Interdisciplinary Studies of Complex Systems
No. 4 (2014) 5–30

Interest Rate Model Selection and Implementation for the Credit Risk Engines

Alexei Kondratyev[1]

Abstract. This paper discusses the interest rate model selection and implementation process for the Monte Carlo credit risk engines. Special attention is paid to the real world model calibration and simulation problems, including development of the robust calibration algorithms for the illiquid interest rate curves and handling of the negative interest rates implied by the forward FX rates for many EM currency pairs.

1 Introduction

Interest rates are the most fundamental building blocks of any Monte Carlo risk engine. All other asset classes and all pricing functions require interest rate curves as an input for discount factors calculation and PVing of future cashflows. Therefore, the choice of interest rate model (or models) is an important task. Once the interest rate model is implemented and embedded into simulation of all other asset classes it is very difficult to replace it unless the risk engine has a highly modular architecture, which, though a desirable system design requirement, is frequently traded off for the risk engine performance optimization.

We only look here at the *parametric models* − the models that can be fully described by a small number of well defined parameters with precise meaning. All the models that take the time series of historical returns as an input (in various shapes and forms) are deliberately excluded from the consideration. The reason for this is the necessity to simulate the interest rate curves over very long time intervals, typically up to the 30-year Monte Carlo tenor. Understanding of potential future scenarios and their driving factors is an integral part of the prudent risk management: only the parametric models have this explanatory power.

We consider five different interest rate models, from the most basic 1-factor models to the now industry standard Libor Market Model. We analyse their relative advantages and drawbacks and discuss the areas of applicability. The derivation of all the model equations and their solutions is covered in Appendix A.

[1] Risk Analytics - Structured Credit Trading, Standard Chartered Bank
alexei.kondratyev@sc.com
The opinions expressed in this paper are those of the author and do not necessarily reflect the views and policies of Standard Chartered Bank.

These models are by far the most widely used in practice and each has its own strengths and weaknesses. There is no single interest rate model that would satisfy all the requirements specified by the risk engine users and developers. Ideally, there should be a pool of implemented and ready to be used models to meet a variety of requirements.

We present here full methodological specifications for two of them, which we believe maximizes the risk engine users utility: 1-factor Hull-White and 2-factor HJM.

2 Model selection

2.1 Ho-Lee model

Historically, the Ho-Lee (HL) model [2, 8, 9] was the first *no-arbitrage model* in the sense that the model is fully consistent with the current term structure of interest rates. The model can be formulated as a stochastic differential equation for the short interest rate (see Appendix A, equations (A.6), (A.7))

$$dr(t) = \theta(t)dt + \sigma W(t) \tag{2.1}$$

with

$$\theta(t) = \dot{f}(0,t) + \sigma^2 t\,,$$

where

σ	is the volatility of the short interest rate;
$f(0,t)$	is the instantaneous forward rate observed at time zero for the period of time $[t, t+dt]$;
$\dot{f}(0,t)$	is the partial derivative of $f(0,t)$ with respect to t;
$W(t)$	is the Brownian motion under the risk neutral measure.

Note that σ is the absolute volatility, i.e. it is a standard deviation of the short interest rate and not a standard deviation of the log-returns of the short interest rate.

The Ho-Lee model is *analytically tractable* in the sense that there are analytical expressions for zero coupon bonds (discount factors) of arbitrary tenor (see Appendix A, equations (A.8), (A.9)).

The model has a high *prediction power* $-$ it presribes a particular type of interest rate dynamics and cannot be calibrated to match the arbitrary volatility structure of the traded instruments such caps and floors.

The Ho-Lee model is also *Markovian.* In practical terms the Markov property means that the evolution of the interes rate curve in the future depends only on the current state and is independent of the path the interest rate curve followed in the past in order to arrive to the current state. This is a useful property that allows us to optimize the Monte Carlo simulation algorithm.

From equation (2.1) we see that the interest rates can become negative, which is a common feature of all Gaussian models. Another drawback is the fact that the interest rates tend to *blow up* $-$ due to the $\sigma^2 t$ term in the drift of equation (2.1) (quadratic term in equation (A.6)), the short interest rate $r(t)$ tends to infinity as $t \to \infty$ with probability 1.

The overall interst rates dynamics is limited to the parallel shifts of the interest rate curve. This a significant limitation if the model is meant to be used for the risk management of the curve trades. On the other hand, the model is easy to calibrate and implement and the generated scenarios are quite intuitive.

2.2 Hull-White model

Similar to the Ho-Lee model, the Hull-White (HW) model [2, 9] can be formulated as a model of the short interest rate (see Appendix A, equations (A.12), (A.14))

$$dr(t) = \kappa\left(\theta(t) - r(t)\right)dt + \sigma dW(t) \tag{2.2}$$

with

$$\theta(t) = f(0,t) + \frac{1}{\kappa}\dot{f}(0,t) + \frac{1}{2}\frac{\sigma^2}{\kappa^2}\left(1 - e^{-2\kappa t}\right),$$

where

σ	is the volatility of the short interest rate;
$f(0,t)$	is the instantaneous forward rate observed at time zero for the period of time $[t, t+dt]$;
$\dot{f}(0,t)$	is the partial derivative of $f(0,t)$ with respect to t;
κ	is a positive constant;
$W(t)$	is the Brownian motion under the risk neutral measure.

Parameter κ is responsible for the exponential decay of the volatility over time and is called the *rate of mean reversion* due to the non-trivial asymptotic probability distributon of $r(t)$ for large t.

The Hull-White model shares many properties with the Ho-Lee model, which is not surprising given that both models belong to the same class of 1-factor HJM models. The model is analytically tractable (see Appendix A, equation (A.15)), Markovian, has high predicting power, fits the current term structure of interest rates by construction (no-arbitrage), allows negative interest rates and is easy to calibrate and implement.

The biggest difference is that in the Hull-White model the short interest rate has a mean and a variance that are bounded independently of t. Equation (A.12) shows that the short interest rate $r(t)$ behaves asymptotically for $t \gg 1$ as

$$r(t) \approx f(0,t) + \sigma \int_0^t e^{-\kappa(t-s)}dW(s) + \frac{1}{2}\frac{\sigma^2}{\kappa^2}.$$

The right hand side of this equation is a Gaussian random variable with mean

$$\mu_\infty(t) = f(0,t) + \frac{1}{2}\frac{\sigma^2}{\kappa^2}$$

and variance

$$\sigma_\infty^2(t) = \sigma^2 \mathrm{E}\left[\left(\int_0^t e^{-\kappa(t-s)}dW(s)\right)^2\right] = \sigma^2 \int_0^t e^{-2\kappa(t-s)}ds \approx \frac{\sigma^2}{2\kappa}$$

for $t \gg 1$. As κ increases, the mean of the short interest rate distribution gets closer to the spot forward rate and the variance of the short interest rate decreases as $1/\kappa$. In other words we observe strong pull towards the forecast rate $f(0,t)$. This makes the Hull-White model superior to the Ho-Lee model and explains its enduring popularity. Another useful property of the Hull-White model is that it can be easily calibrated to swaption prices − a handy property when the model must be calibrated in the risk-neutral environment for the CVA calculation purposes.

2.3 Cox-Ingersoll-Ross model

Historically, the Cox-Ingersoll-Ross (CIR) model [5, 9] preceded the no-arbitrage models, such as the Ho-Lee and the Hull-White models. The CIR model belongs to the class of *equilibrium* models where the term structure of the interest rates is implied by the dynamics of the short interest rate. This, however, gives this model enormous explaining power. The model can be formulated as a model of the short interest rate

$$dr(t) = \kappa \left(\theta - r(t)\right) dt + \sigma \sqrt{r(t)} dW(t) \,, \tag{2.3}$$

where

σ	is the volatility coefficient;
κ	is a positive constant;
θ	is a non-negative constant;
$W(t)$	is the Brownian motion under the risk neutral measure.

The instantaneous short interest rate volatility is proportional to the square root of the short interest rate. This guarantees that the interest rates stay non-negative due to suppressed volatility as the simulated short interest rate tends to zero. But the model has the problem with the existence of the *absorbing state*. For the interest rate process (2.3) to remain positive a Feller condition [6] must be satisfied

$$2\kappa\theta > \sigma^2 \,,$$

in which case the upward drift in (2.3) is sufficiently large for $r(t)$ to never exactly reach zero.

The CIR model is Markovian and analytically tractable. However, the inability to fit the current interest rate curve makes this model less than ideal for the interest rate scenario simulation. At the same time the CIR model is a popular choice when it comes to the simulation of default intensities. The reason for this is that in many cases we have only a few liquid CDS tenors (and frequently only the 5-year CDS spread is available, if at all) and have to deal with very steep piece-wise constant default intensity curves. Since the Gaussian models (HL, HW) would experience significant calibration problems too and would have generated substantial number of negative rates in high volatility, low credit spread environment − the CIR model has a competitive advantage as a simple robust model that does not generate negative rates.

2.4 Heath-Jarrow-Morton model

The N-factor Heath-Jarrow-Morton (HJM) model [2, 7, 13] has the form (see Appendix A, equation (A.3))

$$df(t,T) = \sum_{i=1}^{N} \dot{\sigma}_i(t,T)dW_i(t) + \left(\sum_{i=1}^{N} \dot{\sigma}_i(t,T) \int_t^T \dot{\sigma}_i(t,s)ds \right) dt \,. \tag{2.4}$$

where

$f(t,T)$	is the instantaneous forward rate observed at time t for the period of time $[T, T+dT], t < T$;
$\sigma_i(t,T)$	are components of zero coupon bond volatilities;
$\dot{\sigma}_i(t,T)$	is derivative of $\sigma_i(t,T)$ with respect to T ;
$W_i(t)$	are independent Brownian motions under the risk neutral measure.

We interpret $\sigma_i(t,T), i = 1, 2, \ldots$ as components of the zero coupon bond volatilities that represent the first, the second, ... principal components of the interest rate curve dynamics. The Pricipal Component Analysis (PCA) is one of the most popular methods of the HJM model calibration for the risk factors simulation within the risk engine environment. The PCA would suggest the optimal number of pricipal components we should incorporate into the model and would also dictate a particular functional form of each $\sigma_i(t,T)$.

Using the HJM equation (2.4) we can write for the short interest rate, $r(t) = f(t,t)$,

$$r(t) = f(0,t) + \int_0^t \sum_{i=1}^{N} \dot{\sigma}_i(s,t)dW_i(s) + \int_0^t \left(\sum_{i=1}^{N} \dot{\sigma}_i(s,t) \int_s^t \dot{\sigma}_i(s,u)du \right) ds \,.$$

This equation shows that $r(t)$ is an Itô process under the risk neutral measure, satisfying the stochastic differential equation

$$dr(t) = \mu_r(t)dt + \sigma_r(t)dZ(t) \,,$$

where

$$\begin{aligned}
\mu_r(t) &= \dot{f}(0,t) + \sum_{i=1}^{N} \int_0^t \left((\dot{\sigma}_i(s,t))^2 + \ddot{\sigma}_i(s,t) \int_s^t \dot{\sigma}_i(s,u)du \right) ds \\
&+ \sum_{i=1}^{N} \int_0^t \ddot{\sigma}_i(s,t)dW_i(s) \,, \\
\sigma_r(t) &= \sqrt{\sum_{i=1}^{N} (\dot{\sigma}(t,t))^2} \,, \\
dZ(t) &= \frac{1}{\sigma_r(t)} \sum_{i=1}^{N} \dot{\sigma}(t,t)dW_i(t) \,.
\end{aligned}$$

The last term in expression for $\mu_r(t)$ shows that the drift of the short interest rate is path dependent. This makes this model non-Markovian in general, though a craftful selection of the volatility functions $\sigma_i(t,T)$ can make it Markovian in some special cases − typically in the 1-factor model framework as is the case with the Ho-Lee and the Hull-White models.

Not being Markovian means that the implementation of the multi-factor HJM in the risk engine Monte Carlo simulation is computationally expensive. Another challenge is the discretization of the continuous time. The instantaneous forward rates are a mathematical abstraction: they cannot be directly observed in the market and they are not included into the payoffs of traded derivatives. In real world we only deal with their integrals − the term rates. The necessity to i) discretize the continuous time and ii) perform numerical integration adds to the computational complexity. On the positive side, the calibration through the PCA is straightforward as we shall see in the following sections.

But the most important advantage of the multi-factor HJM model is the variety of scenarios it is able to generate. The model is quite powerful even in the 2-factor setup, where the simulated changes to the steepness of the interest rate curve capture the bulk of the curve trades risk.

2.5 Libor market model

The N-factor Libor Market Model (LMM) [1, 2, 13] has the form (see Appendix A, equation (A.23))

$$\begin{aligned} \frac{dF_n(t)}{F_n(t)} &= \sum_{i=1}^{N} \widetilde{\sigma}_n^i(t) dW_i(t) \\ &+ \sum_{i=1}^{N} \widetilde{\sigma}_n^i(t) \left(\sum_{m=2}^{n} \widetilde{\sigma}_m^i(t) \left(\frac{F_m(t)\Delta T_m}{1+F_m(t)\Delta T_m} \right) \right) dt\,, \end{aligned} \tag{2.5}$$

where

$F_n(t)$	is the forward term rate observed at time t for the period of time $[T_n, T_{n+1}], t \le T_n < T_{n+1}$;
ΔT_n	is the time interval between the two consecutive forward tenors, $\Delta T_n \equiv T_{n+1} - T_n$;
$\widetilde{\sigma}_n^i(t)$	are components of the vector of forward term rate volatilities;
$W_i(t)$	are independent Brownian motions under the T_2-forward measure.

The LMM resolves all major issues with the HJM model, such as the need to simulate the unobservable instantaneous forward rates, potentially explosive growth of the interest rates, and the problem of discretization of continuos time. In particular, one can observe that the LMM drift in (2.5) resembles a discretization of the HJM drift in (2.4) where the integral over continuous time

$$\dot{\sigma}_i(t,T) \int_t^T \dot{\sigma}_i(t,s) ds$$

is approximated by a discrete sum

$$\widetilde{\sigma}_n^i(t) \sum_{m=2}^{n} \widetilde{\sigma}_m^i(t) \left(\frac{F_m(t)\Delta T_m}{1+F_m(t)\Delta T_m} \right) .$$

The modelling of the finite number of directly observable forward term rates is a big advantage of the LMM. On the other hand side, if we want to model M Libor rates the model requires the knowledge of $M \times N$ volatilities. These volatilities can be easily derived from, e.g., caplet volatilities but what should we do with the interest rate curves that do not have a liquid option market? In this respect the LMM compares unfavourably to the HJM model where volatilities can be specified as functions of a handful of parameters. This makes the HJM model easier to calibrate in many cases and gives it higher explaining power.

Similar to the HJM model, the LMM is non-Markovian. This is a common feature of the models with log-normal distributions of state variables. As a result the implementation of the LMM in the risk engine Monte Carlo environment is a non-trivial task due to both computational complexity and the large number of model parameters that need to be properly calibrated, or at least initialised in a consistent manner.

2.6 Which model to choose

The following table summarises the top 10 most important properties of the interest rate models discussed here:

Model characteristic	HL	HW	CIR	HJM	LMM
Rich curve dynamics (multi-factor)	No	No	No	Yes	Yes
Analytically tractable	Yes	Yes	Yes	No	No
Fits the current curve (no-arb)	Yes	Yes	No	Yes	Yes
High prediction power	Yes	Yes	Yes	Yes	No
Easy to calibrate	Yes	Yes	Yes	Yes	No
Easy to implement	Yes	Yes	Yes	No	No
Markov property	Yes	Yes	Yes	No	No
Existence of absorbing state	No	No	Yes	No	No
Can simulate negative rates	Yes	Yes	No	Yes	No
Simulated rates can blow up	Yes	No	No	Yes	No

This table can be used to facilitate the decision making process of model selection. By assigning weights to the various model characteristics, such as simplicity (easy to calibrate, easy to implement), the richness of the simulated scenarios (1-factor vs. multi-factor), etc, it is possible to rank them according to the idividual preferences.

For example, a risk engine developer who is concerned with the risk management of relatively small directional counterparty portfolios comprised of vanilla derivatives may have the following weights for model characteristics on scale from 1 to 20, where 20 is the most and 1 is the least desirable model feature:

Model characteristic	Weight for "Yes"	Weight for "No"
Rich curve dynamics (multi-factor)	2	0
Analytically tractable	10	0
Fits the current curve (no-arb)	12	0
High prediction power	8	0
Easy to calibrate	20	0
Easy to implement	15	0
Markov property	1	0
Existence of absorbing state	0	6
Can simulate negative rates	0	4
Simulated rates can blow up	0	3

As a result, the 1-factor models appear to be well ahead of the multi-factor HJM and LMM with Hull-White model having the highest score – this model is loved for a reason!

Model characteristic	HL	HW	CIR	HJM	LMM
Rich curve dynamics (multi-factor)	0	0	0	2	2
Analytically tractable	10	10	10	0	0
Fits the current curve (no-arb)	12	12	0	12	12
High prediction power	8	8	8	8	0
Easy to calibrate	20	20	20	20	0
Easy to implement	15	15	15	0	0
Markov property	1	1	1	0	0
Existence of absorbing state	6	6	0	6	6
Can simulate negative rates	0	0	4	0	4
Simulated rates can blow up	0	3	3	0	3
Total score	72	75	61	48	27

Note that the very desirable Markov property has received the lowest weight due to the fact that its value is already taken into account by the highest weight assigned to the ease of implementation. Also, the weight distribution is non-linear that reflects significant differences in the relative importance of various model features.

Alternatively, a risk engine developer who is more concerned with the quality of future market scenarios and is not afraid of the associated development, calibration and maintenance costs may have a different set of preferences:

Model characteristic	Weight for "Yes"	Weight for "No"
Rich curve dynamics (multi-factor)	20	0
Analytically tractable	2	0
Fits the current curve (no-arb)	15	0
High prediction power	12	0
Easy to calibrate	8	0
Easy to implement	6	0
Markov property	1	0
Existence of absorbing state	0	10
Can simulate negative rates	0	4
Simulated rates can blow up	0	3

In this case the multi-factor models catch up with the 1-factor models and HJM comes ahead of LMM:

Model characteristic	HL	HW	CIR	HJM	LMM
Rich curve dynamics (multi-factor)	0	0	0	20	20
Analytically tractable	2	2	2	0	0
Fits the current curve (no-arb)	15	15	0	15	15
High prediction power	12	12	12	12	0
Easy to calibrate	8	8	8	8	0
Easy to implement	6	6	6	0	0
Markov property	1	1	1	0	0
Existence of absorbing state	10	10	0	10	10
Can simulate negative rates	0	0	4	0	4
Simulated rates can blow up	0	3	3	0	3
Total score	54	57	36	65	52

It is important to stress that we are looking at the interest rate models from the risk factors simulation point of view rather than trying to evaluate them from the ineterst rate derivatives pricing perspective. The model rankings presented here illustrate the difficulties a risk engine developer is facing while selecting an interest rate model for the Monte Carlo risk engine. A sophisticated model able to give an accurate (non-arbitragable) price of an exotic interest rate derivative may perform poorly when is calibrated to the ineterst rate curve in the illiquid market.

3 Implementation of 1-factor Hull-White model

In the most general case, the Hull-White model specifies the following stochastic process for the short interest rate (see equations (A.14) and (2.2))

$$dr(t) = \kappa \left(\theta(t) - r(t) - \frac{1}{\kappa}\rho(t)\sigma\sigma_x(t) \right) dt + \sigma dW(t) \tag{3.1}$$

with

$$\theta(t) = f(0,t) + \frac{1}{\kappa}\dot{f}(0,t) + \frac{1}{2}\frac{\sigma^2}{\kappa^2}\left(1 - e^{-2\kappa t}\right) ,$$

where

σ	is the volatility of the short interest rate;
$f(0,t)$	is the instantaneous forward rate observed at time zero for the period of time $[t, t+dt]$;
$\dot{f}(0,t)$	is the partial derivative of $f(0,t)$ with respect to t;
κ	is the rate of *mean reversion* ($\kappa > 0$);
$W(t)$	is the Brownian motion under the risk neutral measure;
$\sigma_x(t)$	is the instantaneous volatility of the spot FX rate (FX rate is defined as the number of base currency units per one unit of the interest rate curve currency; for the base currency interest rate curves $\sigma_x(t) \equiv 0$);
$\rho(t)$	is the instantaneous correlation between the spot FX rate process and the Brownian motion that drives the short interest rate process.

The drift term

$$-\rho(t)\sigma\sigma_x(t)$$

is a quanto adjustment for the interest rates simulated in the non-base currencies.

3.1 Simulation scheme

As a rule, the best simulation results are achieved when we simulate a solution to the stochastic differential equation instead of the risk factor increments specified directly by the equation itself [14]. This is especially true when we simulate risk factors over long period of time with large Monte Carlo time steps that inevitably introduce significant discretization errors. Therefore, our first task is to find a solution to equation (3.1). It follows from (A.12) that the solution of this equation has the following form

$$\begin{aligned} r(t) &= f(0,t) + \frac{1}{2}\frac{\sigma^2}{\kappa^2}\left(1 - e^{-\kappa t}\right)^2 - \sigma \int_0^t e^{-\kappa(t-s)}\rho(s)\sigma_x(s)ds \\ &+ \sigma \int_0^t e^{-\kappa(t-s)} dW(s) . \end{aligned}$$

The Monte Carlo simulation is performed on the discrete set of tenor points. Our aim is to specify the dynamics of the short interest rates between any two consecutive Monte Carlo tenors. Let t_n and t_{n+1} be two consecutive Monte Carlo tenors and let us denote $\Delta t_{n+1} = t_{n+1} - t_n$. Then we can write

$$\begin{aligned} r(t_n) &= f(0,t_n) + \frac{1}{2}\frac{\sigma^2}{\kappa^2}\left(1 - e^{-\kappa t_n}\right)^2 - \sigma \int_0^{t_n} e^{-\kappa(t_n - s)}\rho(s)\sigma_x(s)ds \\ &+ \sigma \int_0^{t_n} e^{-\kappa(t_n - s)} dW(s) \end{aligned} \tag{3.2}$$

and

$$\begin{aligned} r(t_{n+1}) &= f(0,t_{n+1}) + \frac{1}{2}\frac{\sigma^2}{\kappa^2}\left(1 - e^{-\kappa t_{n+1}}\right)^2 \\ &- \sigma \int_0^{t_{n+1}} e^{-\kappa(t_{n+1} - s)}\rho(s)\sigma_x(s)ds \\ &+ \sigma \int_0^{t_{n+1}} e^{-\kappa(t_{n+1} - s)} dW(s)\,. \end{aligned} \tag{3.3}$$

From (3.2) and (3.3) we obtain expression for $r(t_{n+1})$ as a function of $r(t_n)$ and the standard Normal random variable ε that drives the short interest rate increment on time interval $[t_n, t_{n+1}]$

$$\begin{aligned} r(t_{n+1}) &= f(0,t_{n+1}) + (r(t_n) - f(0,t_n))e^{-\kappa \Delta t_{n+1}} \\ &+ \left(1 - e^{-\kappa \Delta t_{n+1}}\right)\left(\frac{1}{2}\frac{\sigma^2}{\kappa^2}\left(1 - e^{-\kappa(t_n + t_{n+1})}\right) - \frac{1}{\kappa}\rho(t_n)\sigma\sigma_x(t_n)\right) \\ &+ \sigma\sqrt{\frac{1 - e^{-2\kappa\Delta t_{n+1}}}{2\kappa}}\,\varepsilon\,, \end{aligned} \tag{3.4}$$

where we assumed that ρ and σ_x are piece-wise constant functions of time on the Monte Carlo intervals: $\rho(t) = \rho(t_n)$, $\sigma_x(t) = \sigma_x(t_n)$, $t \in [t_n, t_{n+1}]$. The zero interest rate curve (or, equivalently, discount factors) can be built from the simulated short interest rate (3.4) and expression (A.15) for the zero coupon bonds.

The simulated short interest rate should be subject to flooring at each Monte Carlo tenor to avoid simulation of negative interest rates. Strictly speaking, the flooring of the short interest rate does not guarantee that all zero interest rates are positive. However, in the vast majority of cases the number of Monte Carlo paths that can lead to the negative zero interest rates after the flooring has been applied is negligibly small and does not have a material impact on the zero interest rates distributions.

3.2 Calibration

As always, there is a choice to be made between the risk neutral calibration (e.g. matching the market prices of a range of the traded instruments) and the historical calibration (e.g. matching the moments of the historical distributions of the relevant risk factors). The high level of analytical tractability of the Hull-White model is very helpful in designing the model calibration algorithms in both cases. Here we present an algorithm that aims to calibrate parameters

κ and σ by matching the historical volatilities of zero interest rates (yields of zero coupon bonds).

We start by remembering that the Hull-White model is a special case of the 1-factor HJM model with the zero coupon bond volatility, $\sigma(t,T)$, specified as

$$\dot{\sigma}(t,T) = \sigma e^{-\kappa(T-t)} ,$$

where $\sigma(t,T)$ is a volatility of the zero coupon bond at time t with maturity at time T and $\dot{\sigma}(t,T)$ denotes the first derivative with respect to T. Thus, we have the following expression for the zero coupon bond volatility

$$\sigma(t,T) = \int_t^T \dot{\sigma}(t,s)ds = \frac{\sigma}{\kappa}\left(1 - e^{-\kappa(T-t)}\right) .$$

The historical zero coupon bond volatility can be estimated directly from the historical time series of zero interest rates. Let $\overline{\sigma}(T_1)$ and $\overline{\sigma}(T_2)$ denote the historical volatilities of T_1- and T_2-year zero coupon bonds respectively. Then we can write a system of two equations for two unknowns

$$\begin{aligned}\overline{\sigma}(T_1) &= \frac{\sigma}{\kappa}\left(1 - e^{-\kappa T_1}\right) ,\\ \overline{\sigma}(T_2) &= \frac{\sigma}{\kappa}\left(1 - e^{-\kappa T_2}\right) .\end{aligned}$$

Equation for κ is now

$$\frac{\overline{\sigma}(T_1)}{\overline{\sigma}(T_2)} = \frac{1 - e^{-\kappa T_1}}{1 - e^{-\kappa T_2}}$$

and can be solved numerically. Once the value of κ is found, the value of σ is given by expression

$$\sigma = \frac{\kappa \overline{\sigma}(T_2)}{1 - e^{-\kappa T_2}} .$$

In many cases the best calibration results are achieved for $T_1 = 1$ year and $T_2 = 10$ years.

4 Implementation of 2-factor HJM model

The 2-factor HJM model we want to use for the interest rate curves simulation has the following form (see equations (A.3) and (2.4) with $N = 2$)

$$\begin{aligned} df(t,T) &= \sum_{i=1}^{2} \dot{\sigma}_i(t,T) dW_i(t) \\ &+ \sum_{i=1}^{2} \left(\dot{\sigma}_i(t,T) \int_t^T \dot{\sigma}_i(t,s)ds - \dot{\sigma}_i(t,T)\sigma_x(t)\rho_{x,i}(t) \right) dt , \end{aligned} \tag{4.1}$$

where

$f(t,T)$ is the instantaneous forward rate observed at time t for the period of time $[T, T + dT], t < T$;

$W_i(t)$ are the Brownian motions under the base currency risk neutral measure;

$\sigma_x(t)$ is the instantaneous volatility of the spot FX rate (FX rate is defined as the number of base currency units per one unit of the interest rate curve currency; for the base currency interest rate curves $\sigma_x(t) \equiv 0$);

$\rho_{x,i}(t)$ are the instantaneous correlations between the spot FX rate and the Brownian motions of the corresponding interest rate diffusion components.

The volatility functions (as suggested, e.g., in [2]) are

$$\sigma_1(t,T) = a(T-t)\,, \tag{4.2}$$

$$\sigma_2(t,T) = b(T-t)\left(1-2e^{-\kappa(T-t)}\right)\,, \tag{4.3}$$

where a, b and κ are constants. We require κ to be strictly positive while a and b can take any real values. Parameters a and b are the volatility coefficients that determine the magnitude of the simulated shocks to the interest rate curve. Parameter κ is responsible for the shape of the volatility term structure.

The last term in the right hand side of equation (4.1) is the quanto adjustment for the non-base currency interest rate curves. The necessity of having a quanto adjustment while simulating the non-base currency risk factors under the base currency risk measure is discussed in [11].

4.1 Simulation scheme

In Section 2.4 we mentioned that $\sigma_i(t,T), i = 1,2$ can be interpreted as components of the zero coupon bond volatilities that represent the first and the second principal components of the interest rate curve dynamics. We note that the volatility function (4.2) has the first derivative

$$\dot{\sigma}_1(t,T) = a = const$$

similar to the volatility specification for the Ho-Lee model. On the other hand, the volatility function (4.3) has the first derivative with respect to T that changes sign at some tenor point, implying that the simulated shock would move the short end and the long end of the curve into opposite directions. Therefore, the choice of the volatility functions (4.2) and (4.3) is consistent with the first principal component being responsible for the parallel shift of the interest rate curve and the second principal component being responsible for the change in curve's steepness.

This means that when it comes to the estimation of correlation between the forward interest rates and all the other risk factors, including FX rates, we can say that the first pricipal component holds almost all available information about the correlation structure. In practical terms it means that $W_1(t)$ in (4.1) should be simulated as a Brownian motion correlated with the corresponding FX rate with correlation $\rho_{x,1}(t) \equiv \rho(t)$ while $W_2(t)$ in (4.1) should have $\rho_{x,2}(t) \equiv 0$.

It has been already mentioned that, whenever possible, we should simulate risk factors as solutions to stochastic differential equations, rather than using these equations directly. Thus, we first write equation (4.1) in the integral form

$$
\begin{aligned}
f(t,T) &= f(0,T) + \sum_{i=1}^{2} \int_0^t \dot{\sigma}_i(s,T) dW_i(s) \\
&+ \sum_{i=1}^{2} \int_0^t \left(\dot{\sigma}_i(s,T) \int_s^T \dot{\sigma}_i(s,u) du - \dot{\sigma}_i(s,T)\sigma_x(s)\rho_{x,i}(s) \right) ds \,.
\end{aligned}
$$

Since both $\sigma_1(t,T)$ and $\sigma_2(t,T)$ are functions of $T-t$ we have

$$
\begin{aligned}
f(t,T) &= f(0,T) + \sum_{i=1}^{2} \int_0^t \dot{\sigma}_i(s,T) dW_i(s) \\
&+ \sum_{i=1}^{2} \int_0^t \left(\dot{\sigma}_i(s,T)\sigma_i(s,T) - \dot{\sigma}_i(s,T)\sigma_x(s)\rho_{x,i}(s) \right) ds \,. \qquad (4.4)
\end{aligned}
$$

Now, substituting (4.2) and (4.3) into (4.4) and assuming zero correlation between $W_2(t)$ and the corresponding FX rate process, we find the final expression for $f(t,T)$

$$
\begin{aligned}
f(t,T) &= f(0,T) + a \int_0^t dW_1(s) + a^2 \int_0^t (T-s) ds \\
&+ b \int_0^t \left(1 - 2e^{-\kappa(T-s)}(1-\kappa(T-s)) \right) dW_2(s) \\
&+ b^2 \int_0^t \left(1 - 2e^{-\kappa(T-s)}(1-\kappa(T-s)) \right) (T-s)(1-2e^{-\kappa(T-s)}) ds \\
&- a \int_0^t \sigma_x(s)\rho(s) ds \,. \qquad (4.5)
\end{aligned}
$$

The simulation procedure can be split into three steps. The first step is to calculate all spot instantaneous forward rates from the either linearly or cubic spline interpolated spot zero interest rate curve, which is typically defined on a relatively small set of tenors. A *zero* interest rate, $R(t,T)$, observed at time t for the period of time $[t,T]$ is the yield of a zero coupon bond with maturity at time T.

The second step is to run the Monte Carlo simulation using the instantaneous forward rate process (4.5) while monitoring the simulated instantaneous forward rates for potential negative numbers. A suitable global flooring can be applied, e.g. 5 basis points.

The third step is rebuilding of the corresponding zero interest rate curves that define the simulated interest rate scenarios and can be taken as an input by the pricing functions. The flooring of the instantaneous forward rates guarantees that the forward term rates and the zero interest rates are strictly positive.

To perform the first step we have to decide on what a suitable discretization scheme can be. The natural restriction on discretization of the forward curve is the granularity of Monte Carlo tenors: forward rates should be at least as granular as the Monte Carlo grid.

The instantaneous forward rate $f(t, T)$ has the meaning of an interest rate observed at time t for the period of time $[T, T + dT]$. The forward term rate $F(t, T_1, T_2)$ has the meaning of an interest rate observed at time t for the time interval $[T_1, T_2]$. Thus, the following discretization of $f(t, T)$ and $F(t, T_1, T_2)$ should be used for the simulation of interest rate curves on a given set of Monte Carlo tenors:

$$f(t, T) \to f(t_i, T_n)$$

and

$$F(t, T_1, T_2) \to F(t_i, T_m, T_n)\,,$$

where

t_i	is the Monte Carlo tenor;
T_n	is the forward curve tenor;
$f(t_i, T_n)$	is the interest rate observed at time t_i for the time interval $[T_{n-1}, T_n]$ with $t_i \le T_{n-1} < T_n$;
$F(t_i, T_m, T_n)$	is the interest rate observed at time t_i for the time interval $[T_m, T_n]$ with $t_i \le T_m < T_n$.

The second step follows logically from the chosen discretization scheme. In order to write expression (4.5) in the discrete form we introduce several new notations. Let $\Delta t_l = t_l - t_{l-1}$ be the Monte Carlo time step between the Monte Carlo tenors t_{l-1} and t_l, $l = 0, 1, 2, \ldots, L_{max}$ and $\Delta T_k = T_k - T_{k-1}$ be the time interval between the two consecutive forward curve tenors T_{k-1} and T_k, $k = 0, 1, 2, \ldots, K_{max}$. Further, let $\xi(l)$ be the value of the index of the forward curve tenor $T_{\xi(l)}$ that corresponds to the Monte Carlo tenor t_l: $T_{\xi(l)} = t_l$. Then the discrete version of the instantaneous forward rate process has the form

$$\begin{aligned}
f(t_i, T_n) &= f(0, T_n) + a\sum_{l=1}^{i}\sqrt{\Delta t_l}\,\varepsilon_l^1 + a^2\sum_{l=1}^{i}\sum_{k=\xi(l-1)+1}^{\xi(l)}(T_n - T_k)\Delta T_k \\
&+ b\sum_{l=1}^{i}\Big(1 - 2e^{-\kappa(T_n - t_l)}(1 - \kappa(T_n - t_l))\Big)\sqrt{\Delta t_l}\,\varepsilon_l^2 \\
&+ b^2\sum_{l=1}^{i}\sum_{k=\xi(l-1)+1}^{\xi(l)}\Big[\Big(1 - 2e^{-\kappa(T_n - T_k)}(1 - \kappa(T_n - T_k))\Big) \\
&\times (T_n - T_k)(1 - 2e^{-\kappa(T_n - T_k)})\Big]\Delta T_k \\
&- a\sum_{l=1}^{i}\sigma_x(t_l)\rho(t_l)\Delta t_l\,, \qquad (4.6)
\end{aligned}$$

where ε_l^1 and ε_l^2 are standard Normal random variables. Since a set of the Monte Carlo tenors is a subset of the forward curve tenors we have $\xi(l) \ge l$. If the forward curve tenor's grid coincides with the Monte Carlo simulation grid we have $\xi(l) \equiv l$ and the sum $\sum_{k=\xi(l-1)+1}^{\xi(l)}$ collapses to a single term.

The correlation with other risk factors is achieved through random variables ε_l^1 and ε_l^2: ε_l^2 is an independent standard Normal random variable and

ε_l^1 is a standard Normal random variable correlated with the random variables that drive the increments of all other risk factors.

The third step, rebuilding of the zero interest rate curve, is a straightforward task once all the forward rates have been simulated through (4.6). Forward term rates are calculated as

$$F(t_i, T_m, T_n) = \frac{1}{T_n - T_m} \sum_{k=m+1}^{n} f(t_i, T_k) \Delta T_k \,. \tag{4.7}$$

The zero interest rates, $R(t_i, T_n)$, are interest rates observed at time t_i for the period of time $[t_i, T_n]$. They are a subset of the forward term rates given by (4.7)

$$R(t_i, T_n) = F(t_i, t_i, T_n) \,.$$

4.2 Calibration

The calibration of the 2-factor HJM model specified by equations (4.1)-(4.3) is based on the Principal Component Analysis (PCA) of the zero coupon rates dynamics (see, e.g., [2, 15]).

We start with the historical time series of the zero interest rates. The PCA consists of the decomposing of the interest rate curve deformations into *principal components*, or elementary building blocks such as parallel shift, change in slope, twist, etc. The principal components can be ordered in terms of their relative importance, or explaining power. Thus, the first principal component (parallel shift) is typically able to explain up to 85-90% of the interest rate curve deformations, the second principal component (change in slope) can explain 8-9% of the interest rate curve deformations with all the other principal components (twists, double twists) being responsible for the balance. Because we want to do the PCA on the interest rate curve deformations we should carefully select the curve tenors. Ideally, the PCA should be performed on the interest rate curve specified on 6-10 tenors that include both the short end (3-month, 6-month tenors) and the long end (10-year, 30-year tenors).

Assume that we have M-year historical time series for N interest rate curve tenors. If we further assume that there are 250 daily interest rate curve observations in one year then we have a total of $250 \times M$ daily observations. Let $Y_1(t), Y_2(t), \ldots, Y_N(t)$, $t = 0, 1, 2, \ldots, 250 \times M$, be different daily observations of the zero interest rates of N different tenors over the last M years. For a given observation period Δt measured in days we consider the differences of successive zero interest rates at each curve tenor

$$D_{i,n} = Y_i((n+1)\Delta t) - Y_i(n\Delta t), \quad i = 1, 2, \ldots, N \,,$$

where n ranges over the number of consecutive periods of Δt days in the dataset: with $\Delta t = 1$ day, $n = 1, 2, \ldots, 250 \times M$; with $\Delta t = 5$ days, $n = 1, 2, \ldots, 50 \times M$.

Let C be $N \times N$ empirical covariance matrix

$$C_{ij} = \frac{1}{\widetilde{M}} \sum_{n=1}^{\widetilde{M}} (D_{i,n} - \overline{D}_i)(D_{j,n} - \overline{D}_j) \,,$$

where $\widetilde{M} = \mathrm{int}(M/\Delta t)$ is the number of periods and $\overline{D}_i$ is the sample mean of $D_{i,n}$, $n = 1, 2, \ldots, \widetilde{M}$. The PCA consists of computing the eigenvalues χ_k and eigenvectors $\vec{\nu}^k = (\nu_1^k, \nu_2^k, \ldots, \nu_N^k)$ of C and ranking them according to the magnitudes of the eigenvalues. From the standard matrix theory we have

$$C_{ij} = \sum_{k=1}^{N} \chi_k \nu_i^k \nu_j^k \,.$$

This is consistent with assuming that zero interest rate differences satisfy

$$D_{i,n} = \overline{D}_i + \sum_{k=1}^{N} \sqrt{\chi_k}\, \nu_i^k \varepsilon_{k,n} \,,$$

where $\varepsilon_{k,n}$ are independent standard Normal random variables. The normalized eigenvectors can be viewed as deformation eigenmodes of the zero interest rate curve and the eigenvalues describe the amplitude of each mode.

We interpret the HJM equation (4.1) as a parametric representation of two main principal components: a parallel shift of the whole curve driven by $\sigma_1(t,T)$ and a change in slope of the curve driven by $\sigma_2(t,T)$. Therefore, using reduced set of eigenvectors and the corresponding eigenvalues we obtain a 2-dimensional approximation to the statistics of zero interest rate differences

$$D_{i,n} = \overline{D}_i + \sum_{k=1}^{2} \sqrt{\chi_k}\, \nu_i^k \varepsilon_{k,n} \,.$$

Now we have everything we need to calibrate parameters a, b and κ. Note that the zero coupon bond yield observed at time t for the time interval $T - t$, $Y(t,T)$, has the stochastic differential

$$\begin{aligned} dY(t,T) &= d\left(-\frac{1}{T-t}\ln P(t,T)\right) \\ &= -\frac{1}{T-t}\sum_{k=1}^{2} \sigma_k(t,T)\,dW_k(t) + \mathrm{drift}\,, \end{aligned} \tag{4.8}$$

where $P(t,T)$ is the price of the zero coupon bond and $W(t)$ is a Brownian motion under the risk neutral measure. Let $V_k(T)$, $k = 1, 2$ be smooth interpolation functions associated with the principal components. The PCA suggests that the zero interest rate curve should satisfy the difference equations

$$\begin{aligned} \Delta Y(t,T) &= Y(t+\Delta t, T) - Y(t,T) \\ &= \sum_{k=1}^{2} V_k(T-t)\sqrt{\chi_k}\,\varepsilon_k(t) + \mathrm{mean}\,, \end{aligned} \tag{4.9}$$

where $\varepsilon_k(t)$ are independent standard Normal random variables. The use of $T-t$ in (4.9) is due to the fact that the PCA was done using relative maturities and not the fixed maturities as in HJM equation. By matching stochastic terms in (4.8) and (4.9) we get an estimator for the volatility structure of zero coupon bonds

$$\sigma_k(t,T) = (T-t)V_k(T-t)\sqrt{\frac{\chi_k}{\Delta t}}\,, \quad k = 1,2\,. \tag{4.10}$$

From (4.2), (4.3) and (4.10) we obtain estimators for a, b and κ. Parameter a can be computed directly as

$$a = \frac{1}{N}\sum_{i=1}^{N}\nu_i^1\sqrt{\frac{\chi_1}{\Delta t}}\,.$$

Parameters b and κ can be estimated through the least squares algorithm that minimizes the following expression

$$\sum_{i=1}^{N}\left(b(1-2e^{-\kappa T_i}) - \nu_i^2\sqrt{\frac{\chi_2}{\Delta t}}\right)\,,$$

where $b \in [b_{\min}, b_{\max}]$ and $\kappa \in [\kappa_{\min}, \kappa_{\max}]$. The choice of $b_{\min}$, $b_{\max}$, $\kappa_{\min}$ and $\kappa_{\max}$ can be either interest rate curve specific or they can be global system parameters. The typical values are: $b_{\min} = -a$, $b_{\max} = a$, $\kappa_{\min} = 0$, $\kappa_{\max} = 1$.

The correlation ρ in (4.5) is usually assumed to be constant. Its value can be estimated from the historical time series as a correlation between the zero interest rate of a particular liquid tenor (e.g., 1-year zero interest rate) and the corresponding spot FX rate. Because the increments of interest rates in the HJM model are normally distributed and the spot FX rates are usually simulated through the log-normal process, ρ is a correlation between the historical time series of *absolute* interest rate returns and *relative* FX rate returns (log-returns).

5 Practicalities of interest rates simulation

In practice the risk engine may have to support several interest rate curves for any given currency, for example 6-month and 1-year swap curves for USD. These interest rate curves must be simulated in a consistent manner taking into account high degree of correlations between them. The practical approach is to select a single *primary* interest rate curve for each currency and simulate all other *secondary* interest rate curves as either *ratios* or *spreads* over the corresponding primary curves.

Typically, the primary curves are constructed from a variety of traded fixed income instruments and interest rate derivatives. For example, the short end of the primary interest rate curve can be constructed from the money market deposits and futures rates, the long end can be constructed from the yields on the long-term government bonds, and the middle part can be constructed from the interest rate swap rates. These curves should be simulated directly through one of the interest rate models described in this chapter. The random variables that drive their dynamics (e.g., ε in (3.4) and ε^1 in (4.6)) should be correlated with the random variables that are generated to simulate all other risk factors in the risk engine. It can be done through the Cholesky decomposition of the correlation matrix of primary risk factors.

The secondary curves are usually constructed from a single instrument type. For example, an interest rate curve can be constructed from 6-month interest rate swaps or it can be build from the FX forward rates. In each such case the area of applicability of a secondary curve is quite limited. Why do we need them then? The reason for having several interest rate curves for the same currency is the fact that different curves can be used for forecasting (calculation of the forward rates) and discounting. For example, we would not be able to match the market FX forward prices had we used the 6-month swap rate curves for discounting (interest rate differential).

Why do we need both spread and ratio secondary curves? The answer to this question is the greater flexibility we have in modelling the future market scenarios and in matching the current terms structures of the secondary interest rate curves. Due to the flooring of the simulated primary interest rates they remain positive on all Monte Carlo paths. When the secondary curve is constructed by applying ratios to the simulated primary interest rates it is guaranteed to remainn positive as well, which is a useful property in many cases. At the same time we have situations when the secondary interest rates must be able to take negative values. For example, interest rate curves constructed from the FX forward rates can be negative. These curves should be simulated as spreads over the primary interest rate curves to preserve this property.

Following the notations introduced earlier, let $F(t, T_1, T_2)$ be a forward term rate observed at time t for the time interval $[T_1, T_2]$, $t \le T_1 < T_2$. Let $R(t, T) \equiv F(t, t, T)$ be a zero interest rate observed at time t for the time interval $[t, T]$, $t < T$. Let t_0 denotes the Monte Carlo simulation start date (both primary and secondary zero interest rate curves are known on this date), t denotes the Monte Carlo tenor (counting from t_0) and T denotes the interest rate curve tenor (counting from $t_0 + t$), i.e. t and T are time intervals, not dates. Then the spread zero interest rate, R_s, and the ratio zero interest rate, R_r, can be derived from the simulated primary zero interest rate, R_p, as

$$R_s(t, t{+}T) = R_p(t, t{+}T) + (F_s(t_0, t_0{+}t, t_0{+}t{+}T) - F_p(t_0, t_0{+}t, t_0{+}t{+}T)) \tag{5.1}$$

and

$$R_r(t, t+T) = R_p(t, t+T)\frac{F_r(t_0, t_0+t, t_0+t+T)}{F_p(t_0, t_0+t, t_0+t+T)}\,, \tag{5.2}$$

where the spread, ratio and primary forward term rates, F_s, F_r and F_p, are calculated from the corresponding spot zero interest rate curves

$$F(t_0, t_0+t, t_0+t+T) = \frac{R(t_0, t_0+t+T)(t+T) - R(t_0, t_0+t)t}{T}\,. \tag{5.3}$$

The spot zero interest rates $R(t_0, t_0 + t)$ and $R(t_0, t_0 + t + T)$ can in turn be obtained from the spot zero interest rate curves through the linear or cubic spline interpolation between the curve tenors.

All rates in formulae (5.1)-(5.3) are continuously compounded rates. The real curves, however, may have different compounding. It means that in order to calculate and apply the spreads and ratios we have to i) perform conversions from, e.g., annual compounding to continuous compounding for all curves before the calculation of spreads and ratios and running the Monte Carlo simulation, and ii) perform inverse conversion from continuous compounding to, e.g., annual compounding after the Monte Carlo simulation of the primary curves and the construction of the secondary curves. Transition from the zero rate with compounding frequency n (n compounding periods per year), $\bar{r}$, to the continuously compounded rate, r, is done via the following formulae

$$r = n \ln\left(1 + \bar{r}/n\right) , \qquad \bar{r} = n\left(e^{r/n} - 1\right) .$$

The secondary interest rate curves are practically 100% correlated with their primary curves. This may not be a desired property in all cases. When the aim is to capture the basis risk between the different curves for the same currency, there is a temptention to introduce some randomness into the spreads and ratios. Such attempts should be resisted due to significant calibration issues and the increasing complexity of the simulation process. Treating spreads and ratios as stochastic processes in their own rights can be justified only in exceptional cases. A better solution would be to introduce another primary curve (or curves) for such currencies. Due to the relatively small number of curves where the basis risk is considered to be a material issue, the overall size of the correlation matrix of primary risk factors should remain manageable.

A Derivation of IR model equations

The Heath-Jarrow-Morton Model

We start with an assumption that there are N risk factors modelled by N independent Brownian motions $W_i(t), t > 0, i = 1, \ldots, N$. We assume that zero coupon bond prices (discount factors) $P(t,T)$ satisfy, under the risk-neutral measure, the following equations

$$\frac{dP(t,T)}{P(t,T)} = \sum_{i=1}^{N} \sigma_i(t,T) dW_i(t) + \mu(t,T) dt \,. \tag{A.1}$$

Here $t \leq T$, $\mu(t,T)$ is instantaneous drift and $\sigma_i(t,T)$ are the volatility components of the zero coupon bond prices.

All zero coupon bonds are traded securities. Since zero coupon bonds do not pay dividends, the drift of a risk-neutral measure defined by (A.1) must satisfy

$$\mu(t,T) = r(t) \,,$$

where $r(t)$ is the short interest rate. This is the instantaneous interest rate observed at time t for the period of time $[t, t + dt]$. Therefore, the risk-neutral dynamics for zero coupon bonds (discount factors) should have the form

$$\frac{dP(t,T)}{P(t,T)} = \sum_{i=1}^{N} \sigma_i(t,T) dW_i(t) + r(t,T) dt \,.$$

Applying Itô lemma to $\ln P(t,T)$, we have

$$d\ln P(t,T) = \sum_{i=1}^{N} \sigma_i(t,T)dW_i(t) + r(t)dt - \frac{1}{2}\left(\sum_{i=1}^{N} \sigma_i(t,T)^2\right)dt\,. \qquad \text{(A.2)}$$

Differentiating equation (A.2) with respect to T and taking into account the definition of the instantaneous forward rates, $f(t,T)$,

$$f(t,T) = \lim_{\Delta t \to 0} \frac{1}{\Delta t}\left(\frac{\mathrm{E}\left[\exp\left(-\int_t^T r(s)ds\right)\right]}{\mathrm{E}\left[\exp\left(-\int_t^{T+\Delta t} r(s)ds\right)\right]} - 1\right) = -\frac{\partial \ln P(t,T)}{\partial T}$$

we obtain

$$-df(t,T) = \sum_{i=1}^{N} \dot{\sigma}_i(t,T)dW_i(t) - \left(\sum_{i=1}^{N} \dot{\sigma}_i(t,T)\sigma_i(t,T)\right)dt\,,$$

where $\dot{\sigma}_i(t,T)$ represents derivative of function $\sigma_i(t,T)$ with respect to T. The instantaneous forward rate, $f(t,T)$, has the meaning of an interest rate observed at time t for the period of time $[T, T+dT]$.

Using the fact that $-W_i(t)$ is also a Brownian motion, we obtain stochastic differential equation that describes the evolution of the forward rate curve under the risk-neutral measure

$$df(t,T) = \sum_{i=1}^{N} \dot{\sigma}_i(t,T)dW_i(t) + \left(\sum_{i=1}^{N} \dot{\sigma}_i(t,T)\int_t^T \dot{\sigma}_i(t,s)ds\right)dt\,. \qquad \text{(A.3)}$$

Equation (A.3) is known as the Heath-Jarrow-Morton (HJM) equation [2, 7, 13].

The Ho-Lee Model

The Ho-Lee (HL) model is the simplest special case of the generic HJM framework [2, 8, 9]. This is a 1-factor model ($N = 1$ in equation (A.3)) with a constant variance of forward rates

$$\dot{\sigma}_1(t,T) = \sigma = \mathrm{const}\,.$$

Equation (A.3) now becomes

$$df(t,T) = \sigma dW(t) + \left(\sigma\int_t^T \sigma ds\right)dt\,. \qquad \text{(A.4)}$$

Integrating (A.4) with respect to t we obtain a solution

$$f(t,T) = f(0,T) + \sigma W(t) + \sigma^2\left(Tt - \frac{1}{2}t^2\right)\,. \qquad \text{(A.5)}$$

Equation (A.5) is the Ho-Lee model. The model can be reformulated in the more conventional form as a model of the short interest rate, $r(t) = f(t,t)$,

$$r(t) = f(0,t) + \sigma W(t) + \frac{1}{2}\sigma^2 t^2 \,. \tag{A.6}$$

Differentiating (A.6) with respect to t and denoting $\theta(t) = \dot{f}(0,t) + \sigma^2 t$ we obtain HL equation in its classical textbook form

$$dr(t) = \theta(t)dt + \sigma W(t)\,, \tag{A.7}$$

where

$$\theta(t) = \dot{f}(0,t) + \sigma^2 t\,.$$

From (A.5) we can obtain expression for the zero coupon bond price, $P(t,T)$,

$$\begin{aligned} P(t,T) &= e^{-\int_t^T f(t,s)ds} \\ &= e^{-\int_t^T \left(f(0,s)+\sigma W(s)+\sigma^2\left(st-\frac{1}{2}t^2\right)\right)ds} \\ &= P(0,T)e^{-\int_t^T \sigma W(s)ds} e^{-\frac{1}{2}\sigma^2 T t(T-t)}\,. \end{aligned} \tag{A.8}$$

From (A.8), (A.6) and remembering that $P(0,t) = \exp\left(-\int_0^t f(0,s)ds\right)$ we get

$$P(t,T) = A(t,T)e^{-r(t)(T-t)} \tag{A.9}$$

with

$$\ln A(t,T) = \ln\frac{P(0,T)}{P(0,t)} + f(0,t)(T-t) - \frac{1}{2}\sigma^2 t(T-t)^2\,.$$

The Hull-White Model

The Hull-White (HW) model is another special case of the HJM model [2, 9]. This is a 1-factor HJM model where the volatility function is specified as

$$\dot{\sigma}_1(t,T) = \sigma e^{-\kappa(T-t)}\,,$$

where κ is a positive constant. HJM equation (A.3) becomes

$$df(t,T) = \sigma e^{-\kappa(T-t)}dW(t) + \sigma^2\left(e^{-\kappa(T-t)}\int_t^T e^{-\kappa(s-t)}ds\right)dt\,. \tag{A.10}$$

Integrating (A.10) with respect to t we obtain

$$\begin{aligned} f(t,T) &= f(0,T) + \sigma\int_0^t e^{-\kappa(T-s)}dW(s) \\ &- \frac{1}{2}\frac{\sigma^2}{\kappa^2}\left(1-e^{-\kappa(T-t)}\right)^2 + \frac{1}{2}\frac{\sigma^2}{\kappa^2}\left(1-e^{-\kappa T}\right)^2\,. \end{aligned} \tag{A.11}$$

The Hull-White model (A.11) can be simplified by introducing the state variable

$$Y(t) = \sigma \int_0^t e^{-\kappa(t-s)} dW(s)\,.$$

Then the second term in the righrt hand side of (A.11) can be expressed through $Y(t)$ as

$$\sigma \int_0^t e^{-\kappa(T-s)} dW(s) = e^{-\kappa(T-t)} Y(t)\,,$$

and the short interest rate formulation of the HW model has the following form

$$r(t) = f(0,t) + Y(t) + \frac{1}{2}\frac{\sigma^2}{\kappa^2}\left(1 - e^{-\kappa t}\right)^2 . \tag{A.12}$$

Noting that

$$\begin{aligned}
\frac{\partial Y(t)}{\partial t} &= \frac{\partial}{\partial t}\left(\sigma \int_0^t e^{-\kappa(t-s)} dW(s)\right) = \frac{\partial}{\partial t}\left(e^{-\kappa t}\sigma \int_0^t e^{\kappa s} dW(s)\right) \\
&= -\kappa e^{-\kappa t}\sigma \int_0^t e^{\kappa s} dW(s) + e^{-\kappa t}\frac{\partial}{\partial t}\left(\sigma \int_0^t e^{\kappa s} dW(s)\right) \\
&= -\kappa\left(\sigma \int_0^t e^{-\kappa(t-s)} dW(s)\right) + e^{-\kappa t}\sigma e^{\kappa t} W(t) \\
&= -\kappa Y(t) + \sigma W(t)
\end{aligned} \tag{A.13}$$

and substituting $Y(t)$ from (A.12) into (A.13) we obtain the textbook version of HW equation after differentiation of (A.12) with respect to t

$$\begin{aligned}
dr(t) &= \dot{f}(0,t)dt + dY(t) + \frac{\sigma^2}{\kappa}(e^{-\kappa t} - e^{-2\kappa t})dt \\
&= \left(\dot{f}(0,t) - \kappa Y(t) + \frac{\sigma^2}{\kappa}(e^{-\kappa t} - e^{-2\kappa t})\right) dt + \sigma dW(t) \\
&= \kappa\left(\frac{\dot{f}(0,t)}{\kappa} + f(0,t) - r(t) + \frac{1}{2}\frac{\sigma^2}{\kappa^2}(1 - e^{-\kappa t})^2 + \frac{\sigma^2}{\kappa^2}(e^{-\kappa t} - e^{-2\kappa t})\right) dt \\
&+ \sigma dW(t) \\
&= \kappa\left(\frac{\dot{f}(0,t)}{\kappa} + f(0,t) + \frac{1}{2}\frac{\sigma^2}{\kappa^2}(1 - e^{-2\kappa t}) - r(t)\right) dt + \sigma dW(t) \\
&= \kappa\left(\theta(t) - r(t)\right) dt + \sigma dW(t)\,,
\end{aligned} \tag{A.14}$$

where

$$\theta(t) = f(0,t) + \frac{1}{\kappa}\dot{f}(0,t) + \frac{1}{2}\frac{\sigma^2}{\kappa^2}\left(1 - e^{-2\kappa t}\right) .$$

Following the same computational logic as was the case with the Ho-Lee model, from equations (A.11) and (A.12) we find expression for the zero coupon bond price, $P(t,T)$,

$$P(t,T) = A(t,T)e^{-r(t)B(t,T)} \tag{A.15}$$

with

$$\begin{aligned}\ln A(t,T) &= \ln\frac{P(0,T)}{P(0,t)} + f(0,t)B(t,T) - \frac{1}{4\kappa^3}\sigma^2(e^{-\kappa T} - e^{-\kappa t})^2(e^{2\kappa t} - 1)\,,\\ B(t,T) &= \frac{1 - e^{-\kappa(T-t)}}{\kappa}\,.\end{aligned}$$

The Libor Market Model

The Libor Market Model (LMM) developed by Brace, Gatarek and Musiela [3], Jamshidian [10], Miltersen, Sandmann and Sondermann [12] is based on the HJM framework and can be considered as a "discretization" of the HJM model. An excellent LMM description can be found in [1, 2, 13].

Rather than modelling the unobservable theoretical instantaneous forward rates, $f(t,T)$, LMM models the forward term rates, $F(t,T_n,T_{n+1})$, that have the meaning of interest rates observed at time t for the period of time $[T_n,T_{n+1}], t \le T_n < T_{n+1}$. The starting point is not the log-normal process for the zero coupon bonds as in (A.1). Instead, the log-normal dynamics is assumed for the forward rates themselves.

We start by noting that the forward term rate $F_n(t) \equiv F(t,T_n,T_{n+1})$ is a martingale under the T_{n+1}-forward measure. From the expression for $F_n(t)$

$$F_n(t) = \frac{1}{\Delta T_n}\left(\frac{P(t,T_n)}{P(t,T_{n+1})} - 1\right), \qquad \Delta T_n = T_{n+1} - T_n\,, \tag{A.16}$$

we obtain after applying the Itô lemma [2]

$$dF_n(t) = \frac{1}{\Delta T_n}\frac{P(t,T_n)}{P(t,T_{n+1})}(\sigma(t,T_n) - \sigma(t,T_{n+1}))dW(t)\,, \tag{A.17}$$

where $W(t)$ is a Brownian motion under the T_{n+1}-forward measure and $\sigma(t,T)$ is a volatility of a zero coupon bond $P(t,T)$ as in (A.3). Substituting expression for $P(t,T_n)/P(t,T_{n+1})$ from (A.16) into (A.17) we find

$$\frac{dF_n(t)}{F_n(t)} = \frac{1 + F_n(t)\Delta T_n}{F_n(t)\Delta T_n}(\sigma(t,T_n) - \sigma(t,T_{n+1}))dW(t)\,. \tag{A.18}$$

Our aim is to model the forward term rates directly. In other words we want to work with the volatilities of forward term rates rather than with the volatilities of zero coupon bonds. We can formally write

$$\frac{dF_n(t)}{F_n(t)} = \widetilde{\sigma}_n(t)dW(t)\,, \tag{A.19}$$

where

$$\widetilde{\sigma}_n(t) = \frac{1 + F_n(t)\Delta T_n}{F_n(t)\Delta T_n}(\sigma(t,T_n) - \sigma(t,T_{n+1})) \tag{A.20}$$

is the volatility of the forward term rate $F_n(t)$. We take the forward term rate volatilities $\widetilde{\sigma}_n(t)$ as given. From (A.20) we see that there is a recursive relationship between the zero coupon bond and the forward term rate volatilities

$$\sigma(t,T_{n+1}) = \sigma(t,T_n) - \frac{F_n(t)\Delta T_n}{1 + F_n(t)\Delta T_n}\widetilde{\sigma}_n(t)\,. \tag{A.21}$$

Equations (A.18)-(A.20) specify the process for $F_n(t)$ under the T_{n+1}-forward measure. We want to derive a process for all n under the same fixed measure. The most convenient measure is a T_2-forward measure under which the $F_1(t)$ forward term rate is a martingale. Let us write

$$F_n(t) = \frac{1}{\Delta T_n}\left(\frac{P(t,T_n)}{P(t,T_{n+1})} - 1\right) = \frac{1}{\Delta T_n}\left(\frac{[P(t,T_n)/P(t,T_2)]}{[P(t,T_{n+1})/P(t,T_2)]} - 1\right).$$

After applying the Itô lemma [2] we have

$$\frac{d[P(t,T_n)/P(t,T_2)]}{[P(t,T_n)/P(t,T_2)]} = (\sigma(t,T_n) - \sigma(t,T_2))dW(t),$$

where $W(t)$ is a Brownian motion under the T_2-forward measure. Therefore, we arrive to the following expression

$$\begin{aligned} dF_n(t) &= \frac{1}{\Delta T_n}\frac{P(t,T_n)}{P(t,T_{n+1})}(\sigma(t,T_n) - \sigma(t,T_{n+1}))dW(t) \\ &+ \frac{1}{\Delta T_n}\frac{P(t,T_n)}{P(t,T_{n+1})}(\sigma(t,T_n) - \sigma(t,T_{n+1}))(\sigma(t,T_2) - \sigma(t,T_{n+1}))dt\,. \end{aligned} \tag{A.22}$$

From (A.21) and (A.22) we find

$$\begin{aligned} dF_n(t) &= F_n(t)\widetilde{\sigma}_n(t)dW(t) \\ &+ F_n(t)\left(\sum_{m=2}^{n}\widetilde{\sigma}_n(t)(\sigma(t,T_m) - \sigma(t,T_{m+1}))\right)dt \\ &= F_n(t)\widetilde{\sigma}_n(t)dW(t) \\ &+ F_n(t)\left(\sum_{m=2}^{n}\widetilde{\sigma}_n(t)\widetilde{\sigma}_m(t)\left(\frac{F_m(t)\Delta T_m}{1+F_m(t)\Delta T_m}\right)\right)dt\,. \end{aligned} \tag{A.23}$$

Equation (A.23) is the LMM equation for the forward term rates. Similar to HJM equation (A.3), the forward term rate can be driven by N independent Brownian motions. In this case $\widetilde{\sigma}_n(t)\cdot dW(t)$ and $\widetilde{\sigma}_n(t)\cdot\widetilde{\sigma}_m(t)$ are scalar products of N-dimensional vectors.

Acknowledgements

It is a pleasure to thank Brola Lordkipanidze and Alistair Thomas for interesting discussions and comments.

References

[1] L. B. G. Andersen and V. V. Piterbarg, *Interest Rate Modeling*, Volume 2, Atlantic Financial Press, 2010

[2] M. Avellaneda and P. Laurence, *Quantitative Modelling of Derivative Securities: From Theory to Practice*, Chapman & Hall/CRC, 2000

[3] A. Brace, D. Gatarek and M. Musiela, The Market Model of Interest Rate Dynamics, *Mathematical Finance*, **7**, 2, (1997) 127-55

[4] D. Brigo and F. Mercurio, *Interest Rate Models - Theory and Practice*, 2nd edition, Springer-Verlag, 2006

[5] J. C. Cox, J. E. Ingersoll and S. A. Ross, A Theory of the Term Struture of Interest Rates, *Econometrica*, **53**, (1985) 385-407

[6] W. Feller, Two Singular Diffusion Problems, *Annals of Mathematics*, **54**, (1951) 173-82

[7] D. Heath, R.A. Jarrow and A. Morton, Bond Pricing and the Term Structure of Interest Rates: A New Methodology, *Econometrica*, **60**, 1, (1992) 77-105

[8] T. S. Y. Ho and S.-B. Lee, Term Structure Movements and Pricing Interest Rate Contingent Claims, *Journal of Finance*, **41**, (1986) 1011-29

[9] J. C. Hull, *Options, Futures, and Other Derivatives*, 6th edition, Prentice Hall, 2006

[10] F. Jamshidian, LIBOR and Swap Market Models and Measures, *Finance and Stochastics*, **1**, (1997) 293-330

[11] A. Kondratyev, Naïve Monte Carlo, *The Capco Journal of Financial Transformation*, **37**, (2013) 111-116

[12] K. Miltersen, K. Sandmann and D. Sondermann, Closed Form Solutions for Term Structure Derivatives with LogNormal Interest Rates, *Journal of Finance*, **52**, 1, (1997) 409-30

[13] R. Rebonato, *Interest-Rate Option Models*, 2nd edition, John Wiley & Sons, 1998

[14] R. Seydel, *Tools for Computational Finance*, 2nd edition, Springer-Verlag, 2003

[15] R. S. Tsay, *Analysis of Financial Time Series*, 2nd edition, John Wiley & Sons, 2005