing.
- We let $S_{mn}(t)$ denote the corresponding forward swap rate. In order to lighten up the notation, we will suppress the subscripts mn throughout the remainder of this lecture.

[2] Remember, the default convention on US dollar swaps is a semiannual 30/360 fixed leg versus a quarterly act/360 floating leg.

Approximate valuation of swaptions

- A straightforward calculation shows that, under the forward measure Q_k, the dynamics of the swap rate process can be written in the form:

$$dS(t) = \Omega(t, L)dt + \sum_{m \leq j \leq n-1} \Lambda_j(t, L)dW_j(t), \tag{15}$$

where

$$\Omega = \sum_{m \leq j \leq n-1} \frac{\partial S}{\partial F_j} \Delta_j + \frac{1}{2} \sum_{m \leq i,j, \leq n-1} \rho_{ij} \frac{\partial^2 S}{\partial F_i \partial F_j} C_i C_j, \tag{16}$$

and

$$\Lambda_j = \frac{\partial S}{\partial F_j} C_j. \tag{17}$$

- Not surprisingly, the stochastic differential equation for S has a drift term: the forward swap rate is not a martingale under a forward measure.

Approximate valuation of swaptions

- We shift to the martingale measure Q_{mn} (the swap measure),

$$dW(t) = \frac{\sum_{m \leq j \leq n-1} \Lambda_j(t,F) dW_j(t) + \Omega(t,F)dt}{\nu_{mn}(t)}, \tag{18}$$

where

$$\nu_{mn}(t)^2 = \sum_{m \leq i,j \leq n-1} \rho_{ij} \Lambda_i(t,L) \Lambda_j(t,L). \tag{19}$$

- Then the SDE for the swap rate reads

$$dS(t) = \nu(t)\, dW(t). \tag{20}$$

Approximate valuation of swaptions

- In order to be able to use this dynamics effectively, we have to approximate it by quantities with tractable analytic forms.
- The simplest approximation consists in replacing the values of the stochastic forwards $L_j(t)$ by their initial values L_{j0}. This amounts to "freezing" the curve at its current shape.
- Within this approximation, the coefficients in the diffusion process (15) for the swap rate are deterministic:

$$\Lambda_j(t, L) \approx \Lambda_j(t, L_0), \tag{21}$$

and

$$\Omega(t, L) \approx \Omega(t, L_0). \tag{22}$$

- Let $\nu_0(t)$ denote the value of $\nu(t)$ in this approximation, i.e. $\nu_0(t)$ is given by (19) with all $\Lambda_j(t, L)$ replaced by $\Lambda_j(t, L_0)$.

Approximate valuation of swaptions

- The stochastic differential equation (20) can then be solved in closed form,

$$S(t) = S_0 + \int_0^t \nu_0(s)\, dW(s)\,. \tag{23}$$

- This is a normal model with deterministic time dependent volatility and thus the swaption implied normal volatility ζ_{mn} is approximately given by

$$\begin{aligned} \zeta_{mn}^2 &\approx \frac{1}{T_m} \int_0^t \nu_0(s)^2\, ds \\ &= \frac{1}{T_m} \sum_{m \le j,l \le n-1} \rho_{jl} \int_0^t \Lambda_j(s, F_0)\Lambda_l(s, F_0) ds\,. \end{aligned} \tag{24}$$

- This formula is easy to implement, and it yields reasonably accurate results.
- The frozen curve approximation is the lowest order term in the “small noise expansion”. With some extra work, one can compute higher order terms in that expansion.

Parametrization of the volatility surface

- For the purpose of calibration we require that the deterministic instantaneous CEV volatilities $\sigma_j(t)$ in (8) are piecewise constant.
- In order to help the intuition, we organize constant components as a lower triangular matrix in Table 1.
- Clearly, the problem of determining all the $\sigma_{j,i}$'s is vastly overparametrized. Table 1 contains 7140 parameters (assuming $N = 120$)!

Parametrization of the volatility surface

	$t \in [T_0,\ T_1)$	$t \in [T_1,\ T_2)$	$\ldots$	$t \in [T_{N-1},\ T_N)$
$\sigma_0(t)$	0	0	$\ldots$	0
$\sigma_1(t)$	$\sigma_{1,0}$	0	$\ldots$	0
$\sigma_2(t)$	$\sigma_{2,0}$	$\sigma_{2,1}$	$\ldots$	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\sigma_{N-1}(t)$	$\sigma_{N-1,0}$	$\sigma_{N-1,1}$	$\ldots$	0

Figure: 1. General volatility structure

Parametrization of the volatility surface

- A natural remedy to the overparametrization problem is assuming that the instantaneous volatility is *stationary*, i.e.,

$$\begin{aligned} \sigma_{j,i} &= \sigma_{j-i,0} \\ &\equiv \sigma_{j-i}\,, \end{aligned} \tag{25}$$

  for all $i < j$.
- This assumption appears natural and intuitive, as it implies that the structure of cap volatility will look in the future exactly the same way as it does currently.
- Consequently, the "forward volatility problem" plaguing the traditional terms structure models would disappear.
- Under the stationary volatility assumption, the instantaneous volatility has the structure summarized in Table 2.

Parametrization of the volatility surface

	$t \in [T_0,\ T_1)$	$t \in [T_1,\ T_2)$	$\ldots$	$t \in [T_{N-1},\ T_N)$
$\sigma_0(t)$	0	0	$\ldots$	0
$\sigma_1(t)$	σ_1	0	$\ldots$	0
$\sigma_2(t)$	σ_2	σ_1	$\ldots$	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\sigma_{N-1}(t)$	σ_{N-1}	σ_{N-2}	$\ldots$	0

Figure: 2. Stationary volatility structure

Parametrization of the volatility surface

- It is a good idea to reduce the number of parameters even further, and try to find a parametric fit $\sigma_i = h(T_i)$, $i = 1, \ldots, N-1$.
- A popular (but, by no means the only) choice is the hump function

$$h(t) = (at + b)e^{-\lambda t} + \mu. \tag{26}$$

- Despite its intuitive appeal, the stationarity assumption is not sufficient for accurate calibration of the model.
- The financial reason behind this fact appears to be the phenomenon of mean reversion of long term rates.
- Unlike the Vasicek style models, it is impossible to take this phenomenon into account by adding an Ornstein - Uhlenbeck style drift term to the LMM dynamics as this would violate the arbitrage freeness of the model.
- On the other hand, one can achieve a similar effect by suitably specifying the instantaneous volatility function.

Parametrization of the volatility surface

- In order to implement this idea, we assume that the long term volatility structure is given by $\overline{\sigma}_i = \overline{h}(T_i)$, $i = 1, \ldots, N-1$, where $\overline{h}(t)$ is another hump shaped function.
- We then set

$$\sigma_{j,i} = p_i \sigma_{j-i} + q_i \overline{\sigma}_{j-i}, \tag{27}$$

  i.e. the σ's are mixtures of the short term σ's and the equilibrium $\overline{\sigma}$'s.
- The weights p_i and q_i are parametrized so that $p_i, q_i \geq 0$, $p_i + q_i = 1$, and $p_i \to 0$, as $i \to \infty$.
- In other words, as we move forward in time, the volatility structure looks more and more like the long term limit.

Parametrization of the volatility surface

- This specification is summarized in Table 3.
- The lower triangular matrix in Table 3, LMM's internal representation of volatility, is referred to as the *LMM volatility surface*.
- We leave out the details of this methodology, as that would make the presentation a bit tedious.
- In the final result, we have a parametrization of the volatility surface by a manageable number of parameters $\theta = (\theta_1, \ldots, \theta_d)$ (such as the parameters of the hump functions $h(t)$ and $\overline{h}(t)$, and of the weights p_i), such that $\sigma_{j,i} = \sigma_{j,i}(\theta)$ can be calibrated to the market and has an intuitive shape.

Parametrization of the volatility surface

	$t \in [T_0,\ T_1)$	$t \in [T_1,\ T_2)$	$\ldots$	$t \in [T_{N-1},\ T_N)$
$\sigma_0(t)$	0	0	$\ldots$	0
$\sigma_1(t)$	$p_1\sigma_1 + q_1\overline{\sigma}_1$	0	$\ldots$	0
$\sigma_2(t)$	$p_1\sigma_2 + q_1\overline{\sigma}_2$	$p_2\sigma_1 + q_2\overline{\sigma}_1$	$\ldots$	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\sigma_{N-1}(t)$	$p_1\sigma_{N-1} + q_1\overline{\sigma}_{N-1}$	$p_2\sigma_{N-2} + q_2\overline{\sigma}_{N-2}$	$\ldots$	0

Figure: 3. Approximately stationary volatility structure

Parametrization of the correlation matrix

- The central issue is to calibrate the model, at the same time, to the cap / floor and swaption markets in a stable and consistent manner.
- An important part of this process is determining the correlation matrix $\rho = \{\rho_{jk}\}_{0 \leq j,k \leq N-1}$. The dimensionality of ρ is $N(N+1)/2$, clearly far too high to assure a stable calibration procedure.
- A convenient approach to correlation modeling is to use a parameterized form of ρ_{ij}.
- An intuitive and flexible parametrization is given by the formula:

$$\rho_{ij} = \overline{\rho}_{\min(i,j)} + \left(1 - \overline{\rho}_{\min(i,j)}\right) \exp\left(-\beta_{\min(i,j)} \left| T_i - T_j \right|\right), \tag{28}$$

where

$$\overline{\rho}_k = \rho \tanh\left(\alpha T_k\right), \tag{29}$$

and

$$\beta_k = \beta T_k^{-\kappa}. \tag{30}$$

Parametrization of the correlation matrix

- The meaning of the parameters is as follows: ρ is the asymptotic level of correlations, α is a measure of speed at which ρ is approached, β is a the decay rate of correlations, and κ is an asymmetry parameter.
- Intuitively, positive κ means that two consecutive forwards with short maturities are less correlated than two such forwards with long maturities.
- The parameters in this formula can be calibrated by using, for example, historical data.
- A word of caution is in order: this parametrization produces a matrix that is only approximately positive definite.

Optimization

- In order to calibrate the model we seek instantaneous volatility parameters σ_i so that to fit the at the money caplet and swaption volatilities.
- These can be expressed in terms of the instantaneous volatilities are as follows.
- Let ζ_m denote the at the money implied normal volatility of the caplet expiring at T_m.
- Then, within the frozen curve approximation,

$$\zeta_m(\theta)^2 = \frac{1}{T_m} L_{m0}^{2\beta_m} \sum_{0 \le i \le m-1} \sigma_{m,i}(\theta)^2 \, \delta_i, \tag{31}$$

where $\delta_i = T_{i+1} - T_i$, and θ denotes the set of parameters of the LMM volatility surface.

- This relationship is reasonably accurate, and can be used for calibration.
- However, in practice, one needs to improve on this formula by going beyond the frozen curve approximation.

Optimization

- Similarly, for the at the money implied normal volatility ζ_{mn} of the swaption expiring at T_m into a swap maturing at T_n we have an approximate expression:

$$\zeta_{mn}(\theta)^2 = \frac{1}{T_m} \sum_{0 \le i \le m-1} \sum_{m \le j,l \le n-1} \rho_{jl} \Lambda_{j,i} \Lambda_{l,i} \delta_i, \tag{32}$$

  where $\Lambda_{j,i}$ is the (constant) value of $\Lambda_j(s, L_0)$, the frozen curve approximation to (17), for $s \in [T_i, T_{i+1})$.
- Note that the coefficients $\Lambda_{j,i}$ depend on the parameters of the LMM volatility surface.
- The objective function for optimization is given by:

$$\mathcal{L}(\theta) = \sum_m w_m \left(\zeta_m(\theta) - \overline{\zeta}_m\right)^2 + \sum_{m,n} w_{mn} \left(\zeta_{mn}(\theta) - \overline{\zeta}_{mn}\right)^2, \tag{33}$$

  where $\overline{\zeta}_m$ and $\overline{\zeta}_{mn}$ are the market observed caplet and swaption implied normal volatilities.

Optimization

- The coefficients w_m and w_{mn} are weights which allow the user select the calibration instruments and their relative importance.
- Finally, it is a good idea to add a Tikhonov style regularization in order to maintain stability of the calibration.
- A convenient and computationally efficient choice of the Tikhonov penalty term is the integral of the square of the mean curvature:

$$\frac{1}{2}\lambda \iint_{\text{LMM vol surface}} R(u,v)^2 dudv \tag{34}$$

(of elementary differential geometry of surfaces) of the parameterized LMM volatility surface[3].

- The impact of this penalty term is to discourage regions of extreme curvature (such as a sharp ridge along the diagonal) at the expense of slightly less accurate fit.

[3] For computational efficiency, this integral has to be approximated by a discrete sum.

LMM and smile dynamics

- The classic LMM model has a severe drawback: while it is possible to calibrate it so that it matches at the money option prices, it generally misprices out of the money options.
- The main reason for this is its specification. While the market uses stochastic volatility models in order to price out of the money vanilla options, LMM is incompatible with such models.
- In order to remedy the problem, we describe a model that combines the key features of the LMM and SABR models.
- To this end, we assume that the instantaneous volatilities $C_j(t)$ of the forward rates L_j are of the form

$$C_j(t) = \sigma_j(t) L_j(t)^{\beta_j}, \tag{35}$$

with stochastic volatility parameters $\sigma_j(t)$.

LMM and smile dynamics

- Furthermore, we assume that, under the T_{k+1}-forward measure Q_k, the full dynamics of the forward is given by the stochastic system:

$$\begin{aligned} dL_k(t) &= C_k(t)\, dW_k(t), \\ d\sigma_k(t) &= D_k(t)\, dZ_k(t), \end{aligned} \tag{36}$$

where the diffusion coefficient of the process $\sigma_k(t)$ is of the form

$$D_k(t) = \alpha_k(t)\, \sigma_k(t). \tag{37}$$

- Note that $\alpha_k(t)$ is assumed here to be a (deterministic) function of t rather than a constant.
- This extra flexibility is added in order to make sure that the model can be calibrated to market data.

LMM and smile dynamics

- In addition, we impose the following instantaneous volatility structure:

$$\mathrm{E}\left[dW_j(t)\, dZ_k(t)\right] = r_{jk}\, dt, \tag{38}$$

and

$$\mathrm{E}\left[dZ_j(t)\, dZ_k(t)\right] = \eta_{jk} dt. \tag{39}$$

- The block correlation matrix

$$\Pi = \begin{bmatrix} \rho & r \\ r' & \eta \end{bmatrix} \tag{40}$$

is assumed to be positive definite.

No arbitrage condition for SABR / LMM

- Let us now derive the dynamics of such an extended LIBOR market model under the common forward measure Q_k.
- According to the arbitrage pricing theory, the form of the stochastic differential equations defining the dynamics of the LIBOR forward rates depends on the choice of numeraire.
- Under the T_{k+1}-forward measure Q_k, the dynamics of the forward rate $L_j(t)$, $j \neq k$ reads:

$$dL_j(t) = \Delta_j(t)\,dt + C_j(t)\,dW_j(t).$$

No arbitrage condition for SABR / LMM

- We determine the drifts $\Delta_j(t) = \Delta_j(t, L(t), \sigma(t))$ by requiring absence of arbitrage.
- This is essentially the same calculation as in the derivation of the drift terms for the classic LMM, and we can thus summarize the result as follows.
- In order to streamline the notation, we let $dW(t) = dW^{Q_k}(t)$ denote the Wiener process under the measure Q_k.
- Then, as expected,

$$dL_j(t) = C_j(t) \times \begin{cases} -\sum_{j+1 \le i \le k} \dfrac{\rho_{ji}\delta_i C_i(t)}{1+\delta_i F_i(t)}\, dt + dW_j(t), & \text{if } j < k, \\ dW_j(t), & \text{if } j = k, \\ \sum_{k+1 \le i \le j} \dfrac{\rho_{ji}\delta_i C_i(t)}{1+\delta_i F_i(t)}\, dt + dW_j(t), & \text{if } j > k. \end{cases} \tag{41}$$

No arbitrage condition for SABR / LMM

- Similarly, under the spot measure, the SABR / LMM dynamics reads:

$$dL_j(t) = C_j(t)\left(\sum_{\gamma(t)\leq i\leq j}\frac{\rho_{ji}\delta_i C_i(t)}{1+\delta_i F_i(t)}\,dt + dW_j(t)\right). \tag{42}$$

- Let us now compute the drift term $\Gamma_j(t) = \Gamma_j(t, L(t), \sigma(t))$ for the dynamics of $\sigma_j(t)$, $j \neq k$, under Q_k,

$$d\sigma_j(t) = \Gamma_j(t)\,dt + D_j(t)\,dZ_j(t).$$

- Let us first assume that $j < k$.
- The numeraires for the measures Q_j and Q_k are the prices $P\left(t, T_{j+1}\right)$ and $P(t, T_{k+1})$ of the zero coupon bonds maturing at T_{j+1} and T_{k+1}, respectively.

No arbitrage condition for SABR / LMM

- Since the drift of $L_j(t)$ under Q_j is zero, formula (48) of Lecture Notes #4yields:

$$\begin{aligned}\Gamma_j(t) &= \frac{d}{dt}\left[\sigma_j,\ \log\frac{P(\cdot,T_{j+1})}{P(\cdot,T_{k+1})}\right](t)\\ &= -\frac{d}{dt}\left[\sigma_j,\ \log\prod\nolimits_{j+1\le i\le k}(1+\delta_i F_i)\right](t)\\ &= -\sum\nolimits_{j+1\le i\le k} d\sigma_j(t)\ \frac{\delta_i dF_i(t)}{1+\delta_i F_i(t)}\\ &= -D_j(t)\sum_{j+1\le i\le k}\frac{r_{ji}\delta_i C_i(t)}{1+\delta_i F_i(t)}\ dt.\end{aligned}$$

- Similarly, for $j > k$, we find that

$$\Gamma_j(t) = D_j(t)\sum_{k+1\le i\le j}\frac{r_{ji}\delta_i C_i(t)}{1+\delta_i F_i(t)}\ .$$

No arbitrage condition for SABR / LMM

- This leads to the following system:

$$d\sigma_j(t) = D_j(t) \times \begin{cases} -\sum_{j+1\le i\le k} \dfrac{r_{ji}\delta_i C_i(t)}{1+\delta_i F_i(t)}\, dt + dZ_j(t), & \text{if } j<k, \\ dZ_j(t), & \text{if } j=k, \\ \sum_{k+1\le i\le j} \dfrac{r_{ji}\delta_i C_i(t)}{1+\delta_i F_i(t)}\, dt + dZ_j(t), & \text{if } j>k, \end{cases} \tag{43}$$

under Q_k.

- Similarly,

$$d\sigma_j(t) = D_j(t)\left(\sum_{\gamma(t)\le i\le j} \frac{r_{ji}\delta_i D_i(t)}{1+\delta_i F_i(t)}\, dt + dZ_j(t)\right), \tag{44}$$

under the spot measure.

No arbitrage condition for SABR / LMM

- We now plug in the explicit choices made in (35) and (37). Under the T_{k+1}-forward measure Q_k, the dynamics of the full model reads:

$$dF_j(t) = \sigma_j(t) L_j(t)^{\beta_j} \times \begin{cases} -\sum_{j+1 \le i \le k} \dfrac{\rho_{ji}\delta_i\sigma_i(t) L_i(t)^{\beta_i}}{1+\delta_i F_i(t)}\, dt + dW_j(t), & \text{if } j < k, \\ dW_j(t), & \text{if } j = k, \\ \sum_{k+1 \le i \le j} \dfrac{\rho_{ji}\delta_i\sigma_i(t) L_i(t)^{\beta_i}}{1+\delta_i F_i(t)}\, dt + dW_j(t), & \text{if } j > k, \end{cases} \tag{45}$$

and

$$d\sigma_j(t) = \alpha_j(t)\sigma_j(t) \times \begin{cases} -\sum_{j+1 \le i \le k} \dfrac{r_{ji}\delta_i\sigma_i(t) L_i(t)^{\beta_i}}{1+\delta_i F_i(t)}\, dt + dZ_j(t), & \text{if } j < k, \\ dZ_j(t), & \text{if } j = k, \\ \sum_{k+1 \le i \le j} \dfrac{r_{ji}\delta_i\sigma_i(t) L_i(t)^{\beta_i}}{1+\delta_i F_i(t)}\, dt + dZ_j(t), & \text{if } j > k. \end{cases} \tag{46}$$

No arbitrage condition for SABR / LMM

- These equations are supplemented by the initial conditions:

$$\begin{aligned} L_j(0) &= L_{j0}, \\ \sigma_j(0) &= \sigma_{j0}, \end{aligned} \tag{47}$$

  where L_{j0}'s and σ_{j0}'s are the currently observed values.
- Similarly, under the spot measure Q_0, the dynamics is given by the stochastic system:

$$\begin{aligned} dL_j(t) &= \sigma_j(t) L_j(t)^{\beta_j} \left(\sum_{\gamma(t) \le i \le j} \frac{\rho_{ji} \delta_i \sigma_i(t) L_i(t)^{\beta_i}}{1 + \delta_i F_i(t)} \, dt + dW_j(t) \right), \\ d\sigma_j(t) &= \alpha_j(t) \sigma_j(t) \left(\sum_{\gamma(t) \le i \le j} \frac{r_{ji} \delta_i \sigma_i(t) L_i(t)^{\beta_i}}{1 + \delta_i F_i(t)} \, dt + dZ_j(t) \right). \end{aligned} \tag{48}$$

Discretizing SDEs: single equation

- Numerical solution of a stochastic differential equation amounts to generating paths of the state variables given a path of the stochastic drivers of the system, namely the underlying Brownian motion.
- This requires approximating the continuous time system by a discrete time stochastic system.
- Consider first a one factor SDE,

$$\begin{aligned} dX(t) &= A(t, X(t))dt + B(t, X(t))dZ(t), \\ X(0) &= X_0. \end{aligned} \tag{49}$$

- This is equivalent to

$$X(s) = X(t) + \int_t^s A(u, X(u))du + \int_t^s B(u, X(u))dZ(u). \tag{50}$$

Discretizing SDEs: single equation

- Now, if $f(t,x)$ is twice continuously differentiable, then Ito's lemma states

$$df(t,X(t)) = \mathcal{L}^0 f(t,X(t))dt + \mathcal{L}^1 f(t,X(t))dZ(t)\,, \tag{51}$$

where the operators $\mathcal{L}^i$ are defined by

$$\mathcal{L}^0 = \frac{\partial}{\partial t} + A\frac{\partial}{\partial x} + \frac{1}{2}B^2\frac{\partial^2}{\partial x^2}\,, \tag{52}$$

and

$$\mathcal{L}^1 = B\frac{\partial}{\partial x}\,. \tag{53}$$

Discretizing SDEs: single equation

- Applying Ito's lemma (51) to A yields

$$A(s, X(s)) = A(t, X(t)) + \int_t^s \mathcal{L}^0 A(u, X(u))du + \int_t^s \mathcal{L}^1 A(u, X(u))dZ(u)$$
$$\approx A(t, X(t)) + \mathcal{L}^0 A(t, X(t)) \int_t^s du + \mathcal{L}^1 A(t, X(t)) \int_t^s dZ(u).$$

- We can thus approximate

$$\int_t^{t+\delta} A(s, X(s))ds \approx A(t, X(t))\delta + \mathcal{L}^0 A(t, X(t))I_{(0,0)} + \mathcal{L}^1 A(t, X(t))I_{(1,0)}.$$

- Here

$$I_{(0,0)} = \int_t^{t+\delta} \int_t^s du\, ds, \qquad (54)$$
$$I_{(1,0)} = \int_t^{t+\delta} \int_t^s dZ(u)\ ds,$$

are iterated integrals.

Discretizing SDEs: single equation

- Similarly, we make the following approximation:

$$\begin{aligned} B(s, X(s)) &= B(t, X(t)) + \int_t^s \mathcal{L}^0 B(u, X(u)) du + \int_t^s \mathcal{L}^1 B(u, X(u)) dZ(u) \\ &\approx B(t, X(t)) + \mathcal{L}^0 B(t, X(t)) \int_t^s du + \mathcal{L}^1 B(t, X(t)) \int_t^s dZ(u). \end{aligned}$$

- Therefore,

$$\int_t^{t+\delta} B(s, X(s)) dZ(s) \approx B(t, X(t)) \Delta Z(t) + \mathcal{L}^0 B(t, X(t)) I_{(0,1)} + \mathcal{L}^1 B(t, X(t)) I_{(1,1)},$$

where $\Delta Z(t) = Z(t+\delta) - Z(t)$, and

$$\begin{aligned} I_{(0,1)} &= \int_t^{t+\delta} \int_t^s du\, dZ(s), \\ I_{(1,1)} &= \int_t^{t+\delta} \int_t^s dZ(u)\; dZ(s). \end{aligned} \tag{55}$$

Discretizing SDEs: single equation

- As a result, we obtain the following approximation:

$$\begin{aligned} X(t+\delta) = X(t) + A(t,X(t))\delta + B(t,X(t))\Delta Z(t) + \mathcal{L}^0 A(t,X(t)) I_{(0,0)} \\ + \mathcal{L}^1 A(t,X(t)) I_{(1,0)} + \mathcal{L}^0 B(t,X(t)) I_{(0,1)} + \mathcal{L}^1 B(t,X(t)) I_{(1,1)}. \end{aligned} \tag{56}$$

- Note:

$$\begin{aligned} I_{(0,0)} &= \int_t^{t+\delta} (s-t)ds = \frac{1}{2}\delta^2, \\ I_{(1,1)} &= \int_t^{t+\delta} (Z(s)-Z(t))dZ(s) = \frac{1}{2}\left((\Delta Z)^2 - \delta\right), \\ I_{(0,1)} &= \int_t^{t+\delta} (s-t)dZ(s) = \delta\,\Delta Z - I_{(1,0)}, \\ I_{(1,0)} &= \int_t^{t+\delta} (Z(s)-Z(t))ds. \end{aligned} \tag{57}$$

- In particular, it is very fortuitous that $I_{(1,1)}$ can be computed in a closed, easy to simulate form.

Euler's scheme

- This approximation leads to practical discretization schemes of (49).
- We consider a sequence of times $0 = t_0 < t_1 < \ldots < t_m = T$.
- The first such scheme, Euler's scheme, consists in retaining the first three terms on the right hand side of (56):

$$X_{n+1} = X_n + A(t_n, X_n)\delta_n + B(t_n, X_n)\Delta Z_n, \tag{58}$$

  where $\delta_n = t_{n+1} - t_n$, and $\xi_n \sim N(0, 1)$.
- The random variables ξ_n are assumed independent.

Milstein's scheme

- In the second scheme, *Milstein's scheme*, in addition to the terms present in Euler's scheme, we also retain the last term on the right hand side of (56).
- Note that this term is of order of magnitude δ, while the three discarded terms are of order of magnitude $\delta^{3/2}$ and δ^2.
- Explicitly, Milstein's scheme is given by

$$X_{n+1} = X_n + A(t_n, X_n)\delta_n + B(t_n, X_n)\Delta Z_n + \frac{1}{2} B(t_n, X_n)B'(t_n, X_n)(\Delta Z_n^2 - \delta_n),$$

where $'$ denotes the derivative with respect to x.

Discretizing SDEs: systems of equations

- We now consider an n-dimensional state variable $X \in \mathbb{R}^n$ driven by a d-dimensional Brownian motion $Z(t) \in \mathbb{R}^d$,

$$dX_i(t) = A_i(t, X(t))dt + \sum_{1 \leq a \leq d} B_{ia}(t, X(t))dZ_a(t), \tag{59}$$

  where, for simplicity, we assume that the components of Z are independent.
- This implies that

$$X_i(t+\delta) = X_i(t) + \int_t^{t+\delta} A_i(s, X(s))ds + \sum_{1 \leq a \leq d} \int_t^{t+\delta} B_{ia}(s, X(s))dZ_a(s). \tag{60}$$

- The following calculations generalize the calculations we carried out above for the case of a single factor SDE.

Discretizing SDEs: systems of equations

- If $f(t,x)$ is twice continuously differentiable, then Ito's lemma states

$$df(t,X(t)) = \mathcal{L}^0 f(t,X(t))dt + \sum_{1\leq a\leq d} \mathcal{L}^a f(t,X(t))dZ_a(t)\,, \tag{61}$$

where the operators $\mathcal{L}^i$ are defined by

$$\mathcal{L}^0 = \frac{\partial}{\partial t} + \sum_{1\leq i\leq n} A_i \frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{1\leq i,j\leq n}\sum_{1\leq a\leq d} B_{ia}B_{ja}\frac{\partial^2}{\partial x_i \partial x_j}\,, \tag{62}$$

and

$$\mathcal{L}^a = \sum_{1\leq i\leq n} B_{ia}\frac{\partial}{\partial x_i}\,,\ \text{for } a = 1,\ldots,d. \tag{63}$$

Discretizing SDEs: systems of equations

- Applying Ito's lemma (61) to A_i yields

$$\begin{aligned} A_i(s, X(s)) &= A_i(t, X(t)) + \int_t^s \mathcal{L}^0 A_i(u, X(u))du \\ &\quad + \sum_{1\leq a\leq d} \int_t^s \mathcal{L}^a A_i(u, X(u))dZ_a(u) \\ &\approx A_i(t, X(t)) + \mathcal{L}^0 A_i(t, X(t)) \int_t^s du \\ &\quad + \sum_{1\leq a\leq d} \mathcal{L}^a A_i(t, X(t)) \int_t^s dZ_a(u). \end{aligned}$$

Discretizing SDEs: systems of equations

- We can thus approximate

$$\int_t^{t+\delta} A_i(s, X(s))ds \approx A_i(t, X(t))\delta + \mathcal{L}^0 A_i(t, X(t))I_{(0,0)} + \sum_{1\leq a\leq d} \mathcal{L}^a A_i(t, X(t))I_{(a,0)}.$$

- Here

$$I_{(0,0)} = \int_t^{t+\delta}\int_t^s du\, ds, \qquad I_{(a,0)} = \int_t^{t+\delta}\int_t^s dZ_a(u)\, ds, \tag{64}$$

are iterated integrals.

Discretizing SDEs: systems of equations

- Similarly, we make the following approximation:

$$
\begin{aligned}
B_{ia}(s, X(s)) &= B_{ia}(t, X(t)) + \int_t^s \mathcal{L}^0 B_{ia}(u, X(u))du \\
&\quad + \sum_{1 \le b \le d} \int_t^s \mathcal{L}^b B_{ia}(u, X(u)) dZ_b(u) \\
&\approx B_{ia}(t, X(t)) + \mathcal{L}^0 B_{ia}(t, X(t)) \int_t^s du \\
&\quad + \sum_{1 \le b \le d} \mathcal{L}^b B_{ia}(t, X(t)) \int_t^s dZ_b(u).
\end{aligned}
$$

Discretizing SDEs: systems of equations

- Therefore,

$$\int_t^{t+\delta} B_{ia}(s, X(s))dZ_a(s) \approx B_{ia}(t, X(t))\Delta Z_a(t) + \mathcal{L}^0 B_{ia}(t, X(t))I_{(0,a)} + \sum_{1\leq b\leq d} \mathcal{L}^b B_{ia}(t, X(t))I_{(a,b)},$$

where

$$\begin{aligned} I_{(0,a)} &= \int_t^{t+\delta} \int_t^s du\, dZ_a(s), \\ I_{(a,b)} &= \int_t^{t+\delta} \int_t^s dZ_a(u)\ dZ_b(s). \end{aligned} \tag{65}$$

- The integral $I_{(a,b)}$, for $a \neq b$, is known as the *Levy area*.
- There is no close form expression for the Levy area, and it is computationally expensive to simulate.

Discretizing SDEs: systems of equations

- As a result, we obtain the following approximation:

$$
\begin{aligned}
X_i(t+\delta) = X_i(t) + A_i(t, X(t))\delta + \sum_{1\le a\le d} B_{ia}(t, X(t))\Delta Z_a(t) \\
+ \mathcal{L}^0 A_i(t, X(t)) I_{(0,0)} + \sum_{1\le a\le d} \mathcal{L}^a A_i(t, X(t)) I_{(a,0)} \\
+ \sum_{1\le b\le d} \left(\mathcal{L}^0 B_{ib}(t, X(t)) I_{(0,b)} + \sum_{1\le a\le d} \mathcal{L}^a B_{ib}(t, X(t)) I_{(a,b)} \right),
\end{aligned}
\tag{66}
$$

where

$$
\Delta Z_a(t) = Z_a(t+\delta) - Z_a(t).
$$

Discretizing SDEs: systems of equations

- Note that:

$$
\begin{aligned}
I_{(0,0)} &= \int_t^{t+\delta} (s-t)ds \\
&= \frac{1}{2}\delta^2, \\
I_{(a,a)} &= \int_t^{t+\delta} (Z_a(s) - Z_a(t))dZ_a(s) \\
&= \frac{1}{2}\left((\Delta Z_a)^2 - \delta\right), \\
I_{(0,a)} &= \int_t^{t+\delta} (s-t)dZ_a(s) \\
&= \delta\Delta Z_a - I_{(a,0)}, \\
I_{(a,0)} &= \int_t^{t+\delta} (Z_a(s) - Z_a(t))ds
\end{aligned}
\tag{67}
$$

- Note, in particular, that $I_{(a,a)}$ admits a simple, closed form expression.

Integrability condition

- In order to deal with the Levy areas $I_{(a,b)}$, we impose the following *integrability condition*:

$$\mathcal{L}^a B_{ib} = \mathcal{L}^b B_{ia}, \tag{68}$$

or explicitly

$$\sum_{1 \le k \le n} B_{ka} \frac{\partial B_{ib}}{\partial x_k} = \sum_{1 \le k \le n} B_{kb} \frac{\partial B_{ia}}{\partial x_k} . \tag{69}$$

- Note that then

$$\begin{aligned} I_{(a,b)} + I_{(b,a)} &= \int_t^{t+\delta} \int_t^s \left(dZ_a(u)\, dZ_b(s) + dZ_b(u)\, dZ_a(s) \right) \\ &= \Delta Z_a \, \Delta Z_b. \end{aligned} \tag{70}$$

- In other words, the Levy areas $I_{(a,b)}$ and $I_{(b,a)}$ conspire to add up to a simple, easy to simulate expression!

Integrability condition

- Therefore, when the integrability condition holds, (66) can be written as

$$
\begin{aligned}
&X_i(t+\delta) \\
&\quad = X_i(t) + A_i(t, X(t))\delta + \sum_{1\le a\le d} B_{ia}(t, X(t))\Delta Z_a(t) + \frac{1}{2}\mathcal{L}^0 A_i(t, X(t))\delta^2 \\
&\quad + \sum_{1\le a\le d}\Big(\left(\mathcal{L}^a A_i(t, X(t)) - \mathcal{L}^0 B_{ia}(t, X(t))\right) I_{(a,0)} + \mathcal{L}^0 B_{ia}(t, X(t))\Delta Z_a \delta \\
&\quad + \sum_{1\le a\le d}\Big(\frac{1}{2}\mathcal{L}^a B_{ia}(t, X(t))(\Delta Z_a^2 - \delta) + \sum_{a+1\le b\le d} \mathcal{L}^a B_{ib}(t, X(t))\Delta Z_a \Delta Z_b\Big).
\end{aligned}
\tag{71}
$$

- This approximation leads to the following two discretization schemes.

Euler's scheme

- Euler's scheme is obtained by discarding all but the first three terms on the right hand side of (66):

$$X_{i,n+1} = X_{i,n} + A_{i,n}\delta_n + \sum_{1\leq a\leq d} B_{ia,n}\,\Delta Z_{a,n}. \tag{72}$$

- Euler's scheme is of order of convergence $1/2$ meaning that the approximate solution converges (in a suitable norm) to the actual solution at the rate of $\delta t^{1/2}$, as $\delta t \equiv \max \delta t_n \to 0$.

Milstein's scheme

- Milstein's scheme includes the last term in (71):

$$X_{i,n+1} = X_{i,n} + A_{i,n}\delta_n + \sum_{1\le a\le d} B_{ia,n}\,\Delta Z_{a,n} + \sum_{1\le a\le d}\Big(\frac{1}{2}\,\mathcal{L}^a B_{ia,n}(\Delta Z_{a,n}^2 - \delta_n) + \sum_{a+1\le b\le d} \mathcal{L}^a B_{ib,n}\,\Delta Z_{a,n}\,\Delta Z_{b,n}\Big).$$

- This can be rewritten in a more symmetric form as:

$$X_{i,n+1} = X_{i,n} + \Big(A_{i,n} - \frac{1}{2}\sum_{1\le a\le d}\mathcal{L}^a B_{ia,n}\Big)\delta_n + \sum_{1\le a\le d} B_{ia,n}\,\Delta Z_{a,n} + \frac{1}{2}\sum_{1\le a,b\le d}\mathcal{L}^a B_{ib,n}\,\Delta Z_{a,n}\,\Delta Z_{b,n}. \tag{73}$$

- Milstein's scheme is of order of convergence 1 meaning that the approximate solution converges to the actual solution at the rate of δt, as $\delta t \to 0$.

One factor Brownian motion

- There exist many more of less refined methods for simulating a Wiener process; here we describe two of them.
- The *random walk method* is easy to implement at the expense of being rather noisy.
- It represents a Wiener process as a random walk sampled at a finite set of event dates $t_0 < t_1 < \ldots < t_m$:

$$\begin{aligned} Z(t_{-1}) &= 0, \\ Z(t_n) &= Z(t_{n-1}) + \sqrt{t_n - t_{n-1}}\, \xi_n, \quad n = 0, \ldots, m, \end{aligned} \tag{74}$$

where $t_{-1} = 0$, and where ξ_n are i.i.d. random variables with $\xi_n \sim N(0, 1)$.

One factor Brownian motion

- A good method of generating the ξ_n's is to first generate a sequence of uniform pseudorandom numbers u_n (using, say, the Mersenne twister algorithm), and then set

$$\xi_n = N^{-1}(u_n), \tag{75}$$

  where $N^{-1}(x)$ is the inverse cumulative normal function.
- $N^{-1}(x)$ can be efficiently and accurately computed using e.g. the Beasley-Springer-Moro algorithm, see [2].
- This algorithm is superior to the Box-Muller transform method.

One factor Brownian motion

- The *spectral decomposition method* generally leads to much better performance than the random walk method.
- It assures that the simulated process has the same covariance matrix $\mathcal{C}$ as the Wiener process $Z(t)$ sampled at $t_0, t_1, \ldots, t_m$.
- The covariance matrix is explicitly given by:

$$\begin{aligned} \mathcal{C}_{ij} &= \mathrm{E}\left[Z(t_i)Z(t_j)\right] \\ &= \min(t_i,\, t_j). \end{aligned} \tag{76}$$

- Consider the eigenvalue problem for $\mathcal{C}$:

$$\mathcal{C}E_j = \lambda_j E_j, \qquad j = 0, \ldots, m, \tag{77}$$

with orthonormal E_j's.

One factor Brownian motion

- Since the covariance matrix $\mathcal{C}$ is positive definite, all of its eigenvalues λ_j are nonnegative, and we will assume that

$$\lambda_0 \geq \ \ldots \ \geq \lambda_m \geq 0. \tag{78}$$

- We will denote the n-th component of the vector E_j by $E_j(t_n)$, and consider the random variable

$$Z(t_n) = \sum_{0 \leq j \leq m} \sqrt{\lambda_j}\, E_j(t_n)\xi_j, \tag{79}$$

where ξ_j are, again, i.i.d. random variables with $\xi_j \sim N(0,\, 1)$.

One factor Brownian motion

- These numbers are best calculated by applying the inverse cumulative normal function to a sequence of Sobol numbers.
- Alternatively, one could use a sequence of uniform pseudorandom numbers; this, however, leads to a higher sampling variance.
- Then, for each $n = 0, \ldots, m$, $Z(t_n) \sim N(0, t_n)$, and

$$\begin{aligned} \mathrm{E}\left[Z(t_i)Z(t_j)\right] &= \sum_{0 \le k \le m} \lambda_k E_k(t_i) E_k(t_j) \\ &= \mathcal{C}_{ij}. \end{aligned} \tag{80}$$

One factor Brownian motion

- We can thus regard $Z(t_n)$ a realization of the discretized Wiener process[4].
- For computational efficiency, we may want to truncate (79) at some $p < m$.
- This eliminates the *high frequencies* from $Z(t_n)$, and lowers the variance.
- The price for this doing may be systematically lower accuracy.

[4] This realization of the discretized Wiener process is related to the well known Karhounen-Loeve expansion of the (continuous time) Wiener process.

Multi factor Brownian motion

- We now consider the case of a multi-factor Brownian motion $Z_a(t)$, with

$$\mathrm{E}[dZ_a(t)\, dZ_b(t)] = \rho_{ab} dt.$$

- The Cholesky decomposition of ρ yields

$$\rho = LL^{\mathrm{T}}, \tag{81}$$

where L is a $d \times d$ dimensional, lower triangular matrix.

- For example, if

$$\rho = \begin{bmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{bmatrix}, \tag{82}$$

then

$$L = \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho_{12}^2} \end{bmatrix}, \tag{83}$$

Multi factor Brownian motion

- Now, if $X \in \mathbb{R}^d$ is a vector of independent standard normal variables, then LX is a multivariate normal variable with correlation matrix ρ.
- Indeed,

$$\begin{aligned}
\mathrm{E}[(LX)_a (LX)_b] &= \sum_{0 \le k,l \le d} L_{ak} L_{bl} \mathrm{E}[X_k X_l] \\
&= \sum_{0 \le k,l \le d} L_{ak} L_{bl} \delta_{kl} \\
&= \sum_{0 \le k \le d} L_{ak} L_{bk} \\
&= \sum_{0 \le k \le d} L_{ak} (L^{\mathrm{T}})_{kb} \\
&= (LL^{\mathrm{T}})_{kl} \\
&= \rho_{kl}.
\end{aligned}$$

Discretizing LMM: Euler's scheme

- We choose a sequence of event dates $t_0, t_1, \ldots, t_m$, and denote by $L_{jn} \simeq L_j(t_n)$ the approximate solution.
- We also set

$$\begin{aligned} \Delta_{j,n} &= \Delta_j(t_n, L_n), \\ B_{ja,n} &= B_{ja}(t_n, L_{j,n}), \end{aligned} \tag{84}$$

and $\delta_n = t_{n+1} - t_n$.

- Applied to LMM, Euler's scheme (72) reads:

$$L_{j,n+1} = L_{j,n} + \Delta_{j,n}\delta_n + \sum_{1 \le a \le d} B_{ja,n}\, \Delta Z_{a,n}\ , \tag{85}$$

where, as before, $\Delta Z_{a,n} = Z_a(t_{n+1}) - Z_a(t_n)$ is the discretized Brownian motion.

Discretizing LMM: Milstein's scheme

- Fortunately, LMM is in the category of models which satisfy the integrability condition required for Milstein's scheme to work.
- In order to lighten up the notation, let us define:

$$\Upsilon_{jab,n} \equiv B_{ja}(t_n, L_{j,n}) \frac{\partial B_{jb}(t_n, L_{j,n})}{\partial L_j} . \tag{86}$$

- Then Milstein's scheme (73) applied to the LMM model reads:

$$\begin{aligned} L_{j,n+1} = L_{j,n} + \Big(\Delta_{j,n} - \frac{1}{2} \sum_{1\le a\le d} \Upsilon_{jaa,n}\Big)\delta_n \\ + \sum_{1\le a\le d} B_{ja,n}\,\Delta Z_{a,n} + \frac{1}{2} \sum_{1\le a,b\le d} \Upsilon_{jab,n}\,\Delta Z_{a,n}\,\Delta Z_{n,b} . \end{aligned} \tag{87}$$

Efficient drift calculation

- A bit of a challenge lies in handling the drift terms.
- Because of their complexity, their calculation (at each time step) takes up to 50% of the total computation time.
- On the other hand, they are relatively small as compared to the initial values of the LIBOR forwards, and it would be desirable to develop an efficient methodology for accurate approximate evaluation of the drift terms.
- The first and simplest approach consist in "freezing" the values of $F_j(t)$ at the initial value $F_{j,0} \equiv F_j(0)$.
- We precompute the values

$$\Delta_{j,0} \equiv \Delta_j(t, F_0), \tag{88}$$

and use them for the drift terms throughout the simulation.

- This approximation, *the frozen curve approximation*, is rather crude, and does not perform very well when applied to pricing longer dated instruments.

Efficient drift calculation

- Going one step in the low noise expansion beyond the frozen curve approximation produces satisfying results.
- The second approach is a refinement of the frozen curve approximation, and consists in the following.
- From Ito's lemma,

$$\begin{aligned}\Delta_j(t, F(t)) &= \Delta_{j,0} + \int_0^t \mathcal{L}^a \Delta_j(s, F(s)) dZ_a(s) \\ &\simeq \Delta_{j,0} + \mathcal{L}^a \Delta_j(0, F_0) Z_a(t),\end{aligned} \tag{89}$$

where we have suppressed all terms of order higher than $1/2$.

- We thus arrive at the following approximation, *the order* $1/2$ *approximation*:

$$\Delta_{j,1/2}(t) \equiv \Delta_{j,0} + \mathcal{L}^a \Delta_j(0, F_0) Z_a(t). \tag{90}$$

Efficient drift calculation

- The coefficients $\mathcal{L}^a\Delta_j$ in the formula above are explicitly given by the following expressions.
- Under the forward measure Q_k:

$$\mathcal{L}^a\Delta_j = U_{ja}C_j \times \begin{cases} -\sum\limits_{j+1\le i\le k} \frac{\rho_{ji}\delta_i C_i}{1+\delta_i F_i}\left[U_{ja}\frac{\partial C_j}{\partial F_j} + U_{ia}\left(\frac{\partial C_i}{\partial F_i} - \frac{\delta_i C_i}{1+\delta_i F_i}\right)\right], & \text{if } j<k \\ 0, & \text{if } j=k, \\ \sum\limits_{k+1\le i\le j} \frac{\rho_{ji}\delta_i C_i}{1+\delta_i F_i}\left[U_{ja}\frac{\partial C_j}{\partial F_j} + U_{ia}\left(\frac{\partial C_i}{\partial F_i} - \frac{\delta_i C_i}{1+\delta_i F_i}\right)\right], & \text{if } j>k. \end{cases} \tag{91}$$

- Under the spot measure:

$$\mathcal{L}^a\Delta_j = U_{ja}C_j \sum_{\gamma(t)\le i\le j} \frac{\rho_{ji}\delta_i C_i}{1+\delta_i F_i}\left[U_{ja}\frac{\partial C_j}{\partial F_j} + U_{ia}\left(\frac{\partial C_i}{\partial F_i} - \frac{\delta_i C_i}{1+\delta_i F_i}\right)\right]. \tag{92}$$

Efficient drift calculation

- The *order* $3/4$ *approximation*, uses the next order term in the low noise expansion:

$$\Delta_{j,3/4}(t) \equiv \Delta_{j,0}(t, L_0) + \Gamma_{ja} Z_a(t) + \Omega_j t.$$

- Under the forward measure Q_k, the coefficients Γ_{ja} are given by:

$$\Gamma_{ja} = \begin{cases} -C_j \sum\limits_{j+1 \le i \le k} \frac{\rho_{ji}\delta_i C_i}{1+\delta_i L_i} \left[U_{ja} \frac{\partial C_j}{\partial L_j} + U_{ia} \left(\frac{\partial C_i}{\partial L_i} - \frac{\delta_i C_i}{1+\delta_i L_i} \right) \right], & \text{if } j < k, \\ 0, & \text{if } j = k, \\ C_j \sum\limits_{k+1 \le i \le j} \frac{\rho_{ji}\delta_i C_i}{1+\delta_i L_i} \left[U_{ja} \frac{\partial C_j}{\partial L_j} + U_{ia} \left(\frac{\partial C_i}{\partial L_i} - \frac{\delta_i C_i}{1+\delta_i L_i} \right) \right], & \text{if } j > k, \end{cases}$$

Efficient drift calculation

- The coefficients Ω_j are given by:

$$\Omega_j = \begin{cases} -\sum\limits_{j+1\le i\le k} \frac{\rho_{ji}\delta_i}{1+\delta_i L_i}\left[\Delta_{j,0} C_i \frac{\partial C_j}{\partial L_j} + \Delta_{i,0} C_j \left(\frac{\partial C_i}{\partial L_i} - \frac{\delta_i C_i}{1+\delta_i L_i}\right)\right], & \text{if } j < k, \\ 0, & \text{if } j = k, \\ \sum\limits_{k+1\le i\le j} \frac{\rho_{ji}\delta_i}{1+\delta_i L_i}\left[\Delta_{j,0} C_i \frac{\partial C_j}{\partial L_j} + \Delta_{i,0} C_j \left(\frac{\partial C_i}{\partial L_i} - \frac{\delta_i C_i}{1+\delta_i L_i}\right)\right], & \text{if } j > k. \end{cases}$$

Efficient drift calculation

- Under the spot measure,

$$\Gamma_{ja} = C_j \sum_{\gamma(t)\le i\le j} \frac{\rho_{ji}\delta_i C_i}{1+\delta_i L_i}\Big[U_{ja}\frac{\partial C_j}{\partial L_j} + U_{ia}\Big(\frac{\partial C_i}{\partial L_i} - \frac{\delta_i C_i}{1+\delta_i L_i}\Big)\Big],$$

and

$$\Omega_j = \sum_{\gamma(t)\le i\le j} \frac{\rho_{ji}\delta_i}{1+\delta_i L_i}\Big[\Delta_{j,0} C_i \frac{\partial C_j}{\partial L_j} + \Delta_{i,0} C_j\Big(\frac{\partial C_i}{\partial L_i} - \frac{\delta_i C_i}{1+\delta_i L_i}\Big)\Big].$$

The order $3/4$ approximation leads to excellent accuracy.

- On might easily refine this approach by computing terms of higher order in stochastic Taylor's expansion. This leads, however, to more complex and computationally expensive formulas, and the benefit of using an asymptotic expansion disappears. The order $1/2$ approximation appears to offer the best performance versus accuracy profile.

References


Brigo, D., and Mercurio, F.: *Interest Rate Models - Theory and Practice*, Springer Verlag (2006).


Glasserman, P.: *Monte Carlo Methods in Financial Engineering*, Springer Verlag (2003).


Hagan, P., and Lesniewski, A.: LIBOR market model with SABR style stochastic volatility, working paper (2006).


Kloeden, P. E., and Platen, E.: *Numerical Solution of Stochastic Differential Equations*, Springer Verlag (1992).


Mercurio, F.: Joining the SABR and Libor models together, *Risk*, 2009.


Rebonato, R., McKay, K., and White, R.: *The SABR / LIBOR Market Model: Pricing, Calibration and Hedging for Complex Interest-Rate Derivatives*, Wiley (2009).