Models in Numerix Analytics

February 2015

Numerix LLC

Corporate Headquarters

99 Park Avenue

5th Floor

New York, NY 10016

Tel +1 212 302 2220

This document describes the quantitative models implemented in Numerix analytics products, including details of the stochastic processes and numerical methods available for pricing interest rate, cross-currency, credit, foreign exchange, commodity, inflation and equity derivatives and products. The purpose of this document is to complement the information provided in the user documentation for specific software modules without repeating the description of the user interface.

Contents

Chapter 1

Interest Rate Models

1.1 Short-Rate IR Models

1.1.1 Black Model

The Black model is a basic model of interest rates in which either forward or swap interest rates are treated as lognormal stochastic variables, (see, e.g., [22] or [36]). The model gives access to very rapid analytics for a limited set of instruments, including caps and European swaptions.

1.1.2 Hull-White 1-Factor

The Hull-White model has a normal short-rate process

$$dr(t) = (\theta(t) - \lambda(t)r(t))dt + \sigma(t)dW(t), \tag{1.1.1}$$

where

$r(t)$ is the short rate,
$\lambda(t)$ is the mean reversion,
$\theta(t)$ is a deterministic function introduced to match the initial discount curve,
$\sigma(t)$ is the short-rate volatility,
$dW(t)$ is a Brownian motion (in the risk-neutral measure).

This model is implemented using several methods, including proprietary trees, simulations and PDE solver methods.

The zero-coupon bond (ZCB) dynamics are determined from the short rate $r(t)$,

$$P_T(t) = \mathbb{E}_t \left[e^{-\int_t^T r(u)du} \right], \tag{1.1.2}$$

where $\mathbb{E}_t$ is the expectation at time t in the risk-neutral measure. An analytic closed-form solution exists for $P_T(t)$ ([22],[36]). The model gives rise to normal distributions for integral functionals of the short rate, which allows for an easy link with Black model pricing and efficient calibration.

1.1.3 Hull-White 2- and 3-Factor

For multifactor Hull-White, the short rate is modeled as a sum

$$r(t) = \sum_{i=1}^{d} x_i(t) + \alpha(t), \tag{1.1.3}$$

where $\{x_i(t)\}$ are Markov processes and $\alpha(t)$ is an additional deterministic function that is used to match the original discount curve.

Each auxiliary process obeys the Ornstein-Uhlenbeck equation

$$dx_i(t) = -\lambda_i x_i(t)dt + \sigma_i(t)dW_i(t). \tag{1.1.4}$$

The Brownian motions are correlated,

$$\langle dW_i(t)dW_j(t)\rangle = \rho_{ij}dt. \tag{1.1.5}$$

The cases of dimensionality $d = 2$ and $d = 3$ are implemented in Numerix. As in the one-factor case, all integral functionals of the short rate are normally distributed, and transition probabilities have a Gaussian kernel, which allows for an efficient Green's-function-based implementation.

Calibration

Hull-White 2F interest rate model parameters are calibrated to supplied instruments and forward-rate correlations. This uses the supplied correlation matrix of forward rates to determine the difference between the mean reversions and the ratio of the volatilities of the two factors (which is assumed constant), as well as the correlation between the two factors. It then uses the supplied swaptions to fit time-dependent Hull-White volatilities.

In the HJM formulation of the Hull-White model, changes to the forward rate for instantaneous borrowing at time T are given by

$$\mathsf{d}_t f(t,T) = \sum_i \sigma_t^i \,\mathsf{exp}(-\lambda^i(T-t))\mathsf{dW}_t^i. \tag{1.1.6}$$

The correlation between changes to the rates for borrowing at two different times, T_1 and T_2, is given by

$$\langle \mathsf{d}_t f(t,T_1), \mathsf{d}_t f(t,T_2)\rangle = \frac{\sum_{i,j}\rho_{ij}\sigma_t^i\sigma_t^j \,\mathsf{exp}(-\lambda^i(T_1-t)-\lambda^j(T_2-t))}{\prod_{k=1.2}\sqrt{\sum_{ij}\rho_{ij}\sigma_t^i\sigma_t^j \,\mathsf{exp}(-(\lambda^i+\lambda^j)T_k)}}, \tag{1.1.7}$$

which is a function of σ_t^1/σ_t^2, ρ and $\lambda^1 - \lambda^2$. The correlations between at least four forward rates must be provided.

The correlations are assumed to be ordered as $\langle \mathsf{d}f_{T_i}, \mathsf{d}f_{T_j}\rangle$, $j = i+1,\ldots,n$, $i = 1,\ldots,n-1$ (i.e., the part of the correlation matrix above the diagonal, row by row).

1.1.4 Black-Karasinski and Black-Derman-Toy Models

The Black-Karasinski (BK) model is a one-factor interest rate model where the logarithm of the short rate follows a standard Ornstein-Uhlenbeck process in the risk-neutral measure.

Let $r(t)$ be the short rate at time t. The BK model gives the dynamics of

$$x(t) = \log r(t) \tag{1.1.8}$$

as

$$dx_t = \left(\theta(t) - \lambda x_t\right) dt + \sigma(t) dW_t. \tag{1.1.9}$$

The BK model is implemented using proprietary trees and forward simulations. For backward pricings, the BK model is also implemented using simulations and PDE solver methods.

1.1.5 Ho and Lee Model

The Ho and Lee one-factor model assumes that the short rate is distributed normally with no mean reversion. The model is a simple precursor to Hull-White with

$$dr = \theta(t)dt + \sigma dz, \tag{1.1.10}$$

where $\theta(t)$ is a time-dependent drift that allows the model to be arbitrage free with respect to the initial structure [36].

1.1.6 Spot-Skew Model

The spot-skew model is a one-factor interest rate model that interpolates between Hull-White and Black-Karasinski. It is implemented as a PDE-based backward induction. The short rate $r(t)$ is related to an auxiliary stochastic variable $x(t)$ and a skew parameter β by

$$r(x;\beta) = \frac{1}{\beta}[\exp(\beta x) - 1] + \beta. \tag{1.1.11}$$

The variable $x(t)$ follows a standard Ornstein-Uhlenbeck process in the risk-neutral measure,

$$\mathrm{d}x(t) = (\theta(t) - \lambda x(t))dt + \sigma(t)dW(t). \tag{1.1.12}$$

When $\beta = 0$, this becomes the Hull-White model and when $\beta = 1$, it becomes Black-Karasinski. Varying the skew parameter β allows different smiles to be achieved for caplet/swaption implied volatilities.

1.1.7 Shifted Black-Karasinski Model

In the shifted Black-Karasinski (shifted BK) model, the short rate at time t, $r(t)$ is given by

$$r(t) = e^{x(t)} + \beta,$$

where $\beta \leq 0$ is a constant shift and

$$dx_t = \big(\theta(t) - \lambda x_t\big)\, dt + \sigma(t) dW_t.$$

The shifted BK model acts as a bridge between the BK model and the Hull-White model. At $\beta = 0$, the shifted BK model simplifies to the regular BK model, and as $\beta \to -\infty$, the dynamics get closer to the Hull-White dynamics.

1.1.8 Skewed Black-Karasinski

The Skewed Black-Karasinski model (Skew BK) can be understood in terms of a Hull-White model with optional skewness. A normal underlying $u(t)$ drives the short rate $r(t)$ through the relation

$$r(t) = R(t) + (\exp[S(t)u(t)] - 1)/S(t),$$

where $R(t)$ is the instantaneous yield rate and $S(t)$ is the skewness curve. In the zero-skewness limit, the model is pure Hull-White. By setting $S(t) = 1/R(t)$, a pure Black-Karasinski model is reproduced, up to a trivial scaling of volatility. Note that typical underlying parameters correspond to the *normal* case. For example, typical interest-rate volatilities have 1–1.5% order of magnitude for any skew.

For the one-factor case, the driving underlying $u(t)$ follows the mean-reverting normal process

$$du(t) = (\theta(t) - \lambda(t)\, u(t))\, dt + \sigma(t)\, dW(t).$$

In the two-factor version, the Skewed Black-Karasinski (or Skewed Hull-White) model has two slightly different variations. The first corresponds to the original Hull-White chainwise arrangement in which the first factor drives the short rate. The second variation is the symmetric additive version given, for example, by Brigo and Mercurio [22]. With appropriately correlated Brownian factors dW_0 and dW_1, in the standard "chainwise" variation with mean reversion λ and volatility σ, the model follows

$$\begin{aligned} u(t) &= x_0(t) + c(t), \\ dx_0(t) &= (\lambda_0 x_0(t) + x_1(t))dt + \sigma_0(t) dW_0(t), \\ dx_1(t) &= -\lambda_1 x_1(t) dt + \sigma_1(t) dW_1(t). \end{aligned} \tag{1.1.13}$$

In the "symmetric" variation, the model follows

$$\begin{aligned} u(t) &= x_0(t) + x_1(t) + c(t), \\ dx_0(t) &= -\lambda_0 x_0(t) dt + \sigma_0 dW_0(t), \\ dx_1(t) &= -\lambda_1 x_1(t) dt + \sigma_1 dW_1(t). \end{aligned} \tag{1.1.14}$$

The deterministic additive c(t) is adjusted to match the yield curve.

One can find more details on this model, as well as analytical techniques for swaptions approximation, in [13].

1.1.9 Bounded Short-Rate Model

The bounded short-rate model addresses the boundary requirements of practitioners interested in a bounded evolution of rates, such as for economic scenario generation, where it is necessary for rates to evolve between certain user-defined limits.

Numerix implements a one-factor (1F) version of the bounded short-rate model, which is described in detail in [14]. The dynamics are given by

$$r(t) = f(t,x) = \frac{U(t)e^{\beta(t)\,x} + L(t)\alpha(t)}{e^{\beta(t)\,x} + \alpha(t)}, \tag{1.1.15}$$

with the underlying process satisfying

$$dx(t) = (\phi(t) - a(t)\,x(t))\,dt + \gamma(t)\,dW(t),$$

and where time-dependent functions $L(t)$ and $U(t)$ are the lower and upper boundaries, respectively.

1.1.10 CIR Model

In the original Cox-Ingersoll-Ross (CIR) model [24], the short-rate process is

$$dr = a(c-r)dt + \sigma\sqrt{r}dW.$$

There are four parameters: the volatility σ, the mean reversion a, the mean c and the spot $r_0 = r(0)$. The model is affine and analytically solvable up to European bond options, but it does not adapt to the market yield curve.

Scaled CIR Model

The scaled CIR (SCIR) model is a one-factor short-rate model with unit mean and SDE given by

$$dx = a(t)\,(1-x)\,dt + \sigma(t)\sqrt{x}dW,$$

which is the same as the stochastic volatility factor of Heston model and SVLMM. Both the volatility $\sigma(t)$ and the reversion $a(t)$ are generally time dependent. For constant volatility, the starting (spot) value $x_0 = x(0)$ is an additional adjustable nonnegative parameter with the natural default value $x_0 = 1$.

The short rate r is connected to the underlying stochastic agent by the linear relation

$$r = \lambda(t)x,$$

where $\lambda(t)$ is a deterministic time-dependent nonnegative scaling factor. This factor is not defined by the user explicitly, but automatically adjusted to match the market bond prices.

The time-t value of a zero-coupon bond that matures at T is proportional to an exponential:

$$P(t,T) \propto e^{-s(t,T)x},$$

where s is the solution to the equation

$$\frac{ds}{dt} = \lambda(t) + a(t)s - \frac{1}{2}\sigma(t)^2 s^2$$

with $s(T) = 0$..

This model should not be confused with the CIR++ model introduced by Brigo and Mercurio [22], where the scale λ is constant (as well as the volatility σ and the reversion a) and the bond prices are matched by adding a time-dependent offset $\varphi(t)$ to the short rate. The main drawback of this approach is that the lower bound of the short rate is out of control: $\varphi(t)$ can fluctuate to small positive or negative values, but is never strictly equal to zero.

Scaled CIR2 Model

The SCIR2 model is a two-factor extension of the SCIR model. This is not the most general affine model; rather it is an easily manageable subset, most closely related to the "central tendency" approach by Balduzzi et al. [17]. Compared to a popular model by Brigo and Mercurio [22] (an extension of the Longstaff and Schwartz model [44]), the SCIR2 model more effectively captures the market term structure, including the volatility "hump."

The model can be described in a chain-wise manner: the short rate r is linearly connected to a stochastic agent x, which follows a CIR-like evolution, the drift term of which depends on another stochastic CIR-like process y:

$$\begin{aligned} r &= \lambda(t)x, \\ dx &= a_x(t)\left(1 - \rho(t)\,(1-y) - x\right)dt + \sigma_x(t)\sqrt{x}dW_x, \\ dy &= a_y(t)\,(1-y)\,dt + \sigma_y(t)\sqrt{y}dW_y. \end{aligned}$$

As in the one-factor case, the scaling factor $\lambda(t)$ assumes automatic adjustment to the market bond prices. The parameterization is somewhat similar to that of the two-factor HW or BK models, with two volatilities σ_x, σ_y and two reversions a_x, a_y. The important difference is that the Brownian drives W_x and W_y must be uncorrelated, and the connection of the factors is controlled by the "correlation" coefficient $\rho(t)$ in the drift term for x. The analogy with the statistical correlation is not quite deep, but not totally irrelevant either.

The time-t value of a zero-coupon bond that has maturity T is proportional to an exponential:

$$P(t,T) \propto e^{-s_x(t,T)x - s_y(t,T)y},$$

where s_x and s_y are solutions to

$$\frac{ds_x}{dt} = \lambda(t) + a_x(t)s_x - \frac{1}{2}\sigma_x^2(t)s_x^2$$

$$\frac{ds_y}{dt} = a_y(t)s_y - a_x\rho(t)s_x - \frac{1}{2}\sigma_y^2(t)s_y^2$$

with $s(T) = 0$..

Model Parameters

The SCIR and SCIR2 models are suggested as replacements for the BK and BK2F models, respectively, due to the two substantial advantages: less aggressive high end of the short rate distribution and analytically computed bonds leading to fast and accurate pricing and calibration. The two parameters of SCIR model (volatility σ and reversion a) and the five parameters of the SCIR2 model (volatilities $\sigma_{x,y}$, reversions $a_{x,y}$ and correlation ρ) are similar to those of the BK and BK2F models and of a similar range (all of order of 1). It is important, however, to respect the formal contraints of strictly positive reversions and nonnegative correlation.

Optional Parameters

In the Scaled CIR model, like in the original CIR and BK models, the short rate is strictly positive. Although this is usually quite a desirable feature, the practical market is sometimes slightly arbitrageable (or very close to being so), causing the model to become formally (or effectively) inapplicable because a suitable nonnegative scaling factor $\lambda(t)$ cannot be found.

In order to adapt to such situations, the user is allowed to push down the lower bound of the short rate. This bound is specified by the optional parameter "shift" φ, which extends the linear relation between the short rate and the stochastic agent to

$$r = \varphi + \lambda(t)x$$

Once again, this user-specified parameter φ should not be confused with the automatically adjusted lower bound $\varphi(t)$ in the CIR++ and CIR2++ models [22].

Implementation

The Scaled CIR and CIR2 models are implemented for generic, forward and backward Monte-Carlo methods and can work both standalone and as an IR component of a hybrid model. Also, like any model with strictly positive short rate, they can be used as a credit component in a hybrid model.

1.2 LIBOR Market Models

The direct modeling of market observables such as LIBOR or swap rates naturally leads to the class of *market models* (MM). A standard formulation is due to Brace, Gatarek and Musiela (BGM model, [21]). We recommend the following textbooks on the subject: [4], [20] and [22].

1.2.1 Definition of LIBOR Market Models

A market model is defined for a set of indices $\{I_n(t)\}$. For LIBOR market models (LMMs), the indices are forward LIBOR rates, while for swap market models (SMMs), the indices are swap rates. The set of indices should be sufficiently rich to express the forward discount factors on a given set of dates, $\{T_i\}$. That is, for any T_i, it should be possible to derive the zero-coupon bond price $P(T_i, T_j)$ that is observed on T_i and matures at a later time $T_j \geq T_i$. A diffusion

SDE is imposed on each index. Below, we present technical details concerning the LMM model implemented in Numerix pricing analytics.

The forward LIBOR L_n is a rate starting at T_n and ending at T_{n+1} and is given by

$$L_n(t) = \frac{1}{\delta_n}\left(\frac{P(t,T_n)}{P(t,T_{n+1})} - 1\right), \tag{1.2.1}$$

where $P(t,T_n)$ is the time-t price of a zero-coupon bond with maturity T_n. The definition of an F-factor LMM includes:

- A set of maturities $\{T_n\}$, $n = 0,\ldots,N$, such that $0 < T_0 < T_1 < \cdots < T_N$,
- Initial forward LIBOR rates $\{L_n(0)\}$, $n = 0,\ldots,N$,
- A volatility vector function $\lambda_n(t) = \{\lambda_{n,0}(t),\ldots,\lambda_{n,F-1}(t)\}$ for each $n = 0,\ldots,N-1$,
- Time-independent shifts $\{s_n\}$, $n = 0,\ldots,N-1$.

The LMM SDE is

$$dL_n(t) = \cdots + (L_n(t) + s_n)\,\lambda_n(t)\cdot dW(t). \tag{1.2.2}$$

Here $W(t) = \{W_0(t),\ldots,W_{F-1}(t)\}$ is a set of independent Brownian increments, "$\cdot$" denotes an inner product and "$\cdots$" is a measure-dependent drift. (See Section 1.2.2 for the calculation of this drift term.)

The shifts s_n are subject to the additional conditions

$$-L_n(0) < s_n < \frac{1}{\delta_n}, \qquad n = 0,\ldots,N-1. \tag{1.2.3}$$

Setting shifts to zero gives the classical lognormal version of the model,

$$dL_n(t) = \cdots + L_n(t)\,\lambda_n(t)\cdot dW(t).$$

Numerix also provides a stochastic volatility (SV) version of the LMM that has extra input parameters, including the volatility-of-volatility $\eta(t)$ and a time-independent reversion θ. In this case, the forward LIBOR evolution is given by

$$dL_n(t) = \cdots + (L_n(t) + s_n)\sqrt{z(t)}\,\lambda_n(t)\cdot dW(t), \tag{1.2.4}$$

where the stochastic volatility $z(t)$ is common for all forward LIBOR rates.

The SV process $z(t)$ follows the Cox-Ingersoll-Ross (CIR) process

$$dz(t) = \theta(1-z(t))dt + \eta(t)\sqrt{z(t)}\,dV(t), \qquad z(0) = 1, \tag{1.2.5}$$

with *independent* Brownian motions $W(t)$ and $V(t)$. Andersen and Andreasen [2] use an equivalent form of (1.2.4),

$$dL_n(t) = \cdots + \left\{\beta_n\,L_n(t) + (1-\beta_n)\,L_n(0)\right\}\sqrt{z(t)}\,\sigma_n(t)\cdot dW(t), \tag{1.2.6}$$

where the shifts and volatilities of (1.2.4) and (1.2.6) are related via[1]

$$\lambda_n(t) = \beta_n \, \sigma_n(t) \qquad \text{and} \qquad s_n = L_n(0) \, \frac{1-\beta_n}{\beta_n}. \tag{1.2.7}$$

Here, the shift-adjusted vector volatilities

$$\sigma_n(t) = \{\sigma_{n,0}(t), \cdots, \sigma_{n,F-1}(t)\} \quad \text{for} \quad n = 0, \ldots, N-1 \tag{1.2.8}$$

have the sense and order of lognormal (BS) volatilities (~15%) for any shift.

Note: In Numerix pricing analytics, we mainly use a hybrid input for model construction, i.e., absolute shifts s_n and shift-adjusted volatilities $\sigma_n(t)$. In order to restore the inputs $\{\lambda_n(t), s_n\}$ or $\{\sigma_n(t), \beta_n\}$, one should apply the conversion formulae (1.2.7). Users may also input the "bare" form $\{\lambda_n(t), s_n\}$ corresponding to (1.2.4), which becomes the only possible choice in the case of nonpositive initial Libor rates: the hybrid input $\{\sigma_n(t), \beta_n\}$ is forbidden[2] for $L_n(0) \le 0$.

The LMM volatility $\sigma_{n,f}(t)$ (or $\lambda_{n,f}(t)$) has a large number of free functions, which sometimes makes it convenient to use simplifying assumptions regarding the volatility structure. Volatility decompositions available in Numerix pricing analytics are:[3]

- $\sigma_{n,f}(t) = \bar{\sigma}_f(t)\alpha_f(T_n)$
  This form has a transparent "short-rate" meaning. With a good approximation, a forward curve at time t has F degrees of freedom, each of which is defined by a global factor $X_f(t) = \int_0^t \bar{\sigma}_f(s) dW_f(s)$. Each forward LIBOR has its individual movement coefficients $\alpha_f(T_n)$.
- $\sigma_{n,f}(t) = \gamma_n(t) \, \mathcal{N}_{n,f}$
  This form guarantees constant-in-time correlations between the forward LIBOR increments $C_{nm} = \mathcal{N}_n \cdot \mathcal{N}_m$.

Or, equivalently, one can impose these decompositions on the volatilities $\lambda_{n,f}(t)$.

1.2.2 Implementation of Market Models: Numeraire and Discretization

The drift in the SDEs (1.2.2) and (1.2.4) depends on the measure, which is fixed by the choice of a numeraire. We use a measure associated with the so-called rolling spot (RS) numeraire, which is given by

$$N(t) = \frac{1}{P(T_0, T_1)} \cdots \frac{1}{P(T_{i-1}, T_i)} \, \frac{P(t, T_{i+1})}{P(T_i, T_{i+1})}, \tag{1.2.9}$$

for $T_n \le t \le T_{n+1}$. This numeraire is defined on a discrete set of maturities $\{T_i\}$.

[1] We should mention, however, that the equivalence of the two forms is invalid for a special case of zero initial Libor $L_n(0) = 0$.

[2] Negative initial Libor rates ($L_n(0) < 0$) lead to negative $\beta_n(t)$ (which is counterintuitive), and the case $L_n(0) = 0$ breaks the Andersen and Andreasen form (1.2.6).

[3] See [22] for other examples.

The SDEs for the LMM forward LIBOR rates in the RS measure have the form

$$dL_n(t) = (L_n(t) + s_n) \left(\sum_{j=\eta(t)}^{n} \frac{\delta_j \, (L_j(t) + s_j)}{1 + \delta_j \, L_j(t)} \, z(t) \, \lambda_n(t) \cdot \lambda_j(t) \right) dt + (L_n(t) + s_n) \sqrt{z(t)} \, \lambda_n(t) \cdot dW(t), \quad (1.2.10)$$

where $\eta(t) = m + 1$ for $T_m < t \leq T_{m+1}$. Setting $z(t)$ to one restores the non-SV LMM model.

The implementation is based on simulations and requires a discretization of the evolution prescribed by Eq. (1.2.10). For a given time grid $\{t_k\}$, one has the following exact equation for the LIBOR evolution:

$$L_n(t_{k+1}) = (L_n(t_k) + s_n) \times \exp \left(\int_{t_k}^{t_{k+1}} d_n(t) dt + \int_{k_i}^{t_{k+1}} \sqrt{z(t)} \, \lambda_n(t) \cdot dW(t) - \frac{1}{2} \int_{t_k}^{t_{i+k}} z(t) \, |\lambda_n(t)|^2 dt \right) s_n, \quad (1.2.11)$$

with a drift term equal to

$$\int_{t_k}^{t_{k+1}} d_n(t) dt = \sum_{j=\eta(t)}^{n} \int_{t_k}^{t_{k+1}} \frac{\delta_j \, (L_j(t) + s_j)}{1 + \delta_j \, L_j(t)} \, z(t) \, \lambda_n(t) \cdot \lambda_j(t) dt. \quad (1.2.12)$$

To simulate the given time-step in terms of a Brownian increment $\Delta W_k = W(t_{k+1}) - W(t_k)$, we use standard second-order schemes.

1.2.3 Interpolation and Dynamics of Rates

The simulated LMM permits direct calculation of instruments that have fixing and payment dates that fall on the model time grid because all ratios $P(T_i, T_n)/P(T_i, T_{n+1})$ are known. To evaluate an instrument that depends on on zero-coupon bond values $P(t, T)$ for $t, T \neq T_i$, interpolation is required.

The bond ratios

$$M_n(t) = \frac{P(t, T_n)}{P(t, T_{n+1})} \quad (1.2.13)$$

satisfy an SDE

$$dM_n(t) = \cdots + M_n(t) \, \alpha_n(t) \cdot dW, \quad (1.2.14)$$

with the diffusion term

$$\alpha_n(t) = \frac{\delta_n \, (L_n(t) + s_n)}{1 + \delta_n \, L_n(t)} \sqrt{z(t)} \, \lambda_n(t) \quad (1.2.15)$$

for the shifted BGM model.

We use arbitrage-free lognormal interpolation of the bonds, which is done in two steps. The interpolation in maturity time T is given by

$$P(t, T) = P(t, T_n)^{\frac{T_{n+1} - T}{T_{n+1} - T_n}} \, P(t, T_{n+1})^{\frac{T - T_n}{T_{n+1} - T_n}} \exp \left\{ \frac{1}{2} \frac{T - T_n}{T_{n+1} - T_n} \frac{T_{n+1} - T}{T_{n+1} - T_n} \int_0^t d\tau |\alpha_n(\tau)|^2 \right\} \quad (1.2.16)$$

for $T_n < T < T_{n+1}$. The interpolation in t is done according to

$$P(t,T_n) = P(T_i,T_{i+1})^{\frac{T_{i+1}-t}{T_{i+1}-T_i}} \exp\left\{-\frac{1}{2}\frac{t-T_i}{T_{i+1}-T_i}\frac{T_{i+1}-t}{T_{i+1}-T_i}\int_0^{T_i} d\tau|\alpha_i(\tau)|^2\right\} \times \frac{1}{M_{i+1}(t)\,M_{i+2}(t)\cdots M_{n-1}(t)} \quad (1.2.17)$$

for $T_i < t \le T_{i+1}$.

Given interpolated bonds $P(t,T)$, one can compute any financial rate (index) and its dynamics. We introduce a generic rate $R(t)$ as a ratio of two linear combinations of bonds:

$$R(t) = \frac{\sum_i a_i\,P(t,t_i)}{\sum_j a'_j\,P(t,t'_j)} \quad (1.2.18)$$

for any times t_i, t'_i and coefficients a_i, a'_i. For example, the forward LIBOR is

$$L(t) = \frac{1}{T'-T}\left(\frac{P(t,T)}{P(t,T')}-1\right) = \frac{P(t,T)-P(t,T')}{(T'-T)\,P(t,T')}, \quad (1.2.19)$$

and the forward CMS rate is

$$CMS(t) = \frac{P(t,t_0)-P(t,t_n)}{\sum_{j=1}^n (t_j-t_{j-1})\,P(t,t_j)}. \quad (1.2.20)$$

The SDE for the rate $R(t)$ has the form

$$dR(t) = \cdots + \frac{\sum_j a'_j\,P(t,t'_j)\sum_i a_i\,dP(t,t_i) - \sum_i a_i\,P(t,t_i)\sum_j a'_j\,dP(t,t'_j)}{\left(\sum_j a'_j\,P(t,t'_j)\right)^2}, \quad (1.2.21)$$

or

$$dR(t) = \cdots + R(t)\,\frac{1}{\sum_i a_i\,P(t,t_i)}\left(\sum_i a_i\,dP(t,t_i) - R(t)\sum_j a'_j\,dP(t,t'_j)\right). \quad (1.2.22)$$

The instantaneous lognormal volatility $\sigma_R(t)$ of the process

$$dR(t) = \cdots - R(t)\,\sigma_R(t)\cdot dW(t) \quad (1.2.23)$$

is given by

$$\sigma_R(t) = \frac{1}{\sum_i a_i\,P(t,t_i)}\left(\sum_i a_i\,P(t,t_i)\,\sigma(t,t_i) - R(t)\sum_j a'_j\,P(t,t'_j)\,\sigma(t,t'_j)\right). \quad (1.2.24)$$

An instantaneous correlation at the origin between two rates R and R' (used in the calibration) can be easily calculated using the bracket process

$$\langle dR(t)\,dR'(t)\rangle = R(t)\,R'(t)(\sigma_R(t)\cdot\sigma_{R'}(t))\,dt.$$

Thus,

$$\text{Corr}(R(0),R'(0)) = \frac{(\sigma_R(0)\cdot\sigma_{R'}(0))}{\sqrt{(\sigma_R(0)\cdot\sigma_R(0))\,(\sigma_{R'}(0)\cdot\sigma_{R'}(0))}}. \quad (1.2.25)$$

Note that the rate log-volatilities depend on the initial yield curve and the vector LMM volatilities $\lambda_f(0, T_n)$ at the origin.

1.2.4 European Swaption Pricing under Stochastic Volatility

Here, we describe techniques for European swaption pricing in the SV LMM that is implemented in Numerix pricing analytics[4]. We use the following approximation methods described below:

- **Displaced diffusion approximation methods**
  Approximation of the swap rate with a displaced diffusion (used for both SV and non-SV cases).
- **Asymptotic expansion methods**
  Highly accurate modern method (currently used only for non-SV cases).
- **Rank-one methods**
  Classical BGM nonlinear method (used only for non-SV cases).

Displaced Diffusion Approximation Methods

The pricing is done using a sequence of approximations, via Piterbarg's heuristic way (based on a Hull-White formula for the swap rate) or Markovian projection. It is also possible to use the classical BGM nonlinear approximation. This analytics type is also known as "Rank-One Analytical Swaption Pricing". (See Chapter 6.11 of [22].)

The first step is a derivation to approximate the SDE for a forward swap rate $R(t)$, given the forward LIBOR evolution (1.2.4) or (1.2.6). Namely, we impose a simple displaced diffusion evolution

$$dR(t) \simeq \cdots + \left\{\beta(t)R(t) + (1-\beta(t))R(0)\right\}\, \sqrt{z(t)}\,\sigma(t)\cdot dW(t), \tag{1.2.26}$$

where the time-dependent shift and volatility functions ($\beta(t)$ and $\sigma(t)$, respectively) are determined in an optimal way. Then, both caplets and swaptions can be calculated in their respective martingale measures using the same analytical techniques. The goal of the analytics is to calculate the average $\mathbb{E}[(R(t)-K)^+]$ for a displaced diffusion with SV

$$dR(t) = \left\{\beta(t)R(t) + (1-\beta(t))R(0)\right\}\, \sqrt{z(t)}\,\sigma(t)\cdot dW(t), \tag{1.2.27}$$

where $R(t)$ can mean a forward swap—or a forward LIBOR—rate.

The next step is to average the shift, i.e., replace a time-dependent shift by a constant one,

$$\beta(t) \to \beta. \tag{1.2.28}$$

(This step is absent for caplets as long as we consider using time-independent constant shifts for forward LIBOR rates). After that, the SDE (1.2.27) transforms to

$$dR(t) = \left\{\beta R(t) + (1-\beta)R(0)\right\}\, \sqrt{z(t)}\,\sigma(t)\cdot dW(t) \tag{1.2.29}$$

and can be easily solved analytically.

[4]The non-SV case is restored by $z(t) = 1$.

We will cover the above steps in their natural order:

Swaption approximation formula (1.2.27)

$\Downarrow$

Shift averaging (1.2.28)

$\Downarrow$

Solutions for $E[(R(t)-K)^+]$ of the rate process (1.2.29)

for the time-independent shift.

Swaption approximation formula. Let us fix a forward swap rate $R(t)$ such that $\mathbb{E}[(R(t)-K)^+]$ gives an underlying option payoff in the corresponding swap measure (modulo a constant multiplier). Then one can calculate its derivatives over forward LIBOR rates, i.e., $\partial R(t)/\partial L_n(t)$. Then, the rate SDE in the swap measure looks like

$$dR = \sum_n \frac{\partial R(t)}{\partial L_n(t)}\, dL_n(t) = \sqrt{z(t)} \sum_n \frac{\partial R(t)}{\partial L_n(t)}\, \left\{\beta_n\, L_n(t) + (1-\beta_n)\, L_n(0)\right\} \sigma_n(t)\cdot dW(t). \tag{1.2.30}$$

Our goal is to approximate the above SDE with

$$dR(t) = \left\{\beta(t)R(t) + (1-\beta(t))R(0)\right\}\, \sqrt{z(t)}\, \sigma(t)\cdot dW(t). \tag{1.2.31}$$

Two techniques for estimating the effective volatility and the shift are available in Numerix pricing analytics: Piterbarg's heuristic way [53] and Markovian projection [56] and [10]. They agree for the optimal volatility

$$\sigma(t) = \sum_n \frac{l_n}{R_0}\frac{\partial R_0}{\partial l_n}\sigma_n(t) = \sum_n \frac{\partial(\ln R_0)}{\partial(\ln l_n)}\sigma_n(t) \tag{1.2.32}$$

and differ in the optimal shift expression. The Markovian projection formula

$$\beta(t) \;=\; \frac{\sum_n \left(\frac{1}{2}\frac{\partial|\sigma(t)|^2}{\partial(\ln l_n)} + \frac{\partial(\ln R_0)}{\partial(\ln l_n)}\left(|\sigma(t)|^2 - (1-\beta_n(t))(\sigma(t)\cdot\sigma_n(t))\right)\right)\int_0^t(\sigma_n(\tau)\cdot\sigma(\tau))d\tau}{|\sigma(t)|^2\int_0^t|\sigma(\tau)|^2 d\tau} \tag{1.2.33}$$

gives, in general, a better approximation than the normalized Piterbarg's heuristic expression

$$\beta(t) = \frac{\sum_n \frac{\partial(\ln R_0)}{\partial(\ln l_n)}\,\beta_n(t)(\sigma_n(t)\cdot\sigma(t))}{|\sigma(t)|^2}. \tag{1.2.34}$$

Here we have denoted initial values of Libors $l_n \equiv L_n(0)$, rates $R_0 = R(0)$ and their derivative at origin $\frac{\partial R_0}{\partial l_n} \equiv \frac{\partial R(t)}{\partial L_n(t)}\Big|_{t=0}$.

Shift averaging. Continuing, we follow Piterbarg [54] to replace the time-dependent shift process

$$dR(t) = \left\{\beta(t)\, R(t) + (1-\beta(t))\, R(0)\right\}\, \sqrt{z(t)}\, \sigma(t)\cdot dW(t) \tag{1.2.35}$$

by a flat shift process (less difficult for analytical treatment)

$$dR(t) = \left\{\beta\, R(t) + (1-\beta)\, R(0)\right\}\, \sqrt{z(t)}\, \sigma(t) \cdot dW(t), \tag{1.2.36}$$

where β is represented as a weighted sum,

$$\beta = \int_0^T \beta(t)\, w(t)\, dt. \tag{1.2.37}$$

The weights are given by

$$w(t) = \frac{\nu^2(t)\, \sigma^2(t)}{\int_0^T \nu^2(s)\, \sigma^2(s)\, ds} \tag{1.2.38}$$

for

$$\nu^2(t) = \mathbb{E}[(X_0(t) - R(0))^2\, z(t)], \tag{1.2.39}$$

where the process X_0 is a rough approximation for the initial rate process,

$$dX_0(t) = \sqrt{z(t)}\, \sigma(t) \cdot dW(t), \text{ for } X_0(0) = R(0).$$

Finally,

$$\nu(t) = z_0^2 \int_0^t \sigma^2(s)\, ds + z_0\, \eta^2\, e^{-\theta\, t} \int_0^t \sigma^2(s)\, \frac{e^{\theta\, s} - e^{-\theta\, s}}{2\theta}\, ds. \tag{1.2.40}$$

Solutions for $\mathbb{E}[(R(t)-K)^+]$ of the rate process (1.2.29) for the time-independent shift. To compute the average $\mathbb{E}(R(t)-K)^+$ for the rate process

$$dR(t) = (R(t)+s)\, \sqrt{z(t)}\, \lambda(t) \cdot dW(t), \tag{1.2.41}$$

we distinguish two cases. For the non-SV case ($z(t)=1$), the option pricing is done via a shifted BS formula. For the SV case, we calculate a characteristic function of $\log(R(t)+s)$ using standard Heston ODE techniques and perform inverse Fourier integration numerically to obtain the option price.

Asymptotic Expansion Methods

The idea of this highly accurate modern method is an approximation of the swap rate process with a quadratic volatility one fitting three expansion terms of the corresponding swaption price. See [11] for details. For the moment, this method can only be used for non-SV cases.

Rank-One Method

This is a classical BGM non-linear method, also known as "Rank-One Analytical Swaption Pricing". (See Chapter 6.11 of [22].) The advantage of our implementation is that it supports *general* option pay/exercise dates, although the classical scheme requires them to fall on the model maturity dates T_n. This method can only be used for non-SV cases.

1.2.5 European CMS Instruments Approximation

A full description of the general methodology and derivation of CMS pricing formulas can be found in [6]. A brief and incomplete summary follows.

We use standard notations for a forward swap rate $S(t)$ starting at T_B and ending at T_E,

$$S(t) = \frac{P(t,T_B) - P(t,T_E)}{\sum_{i=B+1}^{E} \delta_{i-1} P(t,T_i)}. \tag{1.2.42}$$

We will refer to the denominator in the swap definition (1.2.42) as a *swap level*, or annuity, and denote it by

$$L_S(t) = \sum_{i=B+1}^{E} \delta_{i-1} P(t,T_i).$$

The martingale measure for the swap rate (1.2.42) is associated with the numeraire

$$N(t) = \frac{L_S(t)}{L_S(0)}. \tag{1.2.43}$$

We can approximate the forward swap rate $S(t)$ in its martingale measure using the Markovian projection elaborated by Piterbarg [55] and Antonov and Misirpashaev [10],

$$dS(t) \simeq \left(S(t)\beta_S(t) + (1-\beta_S(t))S_0\right)\sqrt{z(t)}\,\sigma_S(t)\cdot dW(t), \tag{1.2.44}$$

where $S_0 = S(0)$. To proceed with calculations for the fixing date τ, we should average the skew parameters using the Piterbarg (2005) formulae. For the averaged skew, we will use the same notations without the time parameter, $\beta_S(t) \to \beta_S$. Thus, the swap rate observed at a fixed time τ can be approximated as

$$S(\tau) \simeq \frac{S_0\, e^{y_S(\tau)}}{\beta_S} - \frac{(1-\beta_S)S_0}{\beta_S}, \tag{1.2.45}$$

where the process $y_S(t)$ is defined by

$$y_S(\tau) = -\frac{1}{2}\beta_S^2 \int_0^\tau \left|\sigma_S(t)\right|^2 z(t)dt + \beta_S \int_0^\tau \sigma_S(t)\cdot dW(t)\sqrt{z(t)}. \tag{1.2.46}$$

We will use the formulae above to derive effective approximations for the single rate CMS products and the CMS spread options.

Single-Rate CMS Products

In this section, we consider single-rate CMS products such as the CMS swap and CMS cap.

Using the approach from [6], we can calculate a single CMS payment price $P(0,T)\mathbb{E}_T[S(\tau)]$ and a CMS caplet price $P(0,T)\mathbb{E}_T[(S(\tau)-K)^+]$ for some strike K. We denote the CMS fixing date as τ and the CMS payment date as T.

In general, in order to calculate averages like $P(0,T)E_T[S(\tau)]$, it is preferable to switch to the martingale measure of the underlying process $S(t)$. For our case, the measure is a swap

measure with numeraire (1.2.43). Then, the CMS payment price transforms to

$$\mathbb{E}_T[S(\tau)] = \mathbb{E}[M(\tau)S(\tau)], \tag{1.2.47}$$

where the measure change process is

$$M(t) = \frac{L_S(0)}{P(0,T)} \frac{P(t,T)}{L_S(t)}. \tag{1.2.48}$$

As explained in [6], the process $M(t)$ can, in turn, be approximated by (1.2.45), which permits evaluating the average $\mathbb{E}[M(\tau)S(\tau)]$ using the corresponding characteristic function.

To calculate the single rate CMS option price

$$P(0,T)\mathbb{E}_T[(S(\tau) - K)^+],$$

one can proceed in two ways. The first way consists of a simple average adjustment for the rate $S(t)$ in the T-forward measure,

$$\mathbb{E}_T[(S(\tau) - K)^+] \simeq \mathbb{E}_S \left[\left(\frac{\mathbb{E}_T[S(\tau)]}{S_0} S(\tau) - K \right)^+ \right], \tag{1.2.49}$$

where the average $\mathbb{E}_T[S(\tau)]$ was approximated above and the process $S(\tau)$ inside the outer average of the right-hand side of (1.2.49) is approximated by (1.2.45).

A more accurate way consists of calculus in the swap measure,

$$\mathbb{E}_T[(S(\tau) - K)^+] = \mathbb{E}[(S(\tau) - K)^+ M(\tau)],$$

with an optimal linearization of the measure change process,

$$\mathbb{E}[M(\tau)|S(\tau)] \simeq A(\tau) + B(\tau)S(\tau).$$

The resulting calculus is done via numerical inversion of the Laplace transform, see [6].

Note that both ways have been implemented in Numerix software.

CMS Spread Options

We consider a CMS spread $S(t) = S_1(t) - S_2(t)$, where S_1 and S_2 are some forward CMS rates. A call option on the spread fixed at τ with a payment date T and a strike K has price

$$\mathcal{C}_{sp} = P(0,T)\mathbb{E}_T[(S_1(\tau) - S_2(\tau) - K)^+].$$

The simplest approach is to approximate the above expression directly in the T-forward measure via an average adjustment, analogous to (1.2.49):

$$\begin{aligned} \mathcal{C}_{sp} &= P(0,T)E_T[(S_1(\tau) - S_2(\tau) - K)^+] \\ &\simeq P(0,T)E_T \left[\left(\frac{E_T[S_1(\tau)]}{S_1(0)} S_1^*(\tau) - \frac{E_T[S_2(\tau)]}{S_2(0)} S_2^*(\tau) - K \right)^+ \right], \end{aligned} \tag{1.2.50}$$

where the S_i^* processes are approximated by the Markovian projection as in (1.2.45).

An advanced approximation corresponds to calculus in the spread measure, where the spread process is martingale; see [6] for more details. The spread call option approximation can be written as

$$E_T[(S(\tau)-K)^+] \simeq E_S[(S(\tau)-K)^+ (A(\tau)+B(\tau)\,\Delta S(\tau)+C(\tau)\,(\Delta S(\tau))^+)]. \tag{1.2.51}$$

Calculation of the optimal coefficients A, B, and C, and the resulting average in the right-hand side of (1.2.51) is described in [6]. Note that a nonlinear approximation of the measure change has an experimental foundation and permits obtaining higher quality approximations.

1.2.6 Calibration

In this section, we will address the calibration of market models in Numerix implementations, beginning with some preliminaries, general calibration goals, and strategies.

The calibration goal is to come up with model parameters such that the constructed model fits the calibration input, has smooth volatility parameters and remains attractive and effective for further numerical implementation—for example, has a small number of LIBOR dates (i.e., a small number of LIBOR rates (states) to simulate) or has a small number of factors.

The LMM calibration strategy can include the following steps:

- Given a calibration input, choose the model LIBOR dates T_n (or use a direct input for T_n, optional for toolkit calibration).
- Choose the volatility decomposition form (or some other practical volatility characteristic).
- Choose a European swaption analytical formula.
- Find the volatility (analytically or numerically) to fit the calibration input.

Calibration input

The calibration input includes market information on option prices, correlations and European CMS products. A user can specify one market input type in the form of an arbitrary set of:

- swaptions/caplets
- correlations between different rates
- European CMS instruments (CMS swaps, CMS caps/floors, CMS spread options)

The correlations can be between arbitrary forward LIBOR rates or between arbitrary rates, like CMS, etc. The option input consists of a vector of options (dates and market prices) and their weights (optional). The correlations can be input as:

- *a forward LIBOR rates correlations matrix*:
  - correlation matrix C'_{ij}
  - correlation dates: start dates $\tau_i^{(st)}$ and end dates $\tau_i^{(end)}$

A forward LIBOR with start $\tau_i^{(st)}$ and $\tau_i^{(end)}$ has an instantaneous market correlation at origin equal to C'_{ij} with a forward LIBOR with start $\tau_j^{(st)}$ and $\tau_j^{(end)}$.

- *a general rates correlations matrix*:
  - correlation matrix C'_{ij}
  - rate vector: R_i

  The rates R_i and R_j have an instantaneous market correlation at the origin equal to C'_{ij}.
- *a general rates correlations vector*:
  - correlation vector C'_i
  - two rate vectors: R_i and R'_i

  The rates R_i and R'_i have an instantaneous market correlation at the origin equal to C'_i.

See Section 1.2.3 for more details on rates correlations. If the correlation dates coincide with the BGM maturities, one can use exact formulae for model correlations. Otherwise, a volatility interpolation should be used.

The European CMS input can be:

- *CMS swap*
  The CMS swap consists of individual CMS payments, which are characterized by CMS rate parameters (length/frequency), its fixing date and its payment date.
- *CMS cap/floor*
  The CMS cap/floor is a sequence of individual CMS caplets/floorlets defined with a strike, CMS rate parameters (length/frequency), its fixing date and its payment date.
- *CMS spread option*
  The CMS spread option is a sequence of individual CMS spread options defined with a strike, two CMS rate parameters and a common fixing date and payment date.

Calibration Output

The calibration output for LMM includes maturity dates T_n, volatilities, shifts and the volatility-of-volatility. Numerix pricing analytics permit the calibration of any items from the list above with others being fixed. For example, one can fix maturity and volatility-of-volatility, and calibrate rate volatilities and shifts.

Calculation of the LMM Maturities

The LMM maturities T_n are constructed as follows. First, we set a vector that includes exercise dates of European swaptions, the model origin and the last payment date. Second, we stratify[5] the obtained vector with an interval of one year and remove the first element (the origin).

[5]Stratification of a vector with elements t_i with the stratification interval δt works as follows: For each interval $[t_i, t_{i+1}]$, we *uniformly* insert n extra points. The number n is chosen such that the obtained subintervals $(t_{i+1}-t_i)/(n+1)$ are as close as possible to the stratification interval δt. The stratification does not remove initial points t_i. Namely, if given an interval with $t_{i+1} - t_i \leq \delta t$, the algorithm does not insert extra points in this interval.

Thus, the obtained LMM maturities T_n do not necessarily contain swaption payment dates. This becomes possible due to *interpolated* European analytics.

Remark. This choice of maturities strongly reduces the number of LIBOR rates to simulate with respect to the standard calibration, where the LMM maturities are all exercise **and** payment dates. Note, however, that one can set the LMM maturities *directly* as an optional calibration input. (It can include, for example, both exercise and payment dates of the calibration set.) In general, if an exotic instrument depends on a *fine* structure of the yield curve, say, a spread between 3M LIBOR and a short CMS rate, one can set 3M intervals for the coupon date to let the model distinguish between the two underlying rates. Otherwise, if one chooses 1Y coupon date intervals, the model will have the rates (3M LIBOR and short CMS rate) almost perfectly correlated due to the interpolation.

Numerix calibration volatility choices

A user has the possibility of choosing between the following volatility structures:[6]

- **general volatility**
  the matrix of curves $\sigma_{n,f}(t)$
- **flat volatility**
  time-independent volatilities $\sigma_{n,f}(t) = \sigma_{n,f}$
- **composite volatility**
  $\sigma_{n,f}(t) = \bar{\sigma}_f(t)\,\alpha_f(T_n) \equiv \bar{\sigma}_f(t)\,\alpha_{n,f}$
- **time-independent (flat) correlations**
  $\sigma_{n,f}(t) = \gamma_n(t)\,\mathcal{N}_{n,f}$ for different forms of $\gamma_n(t)$ and two forms of *directions* $\mathcal{N}_{n,f}$. The first direction form is a general form without any parametrization; the second one keeps positive correlations between the model forward LIBOR rates and guarantees that a correlation between $L_n(t)$ and $L_m(t)$ decreases with increasing $|n-m|$ and fixed n. The forms for the volatility module $\gamma_n(t)$ include:
  - *General form*
  - *Flat form*
    $\gamma_n(t) = \gamma_n$
  - *Composite form*
    $\gamma_n(t) = \bar{\sigma}(t)\,\alpha_n$
  - *Homogeneous forms*
    $\gamma_n(t) = \bar{\sigma}(T_n - t)\,\alpha_n$, where the homogeneous part $\bar{\sigma}(\tau)$ can be a general curve or parameterized[7], $\bar{\sigma}(\tau) = (a\,\tau + d)\,\exp(-b\,\tau) + c$.

The time-dependent functions are considered to be *step-constant between exercise dates*. In other words, the curve nodes are the option exercise dates. There are three exceptions to this rule:

[6] We present here decompositions for the shift-adjusted volatility $\sigma_n(t)$ used in the model construction. The volatility $\lambda_n(t)$ can be obtained via (1.2.7).

[7] See [22] on justifications of a heaped decaying volatility representation.

- For a one-factor model, the curve nodes are *only exercise dates where there is more than one option*.
- For the homogeneous general volatility form, the nodes of $\bar{\sigma}(\tau)$ are calculated in the following way. First, for given number of time-steps, we construct the model dates t_i. Then, we calculate all possible combinations $T_n - t_i$, sort them and remove the negative part. The obtained vector of dates can contain a lot of elements if the sets $\{t_i\}$ and $\{T_n\}$ are not commensurable. Thus, we remove the close dates from the sorted nodes such that the final vector τ_k contains the same number of elements as the model dates and is as regular as possible. The volatility curve $\bar{\sigma}(\tau)$ is considered to be step-constant between the nodes τ_k.
- For homogeneous parameterized volatility, we use its natural time-dependence $\bar{\sigma}(\tau) = (a\,\tau + d)\,\exp(-b\,\tau) + c$.

Analytical European Option Pricing

We use the analytical method described in Section 1.2.4.

Analytical European CMS Product Pricing

The analytical methods for the CMS products approximation are described in Section 1.2.5.

Numerical Calibration

Having made the above choices, we look for the model parameters to fit input option prices (or implied volatilities) and input correlations using a highly effective proprietary global solver which makes the calibration procedure universal. It does not depend on the option choice—caplets or swaptions—and supports any new correlation types.

The solver used in the calibration is a multi-dimensional quasi-Newton solver. It is very effective when one or multiple solutions exist.

The initial guesses for the calibrator are as follows.

- SV process (1.2.5):
$$\theta = 0.5, \qquad \eta = 1.$$
- Libor shifts s_n (1.2.4):
$$s_n = \frac{1}{2}(\underline{s_n} + \overline{s_n}),$$
where $\underline{s_n} = -L_n(0)$ (or $\underline{s_n} = 0$ in the case of flat positive shift) and $\overline{s_n} = \frac{1}{T_{n+1} - T_n}$. The shifts are capped uniformly by 0.1.
- Full volatility types:
  - General and flat volatility ($\sigma_{n,f}(t)$ is a matrix of curves or $\sigma_{n,f}(t)$ is constant):
$$\sigma_{n,f} = 0.2 \quad \text{for} \quad f = 0, \quad \sigma_{n,f} = 0 \quad \text{for} \quad f > 0.$$

- Composite volatility ($\sigma_{n,f}(t) = \bar{\sigma}_f(t)\alpha_f(T_n) \equiv \bar{\sigma}_f(t)\alpha_{n,f}$):

$$\bar{\sigma}_f(t) = 1 \quad \text{and} \quad \alpha_{n,f} = 0.2 \quad \text{for} \quad f = 0, \quad \alpha_{n,f} = 0 \quad \text{for} \quad f > 0.$$

- Flat correlations ($\sigma_{n,f}(t) = \gamma_n(t)\mathcal{N}_{n,f}$):

  The matrix $N_{n,f}$ is constructed in the following way. Set

$$\theta_n = \frac{n\bar{\theta}}{M-1},$$

  where M is the number of Libors and $\bar{\theta} = \arccos(\rho^*)$, where

$$\rho^* = \text{fabs}(\text{libor\_dates}[\text{last}] - \text{libor\_dates}[0]) * \text{libor\_decorrelation}.$$

  Then

$$N_{n,0} = \cos(\theta_n), \quad N_{n,1} = \sin(\theta_n), \quad \text{and} \quad N_{n,f} = 0 \quad \text{for} \quad f > 1.$$

  The defaults for the particular volatility cases are:

  - General and flat volatility: $\gamma_n(t) = \gamma_n = 0.2/N$, where N is the number of factors.
  - Composite and homogeneous volatility: $\bar{\sigma}(t) = 1$, $\alpha_n = 0.2$.

1.2.7 Pricing with the Model

Pricing instruments (that are non-path dependent) requires that it be possible to extract any zero-coupon bond $P(t,T)$ to calculate indices and that it be possible to calculate a conditional expectation (rollback) to price callable instruments.

The conditional expectation is calculated using MC least-square minimization for some set of states $\{s_i(t)\}$, i.e.,

$$\mathbb{E}[X \mid F_t] \to \mathbb{E}[X \mid \sigma(s_1(t), s_2(t), \cdots)] = f(s_1(t), s_2(t), \cdots).$$

The form and number of states is controlled by the user. There are two groups of states: interest rate (IR) states and stochastic volatility states.

A stochastic volatility state (if present) is the stochastic volatility $z(t)$. Interest rate (IR) states are zero-coupon bonds or their logarithms. Namely, for any time note t, we choose states being $\{P(t,\tau_1), \ldots, P(t,\tau_M)\}$, where M is the number of IR states and $\{\tau_i\}$ is a set containing the LMM maturities $T_n > t$ which cover the interval $[t, T_N]$ in the most homogeneous way for the given number of IR states, M. For example, for $M = 1$, we use a single state $P(t,T_N)$; for $M = 2$, we use two states $P(t,T_N)$ and $P(t,\eta(t))$ (the first valid bond); for three IR states, we take three bonds: $P(t,T_N)$, $P(t,\eta(t))$, and a bond with maturity at the middle of $[t,T_N]$, etc.

In general, the number of IR states should be adapted to the model type, including the number of factors and volatility type. Neglecting the drift in the LIBOR evolution (1.2.10), we can approximate a forward LIBOR $L_n(t)$ as a function if its "driving factor"

$$Y_n(t) = \int_0^t \sqrt{z(s)}\, \lambda_n(s) \cdot dW(s). \tag{1.2.52}$$

We can calculate a covariance matrix of the driving factors Cov(t),

$$\text{Cov}_{nm}(t) \equiv \mathbb{E}[Y_n(t)\, Y_m(t)] = \int_0^t E[z(s)]\, \lambda_n(s) \cdot \lambda_m(s)\, ds = \int_0^t \lambda_n(s) \cdot \lambda_m(s)\, ds,$$

taking into account independence of the SV and the LIBOR Brownian motion $W(t)$. The covariance matrix Cov(t) rank approximates the number of degrees of freedom of the forward curve at t which, in its turn, gives an idea about the number of IR states.

For the flat volatility type ($\lambda_m(t) = \lambda_m$) and the composite type, the rank is equal to the number of factors (F) and suggests using F IR states. For the other volatility types, the time-dependence of $\lambda_m(t)$ changes this rule, although, in practice, one still has F dominating eigenvalues which gives F number of IR states (or, rarely, $F+1$ IR states to take into account the time dependence of the volatility).

From the instrument pricing prospective, one should adapt the LMM number of factors to the instrument payoffs. Instruments that are sensitive to multiple points of the yield moves should be priced on a model that has sufficiently large degrees of freedom of the yield curve. For example, to price a callable CMS spread instrument, one should use at least a 2-F model and adapt the number of states according to the above rules (2-3 IR states plus one SV state in the presence of SV).

On the other hand, the above estimation of the number of states bears approximate character. Any particular case requires experiments on the instrument pricing, with different numbers of states, and convergence analysis.

1.3 CMS Copula Model

The CMS copula model is used to price multi-rate vanilla options. Such options include those with payoffs

- Spread option: $\left(S_T^1 - S_T^2 - K\right)^+$,
- Dual digital option: $\mathbb{I}_{\omega_1 S_T^1 > \omega_1 B_1} \mathbb{I}_{\omega_2 S_T^2 > \omega_2 B_2}$,
- Conditional CMS spread: $\left(S_T^1 - S_T^2 - K\right)^+ \mathbb{I}_{\omega S_T^3 > \omega B_3}$,

where S_T^j is the jth swap rate fixed at T, K is the strike, B_j is the jth barrier and ω_j is the jth put-call indicator. We assume that each S_t^j is a martingale in its annuity measure with numeraire A_t^j. We recommend using a SABR process with the method "ASK" corresponding to 2013 Risk paper "SABR Spread Its Wings" by Antonov, Spector and Konikov. This approach has virtually no arbitrage for all positive strikes even for very long expiries. To price multi-rate options, we need the joint distribution of the S_T^j in a common measure. For this common measure, we choose the terminal measure $\mathbb{Q}^{T_p}$ and use a zero-coupon risk-free bond that matures at the pay date as numeraire. For each rate, we compute its cumulative distribution function in this measure as

$$F(K) = \mathbb{E}^{T_p}\left[\mathbb{I}_{S_T < K}\right] = \frac{A_0}{D_0^{T_p}} \mathbb{E}^A \left[\frac{\mathbb{I}_{S_T < K} D_T^{T_p}}{A_T}\right].$$

Here, D_t^T is the time-t value of the zero-coupon bond with maturity T. We approximate

$$\frac{D_T^{T_p}}{A_T} \approx G(S_T;T_p) := \frac{(S_T-\chi s)(a(T_p)+bS_T)}{\left(1+\frac{1}{m}(S_T-s)\right)^{T_p-T}\left(1-\frac{1}{\left(1+\frac{1}{m}(S_T-\chi s)\right)^{nm}}\right)},$$

where $a(T_p) \approx 1$ and $b \approx 0$ are adjustments we make to guarantee arbitrage conditions are satisfied, s is the OIS spread implied as a difference of S_0 and the discount-based swap rate, χ is an indicator equal 0 for cash-settled swaptions and 1 for physicallly settled ones, n is the rate's tenor in years and m is the rate's fixed frequency. The approximation we made for all the discount factors is a "yield"-based approximation. Using Carr-Madan replication, we have

$$F(K) = \frac{A_0}{D_0^{T_p}}\left(G(K,T_p)\mathsf{DigPutFV}(K) - G'(K,T_p)\mathsf{PutFV}(K) + \int_0^K G''(x,T_p)\mathsf{PutFV}(x)dx\right),$$

where DigPutFV and PutFV stand for the forward values of digital and vanilla puts, respectively. To connect marginals into a joint distribution, we use Gaussian copula, and write

$$F(S_T^1,\ldots,S_T^n) = N_\rho\left(\Phi^{-1}(F_1(S_T^1)),\ldots,\Phi^{-1}(F_n(S_T^n))\right)$$

for the joint distribution, where $N_\rho(x_1,\ldots,x_n)$ is a multi-variate standard normal distribution with correlation matrix ρ, and Φ^{-1} is the inverse of standard cumulative normal distribution. Now, to compute PV and V_0 of a European derivative with payoff $P(S_T^1,\ldots,S_T^n)$, we can write

$$V_0 = D_0^{T_p}\mathbb{E}^{T_p}\left[P(S_T^1,\ldots,S_T^n)\right],$$

which can be computed using quadrature for two- or three-rate payoffs, or using Monte Carlo for a payoff on a larger basket of rates.

Chapter 2

Credit Models

This chapter describes models for credit derivative pricing and credit portfolio risk management. We begin with single-name models in Section 2.1. Static basket models, which are sufficient for standard CDO tranches and credit portfolios, are described in Section 2.2. Finally, Section 2.3 introduces a dynamic top-down model, which is adequate for instruments such as CDO tranches and forward starting CDOs.

2.1 Single-Name Credit Models

2.1.1 Single-Name Deterministic Credit Model

This is the standard deterministic credit model, where both interest rates and survival probabilities are non-stochastic. The model is mainly used for the credit curve stripping, but can also be used to price linear products such as credit default swaps and risky bonds. The model still includes a certain degree of randomness because the timing and the fact of occurrence of a default remains uncertain; however, there is no dependence on any volatilities.

The inputs to the deterministic model are the survival probability curve $P(t)$ and the risk-free discounting curve $D(t)$ as viewed on the valuation date $t = 0$. The calculations are based on the formulas for two elementary cashflow types.

1. The present value of a deterministic amount A payable on the date $T > 0$ conditional on the absence of default up to time T is given by

$$PV_{\text{no default}} = A\,P(T)D(T). \tag{2.1.1}$$

2. The present value of a deterministic amount A payable on date T_2 conditional on the presence of default in the interval $[T_1, T_2]$ with $0 < T_1 < T_2$ is given by

$$PV_{\text{default}} = A\,D(T)(P(T_1) - P(T_2)). \tag{2.1.2}$$

As an example, consider a CDS with notional A, spread S, assumed recovery rate r, grid of payment dates $T_1 < T_2 < \ldots T_N$, and daycount fractions $\Delta_i = \Delta(T_{i-1}, T_i)$ computed using the appropriate daycount conventions (we set $T_0 = 0$). The fee leg is computed as a present value of

the stream of payments (2.1.1) conditional on the absence of default, plus the interest accrued from the last payment date to default.

$$PV_{\text{fee leg}} = S\,A\sum_i \Delta_i P(T_i)D(T_i) + \text{AI}. \tag{2.1.3}$$

The protection leg can be evaluated as a stream of default-conditional payments of the type (2.1.2) using a suitably frequent auxiliary subdivision $0 = \tau_0 < \tau_1 < \ldots \tau_M = T_N$ of time until CDS maturity,

$$PV_{\text{protection leg}} = (1-r)A\sum_j D(\tau_j)(P(\tau_{j-1}) - P(\tau_j)). \tag{2.1.4}$$

It is often sufficient to use the subdivision $\{T_i\}$ provided by the payment dates of the fee leg. Finally, the accrued interest on the fee leg can also be reduced to a stream of payments of type (2.1.2). Using again the subdivision $\{\tau_j\}$, for every j we let $i(j)$ be the maximal index i such that $T_i \le \tau_j$ and set $T(j) = T_{i(j)}$. The accrued interest is given by

$$\text{AI} = S\,A\sum_j \Delta(T(j),\tau_j)\,D(\tau_j)(P(\tau_{j-1}) - P(\tau_j)). \tag{2.1.5}$$

2.1.2 Short-Rate-Based Stochastic Single-Name Models

Numerix single-name credit models build on the methodology of short-rate interest modeling. The subject of the modeling is the continuously compounded credit spread $s(t)$, which is related to the survival probability $P(t)$ in the same way that the continuously compounded risk-free rate $r(t)$ is related to the discount factor $D(t)$:

$$P(t) = \mathbb{E}\left[\exp\left(-\int_0^t s(\tau)d\tau\right)\right], \tag{2.1.6}$$

$$D(t) = \mathbb{E}\left[\exp\left(-\int_0^t r(\tau)d\tau\right)\right]. \tag{2.1.7}$$

The curves $P(t)$ and $D(t)$, as viewed on the valuation date, can be implied using no-arbitrage considerations from vanilla instruments that are traded on the market. If no stochastic evolution for such curves is introduced, a simple model with deterministic credit spreads and interest rates arises. As already mentioned, this model is sufficient to handle vanilla credit instruments, such as CDS, ASW, TRS and CLN without optionality. For single-name exotics, however, it is necessary to go beyond deterministic spreads and/or interest rates.

Numerix implementations allow choosing any of the standard single-factor short-rate models (HW, BK, BDT, CIR) for credit spreads and risk-free rates. The choice of the models is independent; for example, it is possible to use HW for credit spreads and BK for interest rates. It is also possible to assume deterministic interest rates, in which case the remaining model for the credit spread is one-dimensional, which allows for an efficient implementation on a grid-based tree. With stochastic interest rates, the resulting model is two-dimensional, and the implementation is based on a generic tree. We allow taking into account explicit correlations between the components of a two-dimensional model.

The three basic ingredients for the pricing of single-name credit derivative instruments are:

1. Risky payments: payments conditional on the absence of default
2. Riskless payments: payments independent of the default status
3. Default payments: payments that happen only after default

All payments are discounted to the valuation date using the risk-free discount curve.

To quantify the effect of stochastic credit spreads, it is convenient to introduce the no-default process $\bar{\Lambda}(t)$, which is equal to 1 if no default occurs before t and 0 otherwise. The continuous spread $s(t)$ is the intensity of the no-default process. That is, the probability of the no-default process changing from 1 to 0 during the time interval $[t, t+dt]$ is $s(t)\,dt$. Following [27], the SDE for the no-default process can be written in the form

$$d\bar{\Lambda}(t) = -\bar{\Lambda}(t)\,s(t)\,dt + dM(t), \tag{2.1.8}$$

where dM is a martingale under the chosen measure. Note that if the spread $s(t)$ is stochastic and following HW, BK or BDT evolutions, the jump martingale dM is *independent* of the spread. So taking an average over jumps, or in other words over $dM(t)$, results in the survival probability for the interval $[t_1, t_2]$ given that no default occurred before the time t_1:

$$\mathbb{E}_{jumps}\left[\bar{\Lambda}(t_2)\right] = \bar{\Lambda}(t_1)\ \exp\left(-\int_{t_1}^{t_2} s(\tau)d\tau\right). \tag{2.1.9}$$

This survival probability is a function of the spread between t_1 and t_2; for each realization of the spread, the survival probability is different.

It can be shown that the average over jumps of an arbitrary risky payoff reduces to the additional discounting by the value of the survival probability factors, similar to the intuitive relationships that hold in the non-stochastic case. As a result, a payment of $X(T)$ subjected to a credit risk (i.e., risky payment) seen from time t is

$$\bar{\Lambda}(t)\mathbb{E}\left[X(T)\exp\left(-\int_t^T (r(\tau)+s(\tau))\right)d\tau\right]. \tag{2.1.10}$$

The remaining expectation is over the credit spread $s(t)$ and the risk-free rate $r(t)$.

Risky payments are organized using risky slices and discounted by

$$\exp\left(-\int_t^T (r(\tau)+s(\tau))d\tau\right) \tag{2.1.11}$$

during the rollback. Non-risky payments are done in terms of non-risky slices and discounted using a normal interest-rate discount

$$\exp\left(-\int_t^T r(\tau)d\tau\right) \tag{2.1.12}$$

during rollback. Payments conditional on default can be expressed as linear combinations

of risky slices with corrections for risk-free discount factors, similar to the deterministic formula (2.1.2).

2.2 Static Basket Credit Models

Let N be the number of assets in the underlying portfolio, each of which is in good standing as of the valuation date $t = 0$, but has a nonzero probability of default by the horizon date $t = T$. The standard approach to joint default modeling consists of two steps. In the first step, independent single-name default models are defined. In the second step, a joint distribution of defaults is formed to extend the marginal distributions of defaults fixed in step one. The single-name credit default model for asset a consists of its survival probability curve $s_a(t)$ and a model for loss-given-default. The relevant output that summarizes the loss-given-default model for the purpose of credit portfolio modeling is the weight w_a of each asset in the decomposition of the portfolio loss in terms of asset default indicators $U_a(T)$,

$$L(T) = \sum_a w_a U_a(T). \tag{2.2.1}$$

In the case of several portfolios, a separate weight $w_{a,p}$ needs to be introduced for loss exposure of the asset a in each portfolio p,

$$L_p(T) = \sum_a w_{a,p} U_a(T). \tag{2.2.2}$$

The valuation of synthetic CDO tranches that involve senior slices of the portfolio is also affected by expected values of tranche amortization due to recovery. This requires a second set of weights, w'_a, to express the random recovery as

$$R(T) = \sum_a w'_a U_a(T) \tag{2.2.3}$$

(and similarly in the case of several portfolios). We use a fraction-of-notional recovery model of loss-given-default, in which a recovery rate r_a relates the loss and recovery weights to the asset notional A_a as $w_a = (1 - r_a)A_a$ and $w'_a = r_a A_a$. Because the problem of recovery distribution is an exact duplicate of the problem of loss distribution, we do not need to develop any specific techniques for recovery.

2.2.1 Copula Models and Static CDO Model

The current industry standard of introducing correlations is the Gaussian copula model [43]. When the copula model is used to value single-tranche CDOs, we call it the static CDO model. The specific feature of the single-tranche CDO is that its value is fully determined by the time dependence of options on total loss, $\mathbb{E}[(L_t - d)^+]$ and $\mathbb{E}[(L_t - u)^+]$, where d and u are tranche attachment points. We use the name "static CDO model" for any copula model when we restrict ourselves to the calculation of options on aggregated loss and do not seek to model explicit sequences of individual defaults.

An operational definition of the Gaussian copula model can be given by describing how to simulate the joint distribution of default status of the N assets. We introduce an $N \times N$ matrix

of copula model correlations with entries $\{\rho_{ab}\}$. On each simulation path, a random vector $(Z_1, \ldots, Z_N)$ is drawn from the N-dimensional multivariate normal distribution with mean 0 and covariance matrix $\rho = [\rho_{ab}]$ so that the density of this vector is given by

$$p(Z_1, \ldots, Z_N) = \frac{1}{\sqrt{(2\pi)^N \det \rho}} \exp\left(-\frac{1}{2}\vec{Z}^T \rho^{-1} \vec{Z}\right). \tag{2.2.4}$$

For any time horizon, the generated random numbers are compared with the threshold barriers, $b_a(T) = \mathcal{N}^{-1}(1 - s_a(T))$. If $Z_a \leq b_a$, we declare that the asset a has defaulted by the time T, otherwise that it has remained alive. Here $\mathcal{N}(x) = (1/2\pi)\int_0^x \exp(-0.5t^2)dt$ is the $N(0,1)$ cumulative distribution function, and $\mathcal{N}^{-1}$ is the inverse function to $\mathcal{N}$. The explicit expression for the default indicator variable in this model is $U_a(T) = \theta(b_a(T) - Z_a)$, where $\theta(x)$ is the step function, equal to 1 for $x \geq 0$ and 0 otherwise.

The full correlation matrix is unwieldy, obscures the source of correlations as coming from specific common factors and also makes non-Gaussian extensions less transparent. A factor representation is often superior, in which the correlation matrix is decomposed as $\rho_{ab} = \sum_i \beta_a^{(i)} \beta_b{}^{(i)}$. The numbers $\beta_a^{(i)}$ are called factor loadings. The factorized version of the fundamental simulation for the copula model takes the form

$$Z_a = \sum_i \beta_a^{(i)} X_i + \sqrt{1 - \vec{\beta}_a^2} Y_a. \tag{2.2.5}$$

For a Gaussian copula, each component X_i of the global factor and all drivers Y_a of individual risk factors are independent $N(0,1)$-distributed random variables. It is possible to obtain families of non-Gaussian modifications of the copula model using non-normal distributions for the factors.

Student's t-Distribution for Factors. A popular case involves the student's t-distribution with $\nu > 2$ degrees of freedom for the sake of its fatter tails (polynomial rather than exponential). The density of the student's t-distribution is given by

$$t_\nu(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\Gamma(\frac{\nu}{2})}\left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \tag{2.2.6}$$

where the normalization factor is expressed through the gamma function defined by

$$\Gamma(x) = \int_0^\infty t^{x-1} \exp(-t)dt.$$

Most generally, for a t-copula extension, we could independently choose between the Gaussian distribution and the student's t-distribution for each component of the global factor and each individual factor. We choose, however, to restrict the excessive freedom and use at most two different distributions, one for each component of the global factor (either Gaussian or t-distribution with ν_1 degrees of freedom), another for all individual risk factors (either Gaussian or t-distribution with ν_2 degrees of freedom). Note that we must rescale the t-distribution by a factor of $\sqrt{(\nu-2)/\nu}$ to make it univariate and preserve the meaning of the factor loadings β_a.

NIG distribution for factors. Another popular non-Gaussian choice is the Normal Inverse Gaussian (NIG) distribution. It is conventionally introduced through a rather complex probability density function with four parameters, α, β, δ and μ:

$$\text{NIG}_{\alpha\beta\mu\delta}(x) = \frac{\delta\alpha \exp\left[\delta\gamma + \beta(x-\mu)\right]}{\pi\sqrt{\delta^2 + (x-\mu)^2}} K_1\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right),$$

where K_1 denotes the modified Bessel function of the third kind and $\gamma = \sqrt{\alpha^2 - \beta^2}$. The validity domain of the parameters is $|\beta| < \alpha$, $\delta > 0$. The first four moments of the distribution,

$$\begin{aligned}
\text{mean} &= \mu + \frac{\beta\delta}{\gamma}, \\
\text{variance} &= \frac{\alpha^2\delta}{\gamma^3}, \\
\text{skew} &= 3\frac{\beta}{\alpha\sqrt{\gamma\delta}}, \\
\text{kurtosis} &= 3 + 3\left(1 + 4\frac{\beta^2}{\alpha^2}\right)\frac{1}{\gamma\delta},
\end{aligned}$$

uniquely define the parameters α, β, δ and μ under the condition $\text{kurtosis} > 3 + \frac{5}{3}\text{skew}^2$ and can be used as an alternative parameterization. Note that the mean of the distribution only affects the absolute position of the thresholds and thus is irrelevant. Also, the variance is rescaled to 1 so that the distribution width is governed by the factor loadings. Thus we only have two independent parameters to describe the distribution shape. Most naturally, this is done by specifying skew and kurtosis. The explicit specification only allows dropping the shift parameter μ, while the three remaining parameters α, β and δ are redundant in the sense that some formally different combinations are effectively identical.

RFL model. One more direction for improving the market matching ability of the copula models is the random factor loadings of Andersen and Sidenius [5]. In this approach, the factor loading β_a becomes a function of the global factor X. A user can set up a model with an arbitrary dependence $\beta(X)$, identical for all assets. The default indicator of the asset a is modeled as $\mathbb{I}_{Z_a<b_a}$ where

$$Z_a = \beta(X)X + (\nu + \mu Y_a) \tag{2.2.7}$$

where ν and μ are calibrated in such a way that Z_a has zero mean and unit variance, and b_a is calibrated to asset default probability. In particular, a simple step-like dependence can be used to model the effect of increasing correlation under adverse market conditions. The three parameters of this model are the threshold value of the global factor and the values of the factor loading below the threshold (bad economy) and above the threshold (good economy).

Computational techniques. Any version of factor formulation for the copula model suggests a two-stage procedure for the calculation of expectations [8]. At the first stage, the average is taken over the individual risk factors $\{Y_a\}$ conditionally on a realization of the global factor $\vec{X}$. This is facilitated by independence of the individual risk factors. The second stage is the integration over the global factor (which will be denoted $\mathbb{E}_{\vec{X}}[\ldots]$). This program works with any

choice of the distribution for the factors and any dimensionality of the global factor.

Following the two-stage procedure, we introduce conditional default probabilities $\mu_a(\vec{X})$ which are averages of the default indicators U_a conditional on a certain value of the global factor $\vec{X}$,

$$\mu_a(\vec{X}) = \mathbb{E}[U_a \,|\, \vec{X}]. \tag{2.2.8}$$

For the Gaussian copula, the conditional default is explicitly given by

$$\mu_a(\vec{X}) = \mathcal{N}\left(\frac{b_a - \vec{\beta}_a \vec{X}}{\sqrt{1-\vec{\beta}_a^2}}\right). \tag{2.2.9}$$

For a t-copula with ν_1 degrees of freedom for the scalar global factor, there is an explicit expression

$$\mu_a(X) = T_{\nu_1}\left(\sqrt{\tfrac{\nu_1}{\nu_1-2}}\frac{b_a - \beta_a X}{\sqrt{1-\beta_a^2}}\right), \tag{2.2.10}$$

where $T_\nu(x) = \int_{-\infty}^{x} t_\nu(s)ds$ is the distribution function. However, there is no convenient formula for the barrier b_a, which needs to be evaluated numerically.

The loss L conditional on the global factor $\vec{X}$ is a sum of independent Bernoulli random variables with two-state probability distributions $P_a(L|\vec{X})$,

$$P_a(L|\vec{X}) = \begin{cases} 1-\mu_a(\vec{X}) & \text{for} \quad L=0, \\ \mu_a(\vec{X}) & \text{for} \quad L=w_a. \end{cases} \tag{2.2.11}$$

The probability distribution of any sum of independent random variables is the convolution of the individual distribution functions. (For any two functions f_1 and f_2, their convolution, denoted $f_1 * f_2$, is the function $(f_1 * f_2)(x) = \int f_1(y) f_2(x-y)dy$. The convolution of more than two functions, $f_1 * f_2 * \cdots * f_N$, can be reduced to pairwise convolutions as $f_1 * (f_2 * (f_3 * \ldots * (f_{N-1} * f_N)\ldots))$ or in any other order.)

There are several ways to compute the convolution. The direct way is to use the definition and take the integrals one after another. This approach is especially convenient when all loss weights w_a are equal, or at least commensurate. In this case, the distribution of loss is discrete with the support which is a grid with not too many nodes. This approach becomes less efficient when the loss weights are incommensurate and the grid of possible values of loss is dense. It is particularly unsuitable for the calculation of the sensitivities to loss weights (necessary for VaR decomposition) because a small change in one of the weights gives rise to a radical change in the support of the distribution of loss.

Fortunately, there is a complementary approach to computing the convolution which works better right where the direct approach fails. This approach takes advantage of either Fourier or Laplace transforms to replace the convolution by the product of Fourier or Laplace images and then return to the original space by means of the inverse transform. This approach also serves as the basis for various semi-analytical approximations, including the central limit theorems and the saddle point.

The two-stage procedure of computing the distribution conditional on the global factor $\vec{X}$, and integrating over $\vec{X}$ after that, extends to the case of the joint distribution of loss from several

portfolios. Both the convolution approach and the Fourier/Laplace transform approach formally carry over to the multi-portfolio case as well. Unfortunately, the computational difficulty of the convolution approach grows dramatically because, instead of a linear grid of possible values of loss, the recursive scheme would have to operate on an M-dimensional lattice where M is the number of portfolios. This makes the numerical convolution practically unusable for more than two portfolios. The difficulties of the Laplace transform also increase substantially, so in practice it is necessary to resort to approximations based on central limit theorems.

The normal proxy method consists in computing the conditional average

$$\mathbb{E}\left[\left(\mathcal{L}(L_1, L_2, \cdots, L_M) - K\right)^+ \mid \vec{X}\right] \tag{2.2.12}$$

where

$$\mathcal{L}(L_1, L_2, \cdots, L_M) = \sum_p \left((L_p - k_p)^+ - (L_p - K_p)^+\right), \tag{2.2.13}$$

assuming that the joint distribution of portfolio losses L_p conditional on $\vec{X}$ is multivariate normal with the average

$$\Lambda_p(\vec{X}) = \sum_a \mu_a(\vec{X})\, w_{a,p} \tag{2.2.14}$$

and covariance matrix

$$C_{p\,p'}(\vec{X}) = \sum_a \mu_a(\vec{X})\left(1 - \mu_a(\vec{X})\right) w_{a,p} w_{a,p'}. \tag{2.2.15}$$

The appearance of the multivariate normal distribution is not related to the Gaussian choice for the copula but is a consequence of the central limit theorem. We further omit explicit dependence on the central factor $\vec{X}$ and focus on the calculation of the conditional average with the understanding that the entire procedure is repeated with all values of $\vec{X}$ necessary to get the unconditional average.

A direct computation of the average (2.2.12) over a multivariate normal distribution still requires a Monte Carlo simulation. The simulation noise can be significantly reduced if we split a part that is not correlated with other portfolios from each loss variable L_p, similar to the factor formulation (2.2.5) of the copula model. To this end, we decompose the portfolio loss as

$$L_p = \Lambda_p + D_p + \sigma_p W_p. \tag{2.2.16}$$

Here, $\{D_p\}$ are correlated normal variables that capture systematic portfolio risk. The variables $\{W_p\}$ are independent identically distributed $N(0,1)$ variables that are uncorrelated with $\{D_p\}$ and represent idiosyncratic risk factors for each individual portfolio. The amplitudes of the idiosyncratic terms σ_p are restricted by the requirement for the covariance matrix $\mathbb{E}[D_p D_{p'}]$ to have nonnegative eigenvalues. This requirement can be satisfied by taking all amplitudes equal to the smallest eigenvalue of the covariance matrix $C_{pp'}$.

The average of the variable D_p is zero due to the term Λ_p in the definition (2.2.16). The off-diagonal part of the covariance matrix $\mathbb{E}[D_p D_{p'}]$ is $C_{p,p'}$ while the variances are reduced,

$$\mathbb{E}[D_p^2] = C_{pp} - \sigma_p^2. \tag{2.2.17}$$

This reduces the simulation noise in the variables $\{D_p\}$ in comparison with the initial loss variables $\{L_p\}$, and makes it possible to perform the integration over $\{D_p\}$ using just a few points

$(D_1^{(i)}, D_2^{(i)}, \dots D_M^{(i)})$ in the M-dimensional space, adjusted to match the first and second moments.

The remaining problem is to compute the decorrelated average over independent Gaussian $N(0,1)$ variables W_p

$$\mathbb{E}\left[\left(\mathcal{L}(a_1+\sigma_1 W_1, \cdots, a_M+\sigma_M W_M)-K\right)^+\right], \tag{2.2.18}$$

where the shifted centers $a_p = \Lambda_p + D_p^{(i)}$ are dependent on the integration points $\{D_p^{(i)}\}$. Introducing an auxiliary variable y, the expectation over each W_p is decoupled,

$$\begin{aligned}\mathbb{E}\left[\left(\mathcal{L}(a_1+\sigma_1 W_1, \dots, a_M+\sigma_M W_M)-K\right)^+\right] \\ &= \frac{1}{2\pi i}\int_{C^+}\frac{1}{y^2}\exp(-yK)\prod_p \mathbb{E}\left[e^{y((a_p+\sigma_p W_p-k_p)^+-(a_p+\sigma_p W_p-K_p)^+)}\right]dy \\ &= \frac{1}{2\pi i}\int_{C^+}\frac{1}{y^2}\exp\left(\sum_p \log(\chi_p(y)) - yK\right)dy.\end{aligned}$$

Here

$$\chi_p(y) = \mathcal{N}(z_p^{(1)}) + e^{y\,(K_p-k_p)}\,\mathcal{N}(-z_p^{(2)}) + e^{y\,(a_p-k_p)+\frac{1}{2}y^2\,\sigma_p^2}\left(\mathcal{N}(z_p^{(2)}-y\sigma_p)-\mathcal{N}(z_p^{(1)}-y\sigma_p)\right), \tag{2.2.19}$$

where $z_p^{(1)} = \frac{k_p-a_p}{\sigma_p}$ and $z_p^{(2)} = \frac{K_p-a_p}{\sigma_p}$ are constants. The residual one-dimensional integral over y can be taken using the saddle point technique. This approach is referred to as the method with conditional decoupling and saddle point or as the normal limit.

2.2.2 QMP (Quasi-Multi-Period) Model

For the quasi-multi-period (QMP) model, at time T_0 there are N_0 assets that are observed at discrete times $T_1 < T_2 < \dots < T_f$. A set of scenarios, or paths, is generated. At each timestep, any number of assets can default, so at time T_k only N_k assets will be alive, with $N_0 \geq N_1 \geq N_2 \geq \cdots \geq N_f \geq 0$. Each defaulted asset is excluded from the subsequent steps of the simulation with a random recovery fraction distributed in $[0,1]$ according to a known distribution.

Available distributions of the recovery fraction include uniform, deterministic and beta distributions with given mean and standard deviation. In addition to the parameters of the recovery distribution, the market inputs include N_0 survival probability curves $\{p_i(t)\}$ for single-asset probabilities of default and an $N_0 \times N_0$ matrix of default correlations $\{\rho_{ij}\}$.

In each scenario, the defaults at time T_k are simulated as follows. Assume that the path has already been simulated up to T_{k-1}, which means that the number of remaining assets are known after each previous time step, up to N_{k-1} assets left at T_{k-1}. Let $\rho_{ij}^{(k-1)}$ be the $N_{k-1} \times N_{k-1}$ submatrix of the matrix ρ_{ij} that corresponds to the subset of N_{k-1} assets, which have not defaulted by the time T_{k-1}. Generate N_{k-1} normal numbers Z_i correlated with the matrix $\rho_{ij}^{(k-1)}$. The asset i defaults at time T_k if $Z_i > \mathcal{N}^{-1}(p_i(T_k|T_{k-1}))$ and stays alive otherwise. Here $p_i(T_k|T_{k-1}) = p_i(T_k)/p_i(T_{k-1})$ is the conditional probability of survival at T_k given survival at T_{k-1}, and $\mathcal{N}^{-1}$ is the inverse normal distribution function. This completes the definition of the model. Note that the survival thresholds $B_i(T_k) = \mathcal{N}^{-1}(p_i(T_k|T_{k-1}))$ are available in closed form. This will not be the case for the true multi-period model.

The model is defined in terms of an explicit algorithm that, in principle, can be used as described. However, for a basket of high-quality assets, the statistical weight of the paths with

defaults is small, which makes it impossible to achieve acceptable accuracy within a reasonable time. This problem is solved by using the techniques of importance sampling, where the integration measure is transformed in such a way that the contribution of the paths with defaults is factored into a product of a small analytically computable weight and a large factor determined from a simulation. Specifically, Numerix uses a measure transform where the resulting simulation has at least one default on each path. Note that the importance sampling is merely a technical tool and does not add any assumptions or approximations to the basic assumptions of the QMP model.

2.2.3 TMP (True Multi-Period) Model

The setting of the true multi-period (TMP) model is similar to that for the QMP model. The difference is in the criterion for deciding which of the N_{k-1} assets alive at time T_{k-1} default at the next time step T_k. For the QMP model, the decision is based on the positions of N_{k-1} correlated normal numbers Z_i with respect to a set of barriers determined by the conditional probabilities of survival $p_i(T_k|T_{k-1})$. The distribution of the random numbers Z_i is independent of the numbers used on previous steps. Therefore, the property that a *weak* asset that was close to default on one or several previous time steps has a higher probability of default on the present time step is not captured.

The TMP model seeks to remedy this deficiency by incorporating the history of evolution of each asset. To this end, the variable

$$V_i(T_k) = \sum_{q=1}^{k} Z_i(T_q)\sqrt{(T_q - T_{q-1})}$$

is tracked for each asset i until default or maturity, and its value beyond a certain barrier $B_i(T_k)$ is used as an indicator of default. The initial value is $V_i(T_0) = 0$. At each time T_q, the random normal numbers $Z_i(T_q)$ are correlated with the matrix $\rho_{ij}^{(q)}$, which is a restriction of the initial correlation matrix ρ_{ij} to the subset of non-defaulted assets, while the numbers $Z_i(T_q)$ at different times are uncorrelated.

It follows that the variables $V_i(T_k)$ are normal numbers with mean 0 and variance $T_k - T_0$ and are mutually correlated across assets and times. The survival thresholds $B_i(T_k)$ must be chosen to reproduce correct conditional probabilities of survival $p_i(T_k|T_{k-1})$. Because the correlations $\langle V_i(T_k), V_i(T_l)\rangle = \min(T_k, T_l)$ are non-vanishing, there is no analytical expression for the barriers $B_i(T_k)$, as the problem requires inverting the multivariate normal distribution. In the Numerix implementation, this inversion is performed by numerical matching of the conditional survival probability.

An equivalent form of the TMP model is discussed in [35].

2.2.4 Risk Measures

The value-at-risk (VaR) at a given confidence level α is the quantile of the loss distribution,

$$\mathrm{VaR}_\alpha(L) = \inf\{\ell \geq 0 \,|\, P[L \leq \ell] \geq \alpha\}. \tag{2.2.20}$$

For a continuous distribution, this would mean that the VaR is the threshold value of the loss such that only the fraction $1-\alpha$ of the outcomes leads to loss equal to or larger than the threshold. For a discrete distribution, the fraction of the outcomes above or at the level of VaR defined by Eq. (2.2.20) may be smaller than $1-\alpha$, but this effect of distribution granularity is essentially irrelevant for large portfolios.

The expected shortfall at a given confidence level α is the expected value of the loss in excess of $\mathsf{VaR}_\alpha(L)$,

$$\mathsf{ES}_\alpha(L) = \mathbb{E}\left[L \,|\, L \geq \mathsf{VaR}_\alpha(L)\right]. \tag{2.2.21}$$

(Note that the normalization denominator $1-\alpha$ is already included in the definition of conditional expectation.)

To assess the contributions of individual assets to VaR and ES, it is necessary to be able to compute sensitivities to asset weights, $\partial\mathsf{VaR}_\alpha(L)/\partial w_a$ and $\partial\mathsf{ES}_\alpha(L)/\partial w_a$. Note that VaR and ES are homogeneous of degree 1 with respect to portfolio loss weights w_a. That is,

$$\mathsf{VaR}_\alpha(L; v\cdot w_1, \ldots, v\cdot w_N) = v\cdot \mathsf{VaR}_\alpha(L; w_1, \ldots, w_N)$$

and similarly for ES. It follows that

$$\mathsf{VaR}_\alpha(L) = \sum_a w_a \frac{\partial \mathsf{VaR}_\alpha(L)}{\partial w_a} \tag{2.2.22}$$

with a similar decomposition of ES into marginal contributions from the individual underlying names. Explicit expressions are given in [8].

Currently, the functionality of risk measures is only available with the saddle point method solution of the Gaussian copula. Eventually, it should be expanded to include non-Gaussian copulas and methods other than the saddle point method.

2.3 Markovian Models for Credit Basket Loss Dynamics

In this section, we introduce a two-dimensional Markovian model of basket credit loss. This model is adequate for dynamic-sensitive instruments, such as options on CDO tranches and forward-starting CDOs. We assume that the initial number of the assets in the basket is N_{max} and model their defaults via a counting process N_t with the stochastic intensity λ_t. The dynamics of λ_t are defined through an axillary stochastic variable y_t that is related to the intensity by

$$\lambda_t = w(N)e^{y_t}, \tag{2.3.1a}$$

where $w(N)$ is the scaling factor

$$w(N) = 1 - \frac{N}{N_{\mathsf{max}}}. \tag{2.3.1b}$$

The variable y obeys the stochastic differential equation

$$dy_t = \kappa\left(\rho(N_t, t) - y_t\right)dt + \sigma dW_t, \tag{2.3.1c}$$

where W_t is the standard Brownian motion and the stochastic coefficients κ and σ control the reversion strength and the volatility, respectively. The drift term $\rho(N,t)$ in Eq.(2.3.1c) is a function of both the default number and time. For a fixed number of defaults (between the default events), the dynamics of the default intensity λ_t are similar to that of the short rate in the Black-Karasinski (BK) interest rate model. Arriving defaults, however, contribute to the drift term, thereby providing a back action effect on the stochastic intensity. Back action of the asset defaults on the stochastic intensity was introduced in the work of Errais et al. [28] in the context of a simplified affine credit model. The function $\rho(N,t)$ provides a necessary freedom for matching the distribution of the loss assumed by the market data on CDO tranches.

We assume that the value of loss-given-default (LGD) is equal to the same amount h for all assets in the basket. Many authors, including Brigo et al. [23] and Errais et al. [28], point out the importance of a non-trivial LGD distribution for the market-matching ability of their models. We believe that our model can describe the market data even in its simplest form, with a deterministic LGD, because sufficient flexibility is already built in via the function $\rho(N,t)$.

Note that calibration to the loss distribution will be achieved only by adjusting the function $\rho(N,t)$ in the drift term, but not the multiplier κ or the volatility σ in the diffusion term of Eq. (2.3.1c). The volatility term is kept available to tune the dynamics of the model. Given the potential for the growth in the variety and liquidity of dynamics-sensitive credit instruments, we can envision a scenario where the volatility will be calibrated to simpler instruments (for example, European tranche options) and then used to value more complex ones (for example, Bermudan tranche options). If necessary, constant volatility can be generalized to a term structure. This is similar to the calibration strategy for the classic short rate models of interest rates, including Hull-White (HW), Black-Karasinski, and Cox-Ingersoll-Ross (CIR). For these models, the term structure of volatilities is fitted to the options in a cycle of solver iterations, with the free function of time in the drift term being fitted to the discount curve inside each iteration.

When first introduced in the work [46], the model (3.9.1) used different—CIR-like—dynamics of stochastic intensity. Also, the scaling factor $w(N)$ in Eq.(2.3.1a) was not present. We found that BK-like dynamics provide better flexibility in the pricing of derivatives on senior tranches; it turns out to be more favorable in terms of implementation also. The factor $w(N)$ takes into account that the intensity of defaults scales with the number of survived assets. In particular, it ensures that the default intensity vanishes if all assets in the basket have defaulted. A similar scaling factor appears in the Arnsdorf and Halperin model (2007).

Being a two-dimensional Markovian model, the model (3.9.1) can be treated by any of the standard techniques provided the function $\rho(N,t)$ is fixed. The problem, however, is that the drift function $\rho(N,t)$ is unknown, and it needs to be set via calibration to the CDO tranches. Developing an efficient calibration algorithm thus becomes a crucial task for practical use of the model (3.9.1). An efficient calibration method was suggested and tested in the time-continuous case in the work [46]. Below, we show how this technique can be used in the more common discrete time schemes. In particular, we consider implementations of the calibration procedure on a two-dimensional tree and the case of Monte Carlo simulation.

2.3.1 Calibration Overview

Calibration of the model (3.9.1) aims to find the function $\rho(N,t)$ resulting in the loss distribution $P(L,t)$ that is consistent with the market data on CDO tranches. An effective calibration scheme based on the Markovian projection technique was suggested in [46]. The calibration procedure consists of two steps. First, one finds the so-called *local intensity* $\Lambda(L,t)$ consistent with the loss distribution $P(L,t)$. The local intensity is defined as the intensity of defaults in a simplified model of a one-dimensional Markov chain. The functions $P(L,t)$ and $\Lambda(L,t)$ can be related via the forward Kolmogorov equation,

$$\frac{\partial P(L,t)}{\partial t} = \mathbb{I}_{L\geq h}\Lambda(L-h,t)P(L-h,t) - \Lambda(L,t)P(L,t). \tag{2.3.2}$$

One can show that $P(L,t)$ uniquely defines $\Lambda(L,t)$. The market information on CDO tranches, however, is not complete enough for full determination of the distribution function $P(L,t)$, and, in practice, the local intensity surface is found by making some simplifying assumptions on its form and matching the CDO quotes via iterative solving of Eq.(2.3.2). This procedure will be considered in detail below. For now, we will assume that $\Lambda(L,t)$ is known.

The second step of the calibration procedure is based on the extension of Gyöngy's lemma (1987) to jump processes suggested in [46]. Gyöngy's lemma states that the unvaried loss distributions of the models (3.9.1) and (2.3.2) are the same if the expectation of the intensity λ_t conditional on the loss L_t calculated within the stochastic model (3.9.1) coincides with the local intensity of the Markov chain model (2.3.2). That is, the unvaried loss distributions of (3.9.1) and (2.3.2) are the same if

$$\Lambda(L,t) = \mathbb{E}[\lambda_t \,|\, L_t = L]. \tag{2.3.3}$$

This reduces the task of calibrating the model (3.9.1) to finding the drift function $\rho(L,t)$ which makes Eq.(2.3.3) valid for all losses and at all times. In the continuous-time approach, one can further proceed by expressing the function $\rho(L,t)$ through the full probability distribution function $p(L,\lambda,t)$ and then solving the forward Kolmogorov equation that describes the stochastic model (see [46]).

Below we show that the calibration method that is based on the Markovian projection can be used in the discrete time implementations as well. We also present the corresponding algorithms for the tree and Monte Carlo approaches.

2.3.2 Tree Construction for the Local Intensity Model

We begin with the formulation of the one-dimensional Markov chain model (2.3.2) on a tree. First, we fix the set of times $t_0, t_1, \ldots, t_N$ where t_0 is the start time and $t_N = T$ is the time horizon. At this point, we do not make any restrictions on the time intervals $\Delta_i = t_{i+1} - t_i$ apart from their positivity. In particular, we do not assume that these times are equally spaced. Restrictions on the maximal size of time intervals, however, will appear later from the requirement of numerical stability of the tree. Here and below, we prefer to use the number of defaults $N = L/h$ rather than the loss L as a state variable.

The discrete-time version of Eq.(2.3.2) can be written as

$$P(N,t_{i+1}) = \sum_{N' \geq 0}^{N_{\max}} P(N',t_i)\,\hat{T}(N',N,t_i), \tag{2.3.4}$$

where the transition matrix $\hat{T}$ has the nonzero elements

$$\hat{T}(N,N',t_i) = \begin{cases} \Delta_i\,\Lambda(N,t_i), & N' = N+1, \\ 1-\Delta_i\,\Lambda(N,t_i), & N' = N, \end{cases} \tag{2.3.5}$$

corresponding to the probability to default ($N \to N' = N+1$) and to survive ($N \to N' = N$) within the time interval Δ_i. The requirement of positivity of the survival probability imposes the restriction

$$\Delta_i < 1/\Lambda_i \tag{2.3.6}$$

on the time intervals.

2.3.3 Tree Construction for the Stochastic Model

Now we turn to constructing a two-dimensional tree for the stochastic model (3.9.1). We assume that the nodes of the tree are synchronized in time with those of the one-dimensional tree introduced above, so they belong to the same set $t_0, t_1, \ldots, t_N$. Constructing the tree, one also has to discretize the logarithmic intensity y. We choose the simplest equally-spaced grid

$$y_p = p\,\delta, \tag{2.3.7}$$

where p is an integer. In order to simplify our presentation, we will assume that p runs from $-\infty$ to ∞. In practice, one has to choose some upper and lower bounds such that the probably for values y to fall outside the chosen region can be neglected within the required accuracy. Under these assumptions, the time evolution of the two dimensional distribution $p(\lambda, N)$ is described via the equation

$$p(y,N,t_{i+1}) = \sum_{N'=0}^{N_{\max}} \sum_{y'} \; p(y',N',t_i)\;\hat{T}(y',N';y,N;t_i), \tag{2.3.8}$$

where the transition matrix $\hat{T}$ is yet to be set. Apparently, there are infinitely many tree constructions that converge to the description of the stochastic dynamics (3.9.1) in the limit $\Delta_i \to 0$, so the transition matrix $\hat{T}$ is not defined uniquely at finite Δ_i. Constructing this matrix, we choose to present it in the form of a matrix product

$$\hat{T}(t_i) = \hat{T}_D(t_i) \cdot \hat{T}_Y(t_i), \tag{2.3.9}$$

where the matrix $\hat{T}_D$ is diagonal in the index y, while T_Y is diagonal in the index N. The matrix product denoted by $\cdot$ is defined by

$$A(y_1,N_1;y_2,N_2) = \sum_{N=0}^{N_{\max}} \sum_{y} B(y_1,N_1;y,N)\,C(y,N;y_2,N_2), \tag{2.3.10}$$

for $A = B \cdot C$. The meaning of Eq.(2.3.9) is straightforward: The transition matrix $\hat{T}_D$ is responsible for the default transitions only, while $\hat{T}_Y$ is responsible for the stochastic dynamics of the variable y defined in Eq.(2.3.1c). The elements of $\hat{T}_D$ can be defined similarly to Eq.(2.3.5) with the difference being that the intensity of the defaults is now given by Eq. (2.3.1a):

$$\hat{T}_D(y, N; y', N'; t_i) = \mathbb{I}_{y=y'} \begin{cases} \Delta_i w(N) \exp(y), & N' = N + 1, \\ 1 - \Delta_i \, w(N) \exp(y), & N' = N. \end{cases} \tag{2.3.11}$$

The stochastic process described by Eq. (2.3.1c) is well known, and constructing the corresponding transition matrix $\hat{T}_Y$ seems to be a standard task. However, our case is more subtle because the drift function $\rho(N, t)$ is unknown, and it needs to be set at each time step of the calibration procedure. For this purpose, it is convenient to present the matrix $\hat{T}_Y$ as

$$\hat{T}_Y(t_i) = \hat{T}_S(t_i) \cdot \hat{T}'_Y(t_i), \tag{2.3.12}$$

where $\hat{T}_S(t_i)$ corresponds to the shift term $\kappa\rho(N, t)dt$ in Eq.(2.3.1c) while the matrix $\hat{T}'_Y(t_i)$ describes the transitions related to the term $-y_t\kappa dt + \sigma dW_t$. (Both $\hat{T}_S(t_i)$ and $\hat{T}'_Y(t_i)$ are diagonal in the index N.) Now the matrix T'_Y can be constructed in a standard way. For example, in the simplest trinomial implementation, the nonzero elements of T'_Y can be chosen as

$$\hat{T}'_Y(y_p, N; y_{p'}, N'; t_i) = \mathbb{I}_{N=N'} \times \begin{cases} \frac{\Delta_i \sigma^2}{2\delta^2} - \frac{\kappa \Delta_i y_p}{\delta} \, \mathbb{I}_{p<0}, & p' = p + 1, \\ 1 - \frac{\Delta_i \sigma^2}{\delta^2} - \frac{\kappa \Delta_i |y_p|}{\delta}, & p' = p, \\ \frac{\Delta_i \sigma^2}{2\delta^2} + \frac{\kappa \Delta_i y_p}{\delta} \, \mathbb{I}_{p>0}, & p' = p - 1. \end{cases} \tag{2.3.13}$$

The transition matrix $\hat{T}_S$, which implements the shift operator $\kappa\rho(N, t)dt$ in Eq.(2.3.1c), can be constructed as follows: Suppose we want to shift the distribution function $p(y, N, t_i)$ by the value $a(N, t_i)$ in the y-direction. If $a(N, t_i)$ is a multiple of δ, this shift is simply implemented via the transitions $y_p \to y_{p'} = y_p + a(N, t_i)$ taking place with probability 1. In the general case of an arbitrary $a(N, t_i)$, the result of the shift operation must be redistributed in between the two adjacent values $y_{p'}, y_{p'+1}$ that surround the value $y_p + a(N, t_i)$. The corresponding weights are naturally taken to be proportional to the distances between the point $y_p + a(N, t_i)$ and $y_{p'}, y_{p'+1}$. Thus, we define the elements of the transition matrix $\hat{T}_S$, as

$$\hat{T}_S(N, y_p; N', y_{p'}; t_i) = \mathbb{I}_{N=N'} \times \begin{cases} a_N/\delta - \lfloor a_N/\delta \rfloor, & p' = p + \lfloor a_N/\delta \rfloor + 1, \\ 1 - a_N/\delta + \lfloor a_N/\delta \rfloor, & p' = p + \lfloor a_N/\delta \rfloor. \end{cases} \tag{2.3.14}$$

The shift function $a(N, t_i)$ is a discrete-time analog of the drift term $\kappa\rho(N, t)$, and it needs to be set during the calibration procedure.

2.3.4 Stochastic Model Calibration

The calibration of the stochastic model aims to ensure the consistency of the unvaried loss distributions that are calculated the local and stochastic models,

$$P(N,t_i)=\sum_k p(y_k,N,t_i), \tag{2.3.15}$$

by means of choosing the proper shift function $a(N,t_i)$ in the stochastic model. Below we present a simple bootstrapping algorithm that allows determination of the shift function $a(N,t_i)$ step by step for each time node t_i. As in the time-continuous case, the algorithm is based on the method of Markovian projection.

We begin with derivation of a "tree" analog of Eq. (2.3.3) which ensured consistency of the unvaried loss distributions in the time continuous case. Substituting Eq. (2.3.9) into Eq. (2.3.8), one can write the relation between the probability distributions on the consecutive time steps:

$$p(y,N,t_{i+1})=\sum_{N'=0}^{N_{\max}}\sum_{y'} p(y',N',t_i)\;\hat{T}_D(y',N';y'N;t_i)\,\hat{T}_Y(y',N;y,N;t_i). \tag{2.3.16}$$

Summing both parts of this equation over y and using $\sum_y \hat{T}_Y(y',N;y,N)=1$, we obtain

$$\sum_y p(y,N,t_{i+1})=\sum_{N'=0}^{N_{\max}}\sum_{y'} p(y',N',t_i)\;\hat{T}_D(y',N';y'N;t_i). \tag{2.3.17}$$

Comparing the above equation with Eq. (2.3.4) and taking into account definitions (2.3.5) and (2.3.11), one concludes that Eq. (2.3.15) is satisfied if

$$\Lambda(N,t_i)=\frac{w(N)\sum_y e^y p(y,N)}{\sum_y p(y,N)}. \tag{2.3.18}$$

This equation can be viewed as a straightforward extension of the continuous-time result (2.3.3) to the discrete-time case. Indeed, the right-hand side of this equation represents the expectation of the stochastic intensity (2.3.1a) conditional on the number of defaults.

Now we will show that Eq. (2.3.18) can always be satisfied by choosing proper shifts $a(N,t_i)$ on each step of the bootstrapping procedure. Suppose that the model is calibrated until time t_i. This means that, for all $t_j\le t_i$, the distribution function is known and conditions (2.3.15) and (2.3.18) are satisfied. Let us write Eq. (2.3.8) in a compact matrix form

$$\bar{\vec{p}}(t_{i+1})=\bar{\vec{p}}(t_i)\cdot\hat{T}_D(t_i)\cdot\hat{T}'_Y(t_i)\cdot\hat{T}_S(t_i), \tag{2.3.19}$$

where the vector $\vec{p}(t_i)$ represents the two dimensional distribution function $p(y,N,t_i)$ and the bars over vectors denote transpositions. Now Eq. (2.3.15) is guaranteed to be satisfied at time t_{i+1} because condition (2.3.18) was assumed to be valid at t_i. Thus, we need to show that Eq. (2.3.18) can be satisfied at time t_{i+1} via choosing the proper shift function $a(N,t_i)$. Let us consider the distribution function $\vec{p}(t_i)$ in Eq. (2.3.19) after being multiplied by the transition matrices $\hat{T}_D$ and $\hat{T}'_Y$:

$$\bar{\vec{p}}'(t_i)=\bar{\vec{p}}(t_i)\cdot\hat{T}_D(t_i)\cdot\hat{T}'_Y(t_i). \tag{2.3.20}$$

The expected default intensity, conditional on default number N and calculated with respect to this function,

$$\Lambda'(N,t_i)=\frac{w(N)\sum_y e^y p'(y,N,t_i)}{\sum_y p'(y,N,t_i)}, \tag{2.3.21}$$

does not coincide with $\Lambda(N,t_{i+1})$ in general, and the shift function $a(N,t_i)$ has to be chosen so that $\bar{\vec{p}}(t_{i+1})=\bar{\vec{p}}'(t_i)\cdot\hat{T}_S(t_i)$ will satisfy Eq. (2.3.18). One can check that this can always be done and that the corresponding values of the shifts are

$$a(N,t_i)=S\delta+\frac{1}{e^\delta-1}\left(\frac{\Lambda(N,t_i)}{\Lambda'(N,t_i)}\,e^{-S\,\delta}-1\right), \tag{2.3.22}$$

where the integer $S=\left\lfloor(1/\Delta_\lambda)\ln(\Lambda(N,t_i)/\Lambda'(N,t_i))\right\rfloor$.

The bootstrap calibration of the stochastic model on a two-dimensional tree can be summarized as follows:

1. Starting with the distribution function $\vec{p}(t_0)$ that is assumed to satisfy Eqs. (2.3.15) and (2.3.18), build the distribution function $\vec{p}'(t_0)$ according to Eq. (2.3.20).
2. Find the shift values according to Eqs. (2.3.21) and (2.3.22).
3. Obtain the distribution function on the next time step $\bar{\vec{p}}(t_1)=\bar{\vec{p}}'(t_0)\cdot\hat{T}_S$.
4. Advance the calibration, repeating steps 1–3 until the desired time horizon is reached.

2.3.5 Monte Carlo Implementation

In this section, we present a simulation approach to model (3.9.1) that implements an efficient calibration to the distribution of defaults $P(N,t)$. This method can be used to price path-dependent instruments.

As in the tree approach, the simulation algorithm is formulated on the set of times $t_0,t_1,\ldots,t_N$ of the one-dimensional Markov chain model. At the same time, the variable y does not need to be discretized, and this leads to some simplifications in constructing the algorithm.

We define a simulation pattern $\mathcal{G}_i$ as a set of N_p paths generated from time t_0 to t_i. Each path is formed by the values of the state variables $\{y,N\}$ on the time nodes $t_0,t_1,\ldots,t_i$. The kth path is represented by the set

$$\{y_k(t_0),N_k(t_0)\},\ \{y_k(t_1),N_k(t_1)\},\ldots,\{y_k(t_i),N_k(t_i)\}. \tag{2.3.23}$$

We also define a simulation pattern slice $\mathcal{S}_i$ as a set of N_p state variables

$$\{y_1(t_i),N_1(t_i)\},\ \{y_2(t_i),N_2(t_i)\},\ldots,\{y_{N_p}(t_i),N_{N_p}(t_i)\} \tag{2.3.24}$$

at a fixed time node t_i. The simulation pattern $\mathcal{G}_i$ can, thus, be viewed as a set of pattern slices $\mathcal{S}_1,\mathcal{S}_2,\ldots,\mathcal{S}_i$.

A simulation pattern can be constructed step by step in time. Furthermore, since the model (3.9.1) describes a Markov process, the pattern evolution between the consecutive time steps t_i and t_{i+1} can be formulated in terms of pattern slices

$$\mathcal{S}_{i+1}=\hat{\mathcal{T}}(t_i)\,\mathcal{S}_i, \tag{2.3.25}$$

where the operator $\hat{\mathcal{T}}(t_i)$, which will be specified below, implements the necessary transformations under the slice $\mathcal{S}_i$ in accordance with stochastic dynamics of the model (3.9.1). In analogy with Eqs. (2.3.9, 2.3.12), we present the operator $\hat{\mathcal{T}}$ as an operator product,

$$\hat{\mathcal{T}}(i) = \hat{\mathcal{T}}_S(t_i)\,\hat{\mathcal{T}}'_Y(t_i)\,\hat{\mathcal{T}}_D(t_i), \tag{2.3.26}$$

where the operator $\hat{\mathcal{T}}_D(t_i)$ is responsible for the default transitions, while the operators $\hat{\mathcal{T}}'_Y, \hat{\mathcal{T}}_S$ together describe the stochastic dynamics of the variable y given by Eq. (2.3.1c) with $\hat{\mathcal{T}}_S$ taking care of the drift term $\kappa\rho\,dt$. The operator $\hat{\mathcal{T}}_D$, similarly to the matrix $\hat{T}_D$ defined in Eq. (2.3.11), implements the default transitions

$$\hat{\mathcal{T}}_D: \quad N_k(t_i) \to N_k(t_{i+1}) = N_k(t_i) + 1, \tag{2.3.27}$$

taking place with probability

$$w(N_k(t_i))\,\Delta_i\,e^{y(t_i)}. \tag{2.3.28}$$

Stochastic dynamics of the variable y defined in Eq. (2.3.1c) at a fixed default number coincides with the short rate-dynamics in the Hull-White interest rate model. Making use of this analogy, we define the operator $\hat{\mathcal{T}}'_Y$ as

$$\hat{\mathcal{T}}'_Y: \quad y_k(t_i) \to y_k(t_{i+1}) = y_k(t_i) + \Sigma_i\,\xi_k(t_i), \tag{2.3.29}$$

where $\xi_k(t_i)$ is the normally distributed random variable with unit variance, and the dispersion Σ_i is given by

$$\Sigma_i^2 = \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa(t_i - t_0)}\right). \tag{2.3.30}$$

Finally, the shift operator $\hat{\mathcal{T}}_S$ has the form

$$\hat{\mathcal{T}}_S: \quad y_k(t_i) \to y_k(t_{i+1}) = y_k(t_i) + a(N, t_i), \tag{2.3.31}$$

where the shift values $a(N, t_i)$ must be set during the model calibration as in the case of the tree approach.

The calibration procedure of the simulation model is similar to that used in constructing the tree: The shift function $a(N, t_i)$ in Eq. (2.3.31) has to be chosen so that the expectation of the stochastic intensity, conditioned on the number of defaults calculated within the stochastic model, will coincide with the local intensity $\Lambda(N, t_i)$ obtained within the Markov chain model:

$$\Lambda(N, t_i) = \frac{w(N)\sum\limits_{k=1}^{N_p} \exp(y_k(t_i))\,\mathbb{I}_{N_k(t_i)=N}}{\sum\limits_{k=1}^{N_p} \mathbb{I}_{N_k(t_i)=N}}. \tag{2.3.32}$$

To find the corresponding shifts, we construct the slice

$$\mathcal{S}'_i = \hat{\mathcal{T}}'_Y(t_i)\,\hat{\mathcal{T}}_D(t_i)\,\mathcal{S}_i, \tag{2.3.33}$$

and find the corresponding conditional expectation of the default intensity

$$\Lambda'(N,t_i) = \frac{w(N)\sum\limits_{k=1}^{N_p} \exp(y'_k(t_i))\,\mathbb{I}_{N'_k(t_i)=N}}{\sum\limits_{k=1}^{N_p} \mathbb{I}_{N'_k(t_i)=N}}. \tag{2.3.34}$$

The values of the shifts that make Eq. (2.3.32) valid now can be expressed as

$$a(N,t_i) = \ln\left(\frac{\Lambda(N,t_i)}{\Lambda'(N,t_i)}\right). \tag{2.3.35}$$

The simulation procedure can be summarized as follows:

1. Beginning with the initial pattern slice $\mathcal{S}_0$, construct the slice $\mathcal{S}'_0$ according to Eq. (2.3.33).
2. Find the shift function $a(N,t_i)$ according to Eq.(2.3.35).
3. Find the pattern slice on the next time step, applying the shifts $a(N,t_i)$ to $\mathcal{S}'_0$: $\mathcal{S}_1 = \hat{T}(t_i)\,\mathcal{S}'_0$.
4. Advance the simulation by repeating steps 1–3 till maturity is reached.

Finally, we show that by enforcing relation (2.3.32) at each time step of the simulation algorithm, we guarantee that the distribution of defaults obtained in simulation,

$$P_{\mathsf{MC}}(N,t_i) = \frac{1}{N_p}\sum_{k=1}^{N_p} \mathbb{I}_{N'_k(t_i)=N}, \tag{2.3.36}$$

converges to the distribution of defaults of the Markov chain model, $P(N,t_i)$, in the limit of large number of paths, $N_p \to \infty$. Assuming that the pattern slice at time t_i is fixed, let us consider the expected number of paths with $N(t_i) = N$ at the next time step of the simulation

$$\begin{aligned}\mathbb{E}[\,N_p\,P_{\mathsf{MC}}(N,t_{i+1})\,|\,\mathcal{S}_i\,] &= \mathbb{I}_{N>0}\,\Delta_i w(N-1)\sum_{k=1}^{N_p} e^{y_k}\mathbb{I}_{N_k(t_i)=N-1} + \sum_{k=1}^{N_p}[1-\Delta_i w(N)e^{y_k}]\mathbb{I}_{N_k(t_i)=N} \\ &= N_p\,\mathbb{I}_{N>0}\Delta_i\Lambda(N-1,t_i)\,P_{\mathsf{MC}}\,(N-1,t_i) + N_p\,(1-\Delta_i\Lambda(N,t_i))\,P_{\mathsf{MC}}(N,t_i).\end{aligned} \tag{2.3.37}$$

Dividing all terms in the above equation by N_p and taking the limit as $N_p \to \infty$, we obtain

$$P_{\mathsf{MC}}(N,t_{i+1}) = \mathbb{I}_{N>0}\Delta_i\Lambda(N-1,t_i)\,P_{\mathsf{MC}}\,(N-1,t_i) + (1-\Delta_i\Lambda(N,t_i))P_{\mathsf{MC}}(N,t_i) \tag{2.3.38}$$

as $N_p \to \infty$, which coincides with the Markov chain equations (2.3.4) and (2.3.5). Therefore, we have

$$\lim_{N_p\to\infty} P_{\mathsf{MC}}(N,t_i) = P(N,t_i). \tag{2.3.39}$$

2.3.6 From Market Data to Local Intensity

We now turn to the calibration of the local intensity $\Lambda(N,t_i)$ to the actual market data, that is, to single tranche CDOs. The market data on CDO tranches obviously cannot be complete enough for unique determination of the local intensity surface. The liquid tranche quotes are usually

available for a handful of maturity points only. Therefore, to calibrate $\Lambda(N,t_i)$, one needs to assume a certain functional form for $\Lambda(N,t_i)$ and do a parametric fit to the tranches and the index. The local intensity model turns out to be a rather powerful tool for interpolating the CDO tranche spreads (see [15], [46]).

In the case of constant LGDs, the spreads of the CDO tranches are determined completely by the distribution function of defaults $P(N,t_i)$. The distribution function of defaults, in turn, is determined by $\Lambda(N,t_i)$ through Eqs. (2.3.4) and (2.3.5), which we can write explicitly as

$$P(N,t_{i+1}) = (1-\Delta_i\,\Lambda(N,t_i))\,P(N,t_i) + \mathbb{I}_{N>0}\Delta_i\,\Lambda(N-1,t_i)\,P(N-1,t_i). \tag{2.3.40}$$

We note that the spreads of the index are determined completely by the time dependence of the expected number of defaults

$$\bar{N}(t_i) = \sum_{N=0}^{N_{\text{max}}} N\,P(N,t_i). \tag{2.3.41}$$

Therefore, given the index spreads at different maturities and assuming a suitable interpolation scheme one can recover $\bar{N}(t_i)$. Below, we assume that such procedure was implemented and that $\bar{N}(t_i)$ is known.

Prior to setting a particular parametrization for the local intensity surface, it is convenient to present it in the form

$$\Lambda(N,t) = w(N)\,(\alpha_0(t)+\alpha(N,t)), \tag{2.3.42}$$

where the function $\alpha(N,t)$ is defined such that $\alpha(0,t)=0$ for all t. The factor $w(N)$ captures the effect that the intensity of defaults in a CDO basket naturally scales with the magnitude of the number of survived assets, $N-N_{\text{max}}$ (see [15]).

The function $\alpha_0(t)$ in Eq.(2.3.42) can be expressed explicitly through $\bar{N}_t$ and $P(N,t)$ as

$$\alpha_0(t_i) = \frac{1}{w(\bar{N})}\left(\frac{\bar{N}(t_{i+1})-\bar{N}(t_i)}{\Delta_i} - \sum_{N=0}^{N_{\text{max}}} w(N)\,\alpha(N,t_i)\,P(N,t_i)\right). \tag{2.3.43}$$

The above equation can be derived by multiplying both parts of Eq. (2.3.40) by N and then taking the sum over N from 0 to N_{max}.

Expression (2.3.43) allows finding $\alpha_0(t_i)$ at each time t_i when solving Eq. (2.3.40) step by step in time. This means that function $\alpha_0(t)$ is determined completely by the expected number of defaults $\bar{N}_t$ and the function $\alpha(N,t)$ in Eq. (2.3.42).

Now we turn to the parametrization of the function $\alpha(N,t)$. Suppose that the tranche spreads are defined on the set of maturity times $T_1,T_2,\ldots,T_M$. (Each maturity time is assumed to belong to the set of times of the tree.) The simplest form of the time dependence $\Lambda(N,t_i)$ is piecewise constant,

$$\Lambda(N,t_i) = \begin{cases} \Lambda_1(N), & 0 \le t_i < T_1, \\ \Lambda_2(N), & T_1 \le t_i < T_2, \\ \quad\vdots & \\ \Lambda_M(N), & T_{M-1} \le t_i < T_M. \end{cases} \tag{2.3.44}$$

This form allows setting up a bootstrap calibration to CDO tranches with different maturities. First, one calibrates the model to tranches with maturity T_1, which determines $\Lambda_1(N)$. Then, the model is calibrated to tranches which mature at T_2 by adjusting $\Lambda_2(N)$ and keeping $\Lambda_1(N)$ unchanged, and the procedure is repeated until the final maturity is reached.

To determine $\{\Lambda_i(N)\}$, we first define the vector of losses that corresponds to the detachment points of the set of N_T tranches $l_1, l_2, \ldots, l_{N_T}$. Taking each $\Lambda_i(N)$ to be a constant would lead to a problem because, due to the condition $\alpha(0, t_i) = 0$, one would be forced to take $\Lambda_i(N) = 0$ for the loss interval $(0, l_1)$ that corresponds to the equity tranche. Therefore, we will take a piecewise-linear form for Λ. Introduce the values of the local intensities $\Lambda_1, \Lambda_2, \ldots, \Lambda_{N_T}$ that correspond to the values of losses $l_1, l_2, \ldots, l_{N_T}$. Then define $\Lambda_i(N)$ to be the piecewise linear interpolant

$$\Lambda_i(N) = \Lambda_{i-1} \frac{l_i - hN}{l_i - l_{i-1}} + \Lambda_i \frac{hN - l_{i-1}}{l_i - l_{i-1}}, \qquad l_{i-1} \le Nh < l_i, \tag{2.3.45}$$

where $\Lambda_0 = l_0 = 0$. The values $\Lambda_1, \Lambda_2, \ldots, \Lambda_{N_T}$, however, cannot be found via a bootstrapping scheme because, in adjusting the value of Λ_i, one modifies the solution for $\alpha_0(t_i)$, which breaks the calibration within the preceding loss intervals. Therefore, the values of $\Lambda_1, \Lambda_2, \ldots, \Lambda_{N_T}$ have to be found with the help of a multi-dimensional solver.

The form of the local intensity described by Eqs. (2.3.42), (2.3.44) and (2.3.45) is almost identical to the one used by Arnsdorf and Halperin (2007) with the difference that they do not introduce the term $\alpha_0(t)$ and treat the index as (0–100%) tranche during the calibration. Our scheme allows performing the calibrations to the index and to the tranches separately, and it is thus more suitable in cases where the tranche quotes are unknown at certain maturities where index spreads are available.

The above presented parametrization of the local intensity $\Lambda(N, t)$ works the best when the number of available tranche quotes with different maturities is moderately large. In this case, one can indeed neglect the time dependence of the local intensity within the intervals between the adjacent maturities. This approximation, however, is justified only by the smallness of these intervals, and the calibration to tranches can fail to produce a good fit if the tranche quotes at some intermediate maturities are not available.

An approach that aims to mimic the dynamics of the local intensity in between the tranche maturities was suggested in [46]. In this case, one represents the function $\alpha(N, t)$ as a series expansion in the parameter x,

$$\alpha(N, t_i) = \sum_{p=1}^{N_T} \alpha_p(t_i)\, x^p, \tag{2.3.46}$$

where x is defined as

$$x = \frac{N}{\bar{N}(t_i) + z(t_i)}. \tag{2.3.47}$$

The function z is introduced to regularize the singularity as $t \to 0$ in Eq. (2.3.47). Specifically, we used an exponential function,

$$z(t) = \exp(-\gamma t), \tag{2.3.48}$$

but any monotonic function that decays sufficiently fast with time could be used instead. A representation of the local intensity in terms of the parameter x ensures that for $t \gg \gamma^{-1}$ the local intensity becomes a function of the number of defaults, N, normalized by the expected number of defaults, $\bar{N}(t_i)$. This normalization reflects the fact that the typical number of defaults

naturally grows with time even in the case where the typical local intensity stays constant in time. The main dependence of the local intensity on time is contained in the parameter x. A residual dependence on time in the coefficients $\alpha_p(t_i)$ is included to ensure the matching with the initial condition at $t = 0$, and also for the fit of tranches with different maturities.

The form of the local intensity at the start time can hardly be extracted from the market data, and therefore it is natural to take it to be independent (up to the dependence in the scaling factor $w(N) = w(L/h)$) on the loss L:

$$\Lambda(N, t_0) = w(N)\lambda_0, \tag{2.3.49}$$

where $\lambda_0 = \bar{N}(t_1)/\Delta_i$ is the initial expected default intensity. This initial condition can be satisfied by taking the coefficients α_p in the form

$$\alpha_p(t_i) = \alpha'_p(t_i)\,(1 - z(t_i)), \qquad p \geq 1. \tag{2.3.50}$$

The numerical values of the coefficients $\alpha'_p(t_i)$ can be fitted to a set of tranches with a single maturity. Simultaneous fit to tranches with different maturities can be achieved using an additional term structure of the coefficients α'_p.

2.3.7 Multiname Dynamic Model

The multiname dynamic model[1] is a simple dynamic bottom-up model that has sufficient flexibility for simultaneous calibration to the market of credit default swaps (CDSs) and CDO tranches. The multiname dynamic model is capable of evaluating tranche hedge ratios as well as simple instruments whose values can be expressed through a one-dimensional marginal distribution of the portfolio loss. The model is defined by specifying the form of the default intensity for each underlying asset as a function of time t and the number of defaults $N(t)$ that have been accumulated in the credit portfolio up to t.

The choice of the state variable can be better understood by examining the simplest model within the top-down framework. That model, which in analogy to the local volatility model is sometimes called the local intensity model, represents the one-dimensional Markov chain, where default transitions $N \to N + 1$ are governed by the intensity $\lambda(N(t), t)$, which is a deterministic function of N and t. Remarkably, it can be shown that any top-down credit model with default intensity given by an adaptive stochastic process of a general kind can be brought to the form of the local intensity model as long as instruments under consideration can be expressed through the one-dimensional marginal distribution of the portfolio loss [46]. Thus, a "minimal" multiname dynamic model can be constructed as an extension of the local intensity model by choosing the default intensity of each name to be a deterministic function of $N(t)$.

The multiname dynamic model can also be viewed as a special case of the default contagion model [25]. In the default contagion framework, the default intensity of the kth asset is given by the deterministic function $\lambda_k(\mathbf{n}(t), t)$ of the portfolio state represented by the vector of asset default indicators $\mathbf{n} = (n_1, n_2, ..., n_{N_0})$, where N_0 is the number of assets in the portfolio. Default of the kth asset thereby affects the default intensities of other assets (typically by increasing those default intensities). In its general form, the default contagion model contains a huge

[1] We present only a brief and incomplete description here. A full description can be found in [45].

number of free parameters and, in practice, one deals with default intensities of a particular functional form. Examples of specific models that belong to this framework can be found in papers by Davis & Lo (2001), Jarrow & Yu (2001) and Laurent (2007). The Numerix model corresponds to the special case where each default intensity depends on the total number of defaults $N = n_1 + n_2 + \cdots + n_{N_0}$.

Default contagion models are Markovian and they therefore support the application of many standard numerical techniques, including forward/backward induction and Monte Carlo simulations.

Model Definition

Under the standard assumption of constant recovery coefficients, stripping each CDS allows reproducing the survival probabilities of individual portfolio members. This procedure relies on time interpolation of the implied default intensity, which, in the simplest case, is taken to be piecewise constant. Thus, we assume that the kth portfolio member is characterized by the survival probability $p_k(t)$ and the recovery coefficient R_k. We postulate the default intensity of the kth asset to be

$$\lambda_k(t) = a_k(t) + b_k(t)\, Y(N(t), t), \tag{2.3.51}$$

where $N(t)$ is the number of defaults accumulated in the portfolio up to time t. The function $Y(N,t)$ is the common (systemic) component that serves to model the correlation of default events. The functions $a_k(t)$ and $b_k(t)$ are specific to the kth asset. For a given $Y(N,t)$, functions $a_k(t)$ and $b_k(t)$ are used to fit the model to the survival curves of individual assets. The function $Y(N,t)$ provides a degree of freedom for calibrating the model to CDO tranches. It is clear, however, that the model (2.3.51) has too many free parameters (in addition to the freedom in interpolation of the function $Y(N,t)$) since both functions $a_k(t)$ and $b_k(t)$ cannot be defined uniquely by fitting a single survival curve. For this reason, a particular specification of model (2.3.51) containing is chosen where $a_k(t) = 0$. Thus, there is no extra parametric freedom and the intensity of the kth asset becomes a product of the idiosyncratic component $b_k(t)$ and the systemic component, $Y(N,t)$. This guarantees the positivity of all default intensities by construction.

Model Calibration

The model (2.3.51) has to be calibrated to the market data on CDO tranches and individual CDSs. The functions $b_k(t)$ are used to fit single-name spreads, while the spreads of CDO tranches are adjusted via the function $Y(N,t)$. Calibration to single-name survival curves is performed during the numerical integration of the Focker-Planck equation [45]. Calibration to CDO tranches represents a mixture of bootstrapping and iterative fitting.

Consider calibrating to CDO tranches in the simplified case of a portfolio with homogeneous recovery coefficients $R_k = R$ and notionals $A_k = A$. We express the portfolio loss through the number of defaults as $L = hN$, where $h = A(1 - R)$ is the loss given default. The value of a CDO tranche in this case can be computed directly from the probability density of the number of defaults, $P_B(N,t)$.

We parameterize the surface described by $Y(N,t)$ via the linear interpolation in loss between

the values $Y_k(t)$ defined at the loss levels $l_p(t)$:

$$Y(N,t) = Y_{p-1}(t)\,\frac{l_p(t) - Nh}{l_p(t) - l_{p-1}(t)} + Y_p(t)\,\frac{Nh - l_{p-1}(t)}{l_p(t) - l_{p-1}(t)}, \qquad l_{p-1} \le Nh < l_p. \tag{2.3.52}$$

In the simplest case, the functions $l_k(t)$ can be set to be losses that correspond to the detachment levels of the tranches,

$$l_p(t) = L_p. \tag{2.3.53}$$

We also assume that functions $Y_k(t)$ are constant within the intervals between the tranche maturities:

$$Y_p(t) = Y_{i,p}, \quad T_{i-1} \le t < T_i, \tag{2.3.54}$$

where $T_1, T_2, \ldots, T_M$ are the tranche maturity dates and $T_0 = t_0$ is the start date. This form allows setting up a bootstrap calibration to CDO tranches with different maturities: First, one calibrates the model to tranches with maturity T_1, which determines $Y_{1,p}$. Then, the model is calibrated to tranches that mature at T_2, adjusting $Y_{2,p}$, and keeping $Y_{1,p}$ unchanged. The procedure is repeated until the model is fully calibrated.

In fitting after-crisis tranche quotes with acceptable accuracy, the scheme just described often fails. Typically, the model significantly overprices the senior tranche and underprices the supersenior tranche at five-year maturity. The assumption of time independence of the function $Y(N,t)$ within the intervals between adjacent tranche maturities is one of the key reasons for this failure. The values of Y that correspond to more senior tranches effectively have lower leverage for guiding the dynamics. Scaling the loss levels l_k linearly in time for the first interval as

$$l_k(t) = L_k\,\frac{t - t_0}{T_1 - t_0}, \qquad t_0 < t < T_1, \tag{2.3.55}$$

can resolve this problem and dramatically improve the calibration quality in many cases.

Hedge Ratios in the Multiname Dynamic Model

Under the dynamic definition of hedge ratios, when changing the survival probability of a given asset, one immediately perturbs the default intensities of all other assets. Therefore, the procedure described here is referred to as "contagious".

Calculation of the tranche delta with respect to the kth asset can be summarized as follows:

1. Calibrate the dynamic model to CDO tranches and single-name spreads.
2. Find the perturbed values of the survival probability and default intensity λ of the kth asset corresponding to a 1-bp shift of the spread.
3. Recalibrate the model keeping the functions $\{b_j(t)\}$ and $Y(N,t)$ unchanged (as obtained in Step 1) and adjusting the functions $\{a_j(t)\}$ according to

$$a_j(t_i) = \lambda_j(t_i) - \frac{b_j(t_i)}{p_j(t_i)} \sum_{N=0}^{N_0-1} Y(N,t_i)\,P_j(N,t_i), \tag{2.3.56}$$

where $P_j(N,t_i)$ is the probability that both the jth asset has survived up to time t_i and N assets in total have defaulted.

4. Obtain the tranche delta via

$$\delta_k = \frac{\Delta V^{\text{tr}}}{\Delta V_k^{\text{cds}}}, \tag{2.3.57}$$

where ΔV^{tr} is the difference of the tranche values obtained in Steps 3 and 1, and ΔV_k^{cds} is the change in value of the kth CDS.

2.4 Credit Transition Model

The Credit Transition Model is based on the Jarrow-Landro-Turnbull (JLT) credit rating model and its later modifications by Kijma-Komoribayashi (KK) and Halperin. The model is a Markov chain model for pricing risky derivatives that explicitly utilizes the bond's credit rating to indicate the likelihood of the issuer defaulting on its bond.

2.4.1 Definition

Let X_t be the credit rating of a bond at time t, where $t \in \{t_n, n = 0, 1, \dots\}$ is a discrete time variable. Assume that, in the physical measure, X_t is a time-homogeneous Markov chain on the state space $\{1, 2, \dots, K\}$ of credit ratings, where 1 is the highest rating (e.g., AAA in S&P ratings) and $K-1$ is the lowest rating (e.g., CCC in S&P rankings). The state K represents the issuer defaulting on its bond. The transition matrix Q is given by

$$Q = \begin{pmatrix} q_{1,1} & q_{1,2} & \cdots & q_{1,K-1} & q_{1,K} \\ q_{2,1} & q_{2,2} & \cdots & q_{2,K-1} & q_{2,K} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ q_{K-1,1} & q_{K-1,2} & \cdots & q_{K-1,K-1} & q_{K-1,K} \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix},$$

where $q_{i,j} \in [0,1]$ for all i, j and $\sum_{j=1}^{K} q_{i,j} = 1$ for all i. Note that the last row implies that the default state (K) is absorbing. We also assume that X_t is independent of interest rates. Each $q_{i,j}$ represents the real-world probability that the bond moves from rating i to rating j in a single time step.

The risk-neutral probabilities for transitioning from state i to state j between t_n and t_{n+1} are denoted by $\tilde{q}_{i,j}(t_n, t_{n+1})$. In general, these probabilities will depend on time. The different methods compute $\tilde{q}_{i,j}$ in different ways. The KK method, which tends to be more accurate than JLT and less computationally intense than Helperin, is the default method

2.4.2 Implementation

The original formulation of this model is due to Jarrow, Landro and Turnbull (JLT) [37]. This model was improved upon by Kijma and Komoribayashi (KK) [40] and Halperin [33]. The KK method uses different risk premium adjustments, which leads to a more stable expression for the risk premia. The Halperin method uses a minimal entropy technique for the risk premia, and the result is identical to the expression from KK when the model is calibrated to the entire

set of corporate credit spreads. The Halperin method is more general and flexible than the KK scheme.

Jarrow-Landro-Turnbull Method

In the JLT method, the risk premium adjustments $\pi_{i,j}$ are assumed to be deterministic functions of time and independent of j for $j \neq i$, so

$$\pi_{i,j}(t) = \pi_i(t).$$

The risk-neutral probabilities are

$$\tilde{q}_{i,j}(t_n, t_{n+1}) = \begin{cases} \pi_i(t_n) q_{i,j}, & j \neq i, \\ 1 - \pi_i(t_n)(1 - q_{i,i}), & j = i. \end{cases} \tag{2.4.1}$$

Note that $\tilde{q}_{i,i}$ must be positive, so we have

$$0 < \pi_i(t_n) < \frac{1}{1 - q_{i,i}}$$

for all $i = 1, \ldots, K-1$.

The π_i are then found by calibration. The methods for computation are similar to the KK case in Section 2.4.2. See also [37] and [33] for computation details.

We consider two cases: we are given all the default probability curves $p_i(t)$ for all $i < K$, or we are given only a single curve $p_{i_0}(t)$.

Calibration to All Curves Suppose we are given the default probability curves

$$\tilde{q}_{j,K}(t_0, t_n) = p_j(t_n),$$

and let $\hat{Q} = (Q)_{1 \leq i,j \leq K-1}$ be the upper $(K-1) \times (K-1)$ submatrix of $\tilde{Q}$. Then the expression for $\pi_i(t)$ is

$$\pi_i(t_n) = \frac{1}{q_{i,K}} \left[1 - \sum_{j=1}^{K-1} \hat{Q}^{-1}_{i,j}(t_0, t_n) \left(1 - p_j(t_n) \right) \right] \tag{2.4.2}$$

for $1 \leq i \leq K-1$.

Note in particular the coefficient of $\frac{1}{q_{i,K}}$. In practice, the probability in the denominator is often close to zero, indicating that the JLT method can produce large errors. Additionally, the condition $0 < \pi_i(t_n) < \frac{1}{1-q_{i,i}}$ can be violated. The KK method addresses the first issue.

Calibration to a Single Curve Suppose now that we are only given the default probability curve for one state, i_0:

$$\tilde{q}_{i_0,K}(t_0, t_n) = p_{i_0}(t_n).$$

We assume

$$\pi_i(t) = \begin{cases} 1, & i \neq i_0 \\ \pi(t), & i = i_0. \end{cases}$$

Then

$$\pi(t_0) = \frac{p_{i_0}(t_1)}{q_{i_0,K}}$$

and

$$\pi(t_n) = \frac{p_{i_0}(t_{n+1}) - p_{i_0}(t_n) - \sum_{j=1, j\neq i_0}^{K-1} q_{j,K}}{\tilde{q}_{i_0,i_0}(t_0, t_n)},$$

for $n > 0$. Note that this expression is inductive and $\tilde{q}_{i_0,i_0}(t_0, t_n)$ can be found from the previous steps.

Kijima-Komoribayashi Method

Following [40], we use

$$\tilde{q}_{i,j}(t_n, t_{n+1}) = \begin{cases} \pi_i(t_n) q_{i,j}, & j \neq K, \\ 1 - \pi_i(t_n)(1 - q_{i,K}), & j = K, \end{cases} \tag{2.4.3}$$

for the risk-neutral transition probabilities, where π_i are deterministic functions on the time grid. As in the JLT method, we have constraint

$$0 < \pi_i(t_n) \leq \frac{1}{1 - q_{i,K}} \tag{2.4.4}$$

for all $i = 1, \ldots, K-1$.

The π_i are then found by calibration. We consider two cases: we are given all the default probability curves $p_i(t)$ for all $i < K$, or we are given only a single curve $p_{i_0}(t)$.

Calibration to All Curves Assume that we are given the physical transition probability matrix $Q = (q_{i,j})$ and the risk-neutral default probability curves $p_i(t)$ for all states $i < K$.

The idea is to use an inductive argument to find $\pi_i(t_n)$ for an arbitrary n. We first find $\pi_i(t_0)$ in terms of the inputs, and then we show how to compute $\pi_i(t_n)$ having computed π_i at each previous time. The input data gives us $\tilde{q}_{i,K}(t_0, t_n) = p_i(t_n)$ for all n, which represents the risk-neutral probability that the isuer defaults on its bond by time t_n given that it started in state i at time t_0.

For t_0, we have $\tilde{q}_{i,K}(t_0, t_1) = p_i(t_1)$. By Eq. (2.4.3), we have $\tilde{q}_{i,K}(t_0, t_1) = 1 - \pi_i(t_0)(1 - q_{i,K})$. Equating these then gives

$$\pi_i(t_0) = \frac{1 - p_i(t_1)}{1 - q_{i,K}}.$$

Now, suppose that we have $\pi_i(t_\ell)$ (and therefore $\tilde{q}_{i,j}(t_\ell, t_{\ell+1})$ for all $\ell < n$). To compute $\pi_i(t_n)$, consider

$$\tilde{q}_{i,K}(t_0, t_{n+1}) - \tilde{q}_{i,K}(t_0, t_n) = p_i(t_{n+1}) - p_i(t_n),$$

which is the (risk-neutral) probability of defaulting between times t_n and t_{n+1}, without knowledge of the bond's state at t_n, given that the bond started in state i at t_0. We may also compute this

probability by summing over all states j where we consider the bond to transition from state i to state j between t_0 and t_n and then default from state j between t_n and t_{n+1}. This gives

$$p_i(t_{n+1}) - p_i(t_n) = \sum_{j=1}^{K-1} \tilde{q}_{i,j}(t_0, t_n)\tilde{q}_{j,K}(t_n, t_{n+1}). \tag{2.4.5}$$

Let

$$\hat{Q} = \begin{pmatrix} \tilde{q}_{1,1} & \tilde{q}_{1,2} & \cdots & \tilde{q}_{1,K-1} \\ \tilde{q}_{2,1} & \tilde{q}_{2,2} & \cdots & \tilde{q}_{2,K-1} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{q}_{K-1,1} & \tilde{q}_{K-1,2} & \cdots & \tilde{q}_{K-1,K-1} \end{pmatrix} \tag{2.4.6}$$

be the upper-left $(K-1) \times (K-1)$ submatrix of $\tilde{Q}$. Then the right-hand side of Eq. (2.4.5) is the ith component of $\hat{Q}$ times the column vector whose jth component is $\tilde{q}_{j,K}(t_n, t_{n+1})$. Then inverting Eq. (2.4.5) gives

$$\sum_{j=1}^{K-1} \hat{Q}^{-1}_{i,j} \left(p_j(t_{n+1}) - p_j(t_n)\right) = \tilde{q}_{i,K}(t_n, t_{n+1}),$$

where $\hat{Q}^{-1}_{i,j}$ denotes the i,j component of the matrix $\hat{Q}^{-1}$.

Now, by definition we have $\hat{q}_{i,K}(t_n, t_{n+1}) = 1 - \pi_i(t_n)(1 - q_{i,K})$, so we find

$$\pi_i(t_n) = \frac{1}{1 - q_{i,K}} \left(1 - \sum_{j=1}^{K-1} Q^{-1}_{i,j}(t_0, t_n) \left(p_j(t_{n+1}) - p_j(t_n)\right)\right). \tag{2.4.7}$$

We may write $p_j(t_n) = \tilde{q}_{j,K}(t_0, t_n) = 1 - \sum_{\ell=1}^{K-1} \tilde{q}_{j,\ell}(t_0, t_n)$ and compute

$$\begin{aligned} \sum_{j=1}^{K-1} \hat{Q}^{-1}_{i,j}(t_0, t_n)\tilde{q}_{j,K}(t_0, t_n) &= \sum_{j=1}^{K-1} \hat{Q}^{-1}_{i,j}(t_0, t_n) - \sum_{\ell=1}^{K-1}\sum_{j=1}^{K-1} \hat{Q}^{-1}_{i,j}(t_0, t_n)\hat{q}_{j,\ell}(t_0, t_n) \\ &= \sum_{j=1}^{K-1} \hat{Q}^{-1}_{i,j}(t_0, t_n) - \sum_{\ell=1}^{K-1} \delta_{i,\ell} \\ &= \sum_{j=1}^{K-1} \hat{Q}^{-1}_{i,j}(t_0, t_n) - 1 \end{aligned}$$

to give

$$\pi_i(t_n) = \frac{1}{1 - q_{i,K}} \sum_{j=1}^{K-1} \hat{Q}^{-1}_{i,j}(t_0, t_n) \left(1 - p_j(t_{n+1})\right). \tag{2.4.8}$$

Note that this has the advantage over the JLT method of having $1 - q_{i,K} \approx 1$ in the denominator instead of $q_{i,K} \approx 0$. We therefore expect the KK method to be more accurate and stable.

Calibration to a Single Curve Suppose now that we are given the physical transitional probability matrix $Q = (q_{i,j})$ and the risk-neutral default probability curve $p_{i_0}(t)$ for only the current

state i_0. We assume

$$\pi_i(t) = \begin{cases} 1, & i \neq i_0 \\ \pi(t), & i = i_0. \end{cases}$$

We use the same induction idea as in the case when we have all the default probability curves, but in this case we only need to compute a single curve π instead of all π_i.

For t_0, we have $p_{i_0}(t_1) = \tilde{q}_{i_0,K}(t_0,t_1) = 1 - \pi_{i_0}(t_0)(1-q_{i_0,K}) = 1 - \pi(t_0)(1-q_{i_0,K})$. Thus,

$$\pi(t_0) = \frac{1 - p_{i_0}(t_1)}{1 - q_{i_0,K}}.$$

Now assume we have computed $\pi(t_\ell)$ for all $\ell < n$. Then as before,

$$\begin{aligned} p_{i_0}(t_{n+1}) - p_{i_0}(t_n) &= \tilde{q}_{i_0,K}(t_0,t_{n+1}) - \tilde{q}_{i_0,K}(t_0,t_n) \\ &= \sum_{j=1}^{K-1} \tilde{q}_{i_0,j}(t_0,t_n)\tilde{q}_{j,K}(t_n,t_{n+1}) \\ &= \sum_{j=1}^{K-1} \tilde{q}_{i_0,j}(t_0,t_n)(1 - \pi(t_n)(1-q_{j,K})). \end{aligned}$$

Solving for $\pi(t_n)$ gives

$$\pi(t_n) = \frac{\sum_{j=1}^{K-1} \tilde{q}_{i_0,j}(t_0,t_n) - p_{i_0}(t_{n+1}) + p_{i_0}(t_n)}{\sum_{j=1}^{K-1} \tilde{q}_{i_0,j}(t_0,t_n)(1-q_{j,K})}. \tag{2.4.9}$$

Halperin Method

Halperin [33] uses an entropy method to determine $\tilde{q}_{i,j}(t_n,t_{n+1})$. We minimize an entropy functional that measures the distance between the given probabilities $q_{i,j}$ and the risk-neutral probabilities $\tilde{q}_{i,j}$.

Again we consider calibrating to all curves and calibrating to only a single curve. In the case of calibration to all curves, the Halperin and KK methods agree. When calibrating to only a single curve, the expressions for the risk-neutral probabilities are different. For most cases of practical interest, the Halperin method provides little advantage over the KK method while requiring more computation, so the KK method is preferred.

Calibration to All Curves Assume that we are given the physical transition probability matrix $Q = (q_{i,j})$ and the risk-neutral default probability curves $p_i(t)$ for all states $i < K$.

Define

$$H = \int \tilde{q}(x) \log\left(\frac{\tilde{q}(x)}{q(x)}\right) dx.$$

The functional H is the Kullback-Leibler relative entropy and serves as a measure of the distance between q and $\tilde{q}$. Note that when $q = \tilde{q}$ a.e., H is zero. We use a Lagrange multiplier method to find the $\tilde{q}_{i,j}$ that minimize H at each point in time grid $\{t_n\}$.

Let $p_i(t)$ denote the input default probabilities. At t_0, the constraints are

$$\sum_j \tilde{q}_{i,j}(t_0,t_1) = 1, \tag{2.4.10}$$

$$\tilde{q}_{i,K}(t_0,t_1) = p_i(t_0), \tag{2.4.11}$$

for all i. This gives the Lagrangian

$$L = -\sum_j \tilde{q}_{i,j}(t_0,t_1)\log\left(\frac{\tilde{q}_{i,j}(t_0,t_1)}{q_{i,j}}\right) + \lambda_i\left(\sum_j \tilde{q}_{i,j}(t_0,t_1) - 1\right) + \mu_i\left(\tilde{q}_{i,K}(t_0,t_1) - p_1(t_0)\right), \tag{2.4.12}$$

where $\{\lambda_i\}$ and $\{\mu_i\}$ are Lagrange multipliers. Setting $\nabla L = 0$ (where the variables are $\{\tilde{q}_{i,j}(t_0,t_1)\}$, $\{\lambda_i\}$ and $\{\mu_i\}$) gives the system

$$0 = \frac{\partial L}{\partial \lambda_i} = \sum_j \tilde{q}_{i,j}(t_0,t_1) - 1, \tag{2.4.13}$$

$$0 = \frac{\partial L}{\partial \mu_i} = \tilde{q}_{i,K}(t_0,t_1) - p_i(t_0), \tag{2.4.14}$$

$$0 = \frac{\partial L}{\partial \tilde{q}_{i,j}(t_0,t_1)} = -\log\left(\frac{\tilde{q}_{i,j}(t_0,t_1)}{q_{i,j}}\right) - 1 + \lambda_i + \mu_i \mathbb{I}_{j=k}. \tag{2.4.15}$$

Rewrite this system as

$$\tilde{q}_{i,K}(t_0,t_1) = p_i(t_0), \tag{2.4.16}$$

$$\sum_{j\neq K} \tilde{q}_{i,j}(t_0,t_1) = 1 - p_i(t_0), \tag{2.4.17}$$

$$\tilde{q}_{i,j}(t_0,t_1) = q_{i,j}e^{\lambda_1 + \mu_1 \mathbb{I}_{j=K} - 1}. \tag{2.4.18}$$

Combining Equations (2.4.17) and (2.4.18) and using $\sum_j q_{i,j} = 1$ gives

$$1 - p_i(t_0) = e^{\lambda_i} - 1 \sum_{j\neq K} q_{i,j} = e^{\lambda_i - 1}\left(1 - q_{i,K}\right),$$

which gives

$$e^{\lambda_i - 1} = \frac{1 - p_i(t_0)}{1 - q_{i,K}},$$

and therefore

$$\tilde{q}_{i,j}(t_0,t_1) = \begin{cases} \frac{q_{i,j}(1-p_i(t_0))}{1-q_{i,K}} & j \neq K, \\ p_i(t_0), & j = K. \end{cases}$$

Note that this is the same as the expression for the KK probabilities in Eq. (2.4.3), with

$$\pi_i(t_0) = \frac{1 - p_i(t_0)}{1 - q_{i,K}}.$$

Now, assume we know $\tilde{q}_{i,j}$ at all points up to time t_n. We will find $\tilde{q}_{i,j}(t_n,t_{n+1})$ using the same method as above. Since we know each of the one-step transition probabilities up to time t_n, we know the entire cumulative transition matrix $\tilde{Q}$ at t_n. We then seek to update this matrix

by computing the one-step transition probabilities $\tilde{q}_{j,\ell}(t_n, t_{n+1})$. (We have the probabilities that they move from i to j, so we need to compute the probabilities that they then move from j to some other state ℓ.) Given that the bond was in the state i at time t_0, the one-step transition probabilities $\tilde{q}_{j,\ell}(t_n, t_{n+1})$ should minimize the conditional Kullback-Leibler entropy

$$H = \sum_{j \neq k} \tilde{q}_{i,j}(t_0, t_n) \sum_{\ell} \tilde{q}_{j,\ell}(t_n, t_{n+1}) \log\left(\frac{\tilde{q}_{j,\ell}(t_n, t_{n+1})}{q_{j,\ell}}\right).$$

The constraints are

$$\sum_{j} \tilde{q}_{i,j}(t_n, t_{n+1}) = 1, \tag{2.4.19}$$

$$\sum_{j=1}^{K-1} \tilde{q}_{i,j}(t_0, t_n)\tilde{q}_{j,K}(t_n, t_{n+1}) = p_i(t_{n+1}) - p_i(t_0), \tag{2.4.20}$$

and the Lagrangian is

$$\begin{aligned} L = &- \sum_{i,j \neq K} \tilde{q}_{i,j}(t_0, t_n) \sum_{\ell} \tilde{q}_{j,\ell}(t_n, t_{n+1}) \log\left(\frac{\tilde{q}_{j,\ell}(t_n, t_{n+1})}{q_{j,\ell}}\right) \\ &+ \sum_{i,j \neq K} \lambda_j \tilde{q}_{i,j}(t_0, t_n) \left(\sum_{\ell} \tilde{q}_{j,\ell}(t_n, t_{n+1}) - 1\right) \\ &+ \sum_{i \neq K} \mu_i \left(\sum_{j \neq K} \tilde{q}_{i,j}(t_0, t_n)\tilde{q}_{j,K}(t_n, t_{n+1}) - p_i(t_{n+1}) + p_i(t_n)\right). \end{aligned} \tag{2.4.21}$$

Note that we sum over all states $i \neq K$.

Setting $\nabla L = 0$ gives the system

$$0 = \frac{\partial L}{\partial \lambda_j} = \sum_{i \neq K} \tilde{q}_{i,j}(t_0, t_n) \left(\sum_{\ell} \tilde{q}_{j,\ell}(t_n, t_{n+1}) - 1\right), \tag{2.4.22}$$

$$0 = \frac{\partial L}{\partial \mu_i} = \sum_{j \neq K} \tilde{q}_{i,j}(t_0, t_n)\tilde{q}_{j,K}(t_n, t_{n+1}) - p_i(t_{n+1}) + p_i(t_n), \tag{2.4.23}$$

$$0 = \frac{\partial L}{\partial \tilde{q}_{j,\ell}(t_n, t_{n+1})} = -\sum_{i \neq K} \tilde{q}_{i,j}(t_0, t_n) \left(\log\left(\frac{\tilde{q}_{j,\ell}(t_n, t_{n+1})}{q_{j,\ell}}\right) + 1\right) \tag{2.4.24}$$

$$+ \lambda_j \sum_{i \neq K} \tilde{q}_{i,j}(t_0, t_n) + \mathbb{I}_{j=k} \sum_{i \neq K} \mu_i \tilde{q}_{i,j}(t_0, t_n). \tag{2.4.25}$$

We may rewrite this as

$$\sum_{\ell} \tilde{q}_{j,\ell}(t_n, t_{n+1}) = 1, \tag{2.4.26}$$

$$\sum_{j=1}^{K-1} \tilde{q}_{i,j}(t_0, t_n)\tilde{q}_{j,K}(t_n, t_{n+1}) = p_i(t_{n+1}) - p_i(t_n), \tag{2.4.27}$$

$$\tilde{q}_{j,\ell}(t_n, t_{n+1}) = q_{j,\ell} e^{\lambda_j - 1 + \mathbb{I}_{\ell=K} \frac{\sum_{m \neq K} \mu_m \tilde{q}_{m,j}(t_0, t_n)}{\sum_{m \neq K} \tilde{q}_{m,j}(t_0, t_n)}}. \tag{2.4.28}$$

Note from the last equation that we are guaranteed that $\tilde{q}_{j,\ell} \geq 0$ for all j, ℓ and at all points in the time grid. Summing the third equation over ℓ and using the first equation with $\sum_\ell q_{j,\ell} = 1$ gives

$$e^{\lambda_j - 1} = \frac{1}{1 - q_{j,K} + q_{j,K} e^{\frac{\sum_{m \neq K} \mu_m \tilde{q}_{m,j}(t_0,t_n)}{\sum_{m \neq K} \tilde{q}_{m,j}(t_0,t_n)}}}.$$

The solution of the system is then

$$\tilde{q}_{j,\ell}(t_n, t_{n+1}) = \frac{q_{j,\ell} e^{\mathbb{I}_{\ell=K} \frac{\sum_{m \neq K} \mu_m \tilde{q}_{m,j}(t_0,t_n)}{\sum_{m \neq K} \tilde{q}_{m,j}(t_0,t_n)}}}{1 - q_{j,K} + q_{j,K} e^{\frac{\sum_{m \neq K} \mu_m \tilde{q}_{m,j}(t_0,t_n)}{\sum_{m \neq K} \tilde{q}_{m,j}(t_0,t_n)}}}.$$

Reindexing and writing in terms of the premia gives

$$\tilde{q}_{i,j}(t_n, t_{n+1}) = \begin{cases} \pi_i(t_n) q_{i,j}, & j \neq K, \\ 1 - \pi_i(t_n)(1 - q_{i,K}), & j = K, \end{cases} \tag{2.4.29}$$

where

$$\pi_i(t_n) = \left[1 - q_{i,K} + q_{i,K} e^{\frac{\sum_{m \neq K} \mu_m \tilde{q}_{m,i}(t_0,t_n)}{\sum_{m \neq K} \tilde{q}_{m,i}(t_0,t_n)}} \right]^{-1}. \tag{2.4.30}$$

Note that the Lagrange multipliers $\{\mu_j\}$ appear in the expression for $\pi_i(t_n)$ in Eq. (2.4.30), and we do not have an explicit formula for them. However, we can put Eq. (2.4.29) into Eq. (2.4.20) and solve for $\pi_i(t_n)$ to obtain

$$\pi_i(t_n) = \frac{1}{1 - q_{i,K}} \sum_{j=1}^{K-1} \hat{Q}^{-1}_{i,j}(t_0, t_n) \left(1 - p_j(t_{n+1})\right). \tag{2.4.31}$$

Calibration to A Single Curve Suppose now that we are given the physical transitional probability matrix $Q = (q_{i,j})$ and the risk-neutral default probability curve $p_{i_0}(t)$ for only the current state i_0.

We again construct an entropy functional and use Lagrange multipliers to find its extrema. Again the entropy functional is

$$H = \int \tilde{q}(x) \log\left(\frac{\tilde{q}(x)}{q(x)}\right) dx,$$

and the initial constraints are

$$\sum_j \tilde{q}_{i,j}(t_0, t_1) = 1 \tag{2.4.32}$$

$$\tilde{q}_{i_0,K}(t_0, t_1) = p_{i_0}(t_0). \tag{2.4.33}$$

When $i = i_0$, the Lagrangian is given by

$$L = -\sum_j \tilde{q}_{i_0,j}(t_0, t_1) \log\left(\frac{\tilde{q}_{i_0,j}(t_0, t_1)}{q_{i_0,j}}\right) + \lambda_{i_0} \left(\sum_j \tilde{q}_{i_0,j}(t_0, t_1) - 1\right) + \mu \left(\tilde{q}_{i_0,K}(t_0, t_1) - p_{i_0}(t_0)\right).$$

Setting $\nabla L = 0$ and solving in the same manner as the multiple curve calibration case gives

$$\tilde{q}_{i_0,j}(t_0,t_1) = \begin{cases} \frac{q_{i_0,j}(1-p_{i_0}(t_0))}{1-q_{i_0,K}}, & j \neq K, \\ p_{i_0}(t_0), & j = K. \end{cases}$$

This agrees with the KK result with $\pi(t_0) = \frac{1-p_{i_0}(t_0)}{1-q_{i_0,K}}$.

For $i \neq i_0$, we only have the first constraint and the Lagrangian is

$$L = -\sum_j \tilde{q}_{i,j}(t_0,t_1) \log\left(\frac{\tilde{q}_{i,j}(t_0,t_1)}{q_{i,j}}\right) + \lambda_i \left(\sum_j \tilde{q}_{i,j}(t_0,t_1) - 1\right).$$

Setting $\nabla L = 0$ gives

$$\sum_j \tilde{q}_{i,j} = 1, \tag{2.4.34}$$

$$\tilde{q}_{i,j}(t_0,t_1) = q_{i,j} e^{\lambda_i - 1}, \tag{2.4.35}$$

which leads to $\sum_j q_{i,j} = e^{1-\lambda_i}$. But $\sum_j q_{i,j} = 1$, so $\lambda_i = 1$ for all $i \neq i_0$ and

$$\tilde{q}_{i,j}(t_0,t_1) = q_{i,j},$$

which is consistent with the KK method where $\pi_i(t_0) = 1$ for $i \neq i_0$.

Now, assume we know $\tilde{q}(t_\ell, t_{\ell+1})$ for all $\ell < n$. At t_{n+1} we have the constraints

$$\sum_\ell \tilde{q}_{j,\ell}(t_n,t_{n+1}) = 1, \tag{2.4.36}$$

$$\sum_{j=1}^{K-1} \tilde{q}_{i_0,j}(t_0,t_1)\tilde{q}_{j,K}(t_n,t_{n+1}) = p_{i_0}(t_{n+1}) - p_{i_0}(t_n). \tag{2.4.37}$$

The Lagrangian is

$$\begin{aligned} L = &-\sum_{i,j\neq K} \tilde{q}_{i,j}(t_0,t_n) \sum_\ell \tilde{q}_{j,\ell}(t_n,t_{n+1}) \log\left(\frac{\tilde{q}_{j,\ell}(t_n,t_{n+1})}{q_{j,\ell}}\right) \\ &+ \sum_{i,j\neq K} \lambda_j \tilde{q}_{i,j}(t_0,t_n) \left(\sum_\ell \tilde{q}_{j,\ell}(t_n,t_{n+1}) - 1\right) \\ &+ \mu \left(\sum_{j\neq K} \tilde{q}_{i_0,j}(t_0,t_n)\tilde{q}_{j,K}(t_n,t_{n+1}) - p_{i_0}(t_{n+1}) + p_{i_0}(t_n)\right). \end{aligned} \tag{2.4.38}$$

Then setting $\nabla L = 0$ leads to

$$\sum_{\ell} \tilde{q}_{j,\ell}(t_n, t_{n+1}) = 1,$$

$$\sum_{j=1}^{K-1} \tilde{q}_{i_0,j}(t_0, t_n)\tilde{q}_{j,K}(t_n, t_{n+1}) = p_{i_0}(t_{n+1}) - p_{i_0}(t_n),$$

$$\tilde{q}_{j,\ell}(t_n, t_{n+1}) = q_{j,\ell} e^{\lambda_j - 1 + \mu \frac{\tilde{q}_{i_0,j}(t_0,t_n)}{\sum_{i\neq K} \tilde{q}_{i,j}(t_0,t_n)} \mathbb{I}_{\ell=K}}.$$

Applying the first constraint to the third equation, together with $\sum_\ell q_{j,\ell} = 1$ gives

$$e^{\lambda_j - 1} = \left(1 - q_{j,K} + q_{j,K} e^{\mu \frac{\tilde{q}_{i_0,j}(t_0,t_n)}{\sum_{i\neq K} \tilde{q}_{i,j}(t_0,t_n)}}\right)^{-1}.$$

Then we have

$$\tilde{q}_{j,\ell}(t_n, t_{n+1}) = \frac{q_{j,\ell} e^{\mu \frac{\tilde{q}_{i_0,j}(t_0,t_n)}{\sum_{i\neq K} \tilde{q}_{i,j}(t_0,t_n)} \mathbb{I}_{\ell=K}}}{1 - q_{j,K} + q_{j,K} e^{\mu \frac{\tilde{q}_{i_0,j}(t_0,t_n)}{\sum_{i\neq K} \tilde{q}_{i,j}(t_0,t_n)}}}. \tag{2.4.39}$$

which we may write as

$$\tilde{q}_{j,\ell}(t_n, t_{n+1}) = \begin{cases} \pi_j(t_n) q_{j,\ell}, & \ell \neq K, \\ 1 - \pi_j(t_n)(1 - q_{j,K}), & \ell = K, \end{cases} \tag{2.4.40}$$

where

$$\pi_j(t_n) = \frac{1}{1 - q_{j,K} + q_{j,K} e^{\mu \frac{\tilde{q}_{i_0,j}(t_0,t_n)}{\sum_{i\neq K} \tilde{q}_{i,j}(t_0,t_n)}}}. \tag{2.4.41}$$

As before, instead of solving for the Lagrange multiplier μ, we put Eq. (2.4.40) into the constraint

$$\sum_{j=1}^{K-1} \tilde{q}_{i_0,j}(t_0, t_n)\tilde{q}_{j,K}(t_n, t_{n+1}) = p_{i_0}(t_{n+1}) - p_{i_0}(t_n)$$

to give

$$p_{i_0}(t_{n+1}) - p_{i_0}(t_n) = \sum_{j=1}^{K-1} \tilde{q}_{i_0,j}(t_0, t_n) \left(1 - \frac{1 - q_{j,K}}{1 - q_{j,K} + q_{j,K} e^{\mu \frac{\tilde{q}_{i_0,j}(t_0,t_n)}{\sum_{i\neq K} \tilde{q}_{i,j}(t_0,t_n)}}}\right). \tag{2.4.42}$$

This is a continuous nonlinear equation in μ, and one can see that as $\mu \to \infty$, the right-hand side of Eq. (2.4.42) approaches

$$\sum_{j=1}^{K-1} \tilde{q}_{i_0,j}(t_0, t_n) = 1 - p_{i_0}(t_n) \geq p_{i_0}(t_{n+1}) - p_{i_0}(t_n).$$

As $\mu \to -\infty$, the right-hand-side approaches zero, and by the assumption that the $\tilde{q}_{i,j}$ are positive, we have $p_{i_0}(t_{n+1}) - p_{i_0}(t_n) > 0$. Thus, the intermediate value theorem guarantees a solution μ^* to Eq. (2.4.42). Putting μ^* into Eq. (2.4.41) gives $\tilde{q}_{j,\ell}(t_n, t_{n+1})$ for all j and ℓ by Eq. (2.4.40).

Note that this is different from the KK probabilities, where $\pi_i(t_n) = 1$ for all $i \neq i_0$. Therefore, the Halperin method gives a different result in the case where we only have one input curve.

2.4.3 Pricing

There are two methods to price credit-sensitive securities using the credit transition model: on a lattice or with Monte Carlo simulation. For vanilla instruments or instruments with early exercise conditions, the lattice method is preferred, whereas for strongly path-dependent securities, Monte Carlo simulation must be used.

Lattice Method

In a lattice method, we need the conditional expectations

$$V_t = \mathbb{E}\left[V_T \,|\, \mathcal{F}_t\right].$$

To compute V_t, we first compute $\tilde{Q}(t,T)$ by

$$\tilde{Q}(t,T) = \tilde{Q}(t_{\ell-1}, t_\ell)^{\frac{t_\ell - t}{t_\ell - t_{\ell-1}}} \left(\prod_{j=\ell}^{r-1} \tilde{Q}(t_j, t_{j+1})\right) \tilde{Q}(t_r, t_{r+1})^{\frac{T - t_r}{t_{r+1} - t_r}}, \tag{2.4.43}$$

where

$$\ell = \min\{j : t_j \geq t\},$$
$$r = \max\{j : t_j < T\}.$$

The matrix powers in (2.4.43) can be computed using a Schur-Parlett algorithm (G. Golub (1996)) for computing functions of matrices. Given $\tilde{Q}(t,T)$, we can find V_t by

$$V_t = \tilde{Q}(t,T)' V_T,$$

where $'$ denotes a transpose.

Simulation Method

There are two options to simulate the paths of X_t: step-by-step simulation and first-even simulation. In the step-by-step method, we simulate X_t from one time to the next by sampling directly from the discrete conditional distribution, which is a row-transitional matrix and is computed in the same way as in the lattice method. In the first-event method, we simulate the time-to-state-change exponential random variables and take the shortest one to determine the next time and state.

2.5 Black-Karasinski Credit Model

The Black-Karasinski (BK) credit model is a one-factor model for a hazard (intensity) rate. Let $\lambda \geq 0$ be the hazard rate, and set

$$x(t) = \log \lambda(t).$$

The BK model gives the dynamics of x as

$$dx_t = \left(\theta(t) - \kappa x_t\right) dt + \sigma(t) dW_t$$

for some parameters θ, κ, and σ. Here κ represents the rate of mean reversion, σ represents the volatility, and θ/κ is the long-term mean. The parameter θ is used to fit the current yield curve.

The BK model can be calibrated to CDS options.

2.6 Cox-Ingersoll-Ross Credit Model

The Cox-Ingersoll-Ross (CIR) credit model is a one-factor model for a hazard (intensity) rate. It is analogous to the scaled CIR model for interest rates. The CIR credit model gives the hazard rate λ by

$$\lambda_t = \beta(t) x_t,$$

where $\beta(t) \geq 0$ is deterministic and x_t satisfies

$$dx_t = a(t)\left(1 - x_t\right) dt + \sigma(t)\sqrt{x_t} dW_t,$$

for some volatility σ and mean reversion a, both of which are generally time dependent.

Chapter 3

Foreign Exchange and Equity Models

In this chapter, we describe the Numerix single-foreign-exchange and equity models. Below, we suppose that the interest rates are deterministic; we will address the hybrid models with stochastic interest rates in another chapter.

The interface for the standard equity models and FX models is identical, except for the fact that the FX models take an object representing the foreign interest rate, whereas equity models take a continuous dividend curve object. We use $u(t)$ for both the foreign interest rate in FX models and the dividend curve in EQ models.

The user has a choice in discrete dividend treatments between the two most popular ways described in Bos, Gairat and Shepeleva [57]. Namely, their Model 1 and Model 2 preserve analytical solvability, which means that if a European option with zero dividends is analytically solvable, then a discrete dividend modification remains solvable.

Below, we list the equity (FX) models that are available in Numerix products in terms of the SDEs on S_t that contain *continuous* dividend (foreign) yield $u(t)$. For example, the SDE for the Black-Scholes model is

$$dS_t = (r(t) - u(t))S_t dt + \sigma(t)S_t dW_t.$$

Let d_i denote a discrete equity dividend paid at t_i. In order to adapt this model to *discrete* dividends, one should set $u(t) = 0$ and make some transformations to obtain equity price e_t in the presence of discrete dividends.

- **Model 1**
  First, given a model time horizon T, one reduces the initial value of the driving process S_0 by $D_0(T)$ (so $S_0 \to S_0 - D_0(T)$), where

  $$D_t(T) = \sum_i \mathbb{I}_{t \le t_i < T}\ d_i\, P(t, t_i)$$

  is the dividend payment stream starting from time t and $P(t, t')$ is the price of a zero coupon bond maturing at t'. Then, the equity in the presence of discrete dividends is

  $$e_t = S_t + D_t(T).$$

This model keeps positive equity prices and holds valid for stochastic interest rates.[1] The main drawback of Model 1 is its dependence of the model time-horizon T.

- **Model 2**
  Here, the zero-dividend process S_t is considered as a sum of the equity price e_t and the rolled-forward PV of already-paid dividends,

$$D(t) = \sum_i \mathbb{I}_{t_i<t}\ d_i e^{\int_{t_i}^{t} r(s)ds}.$$

  In other words, the equity price at time t is

$$e_t = S_t - D(t).$$

  Here, usage of stochastic interest rates is problematic and a positivity of the equity process is not guaranteed.

Let us stress again that if the zero-dividend model S_t is solved, then one can also solve the discrete dividends model e_t in the same manner.

3.1 Black-Scholes Model

This is the standard Black-Scholes Model

$$dS_t = (r(t) - u(t))S_t dt + \sigma(t)S_t dW_t, \tag{3.1.1}$$

where r is the interest rate, u is the continuous dividend yield (or foreign interest rate for FX models) and σ is the volatility.

3.2 Dupire's Local Volatility Model

This standard local volatility model is given by

$$dS_t = (r(t) - u(t))S_t dt + \sigma_{loc}(S_t, t)S_t dW_t, \tag{3.2.1}$$

where r is the interest rate, u is the continuous dividend yield (or foreign interest rate for FX models) and σ_{loc} is the local volatility.

Dupire's formula is the standard method used to extract the local volatility $\sigma_{loc}(S_t, t)$ from the implied volatilities $\sigma_{imp}(K, T)$ of European options with different strikes and maturities. As described by Gatheral [30], we store the implied volatility surface in the form $w(x, T)$, where

$$w = \sigma_{imp}^2 T, \tag{3.2.2}$$

$$x = \ln\left(\frac{K}{F_T}\right). \tag{3.2.3}$$

[1] For a deterministic yield curve, the bond prices are $P(t, t_i) = e^{-\int_t^{t_i} r(s)ds}$.

In this parametrization, Dupire's formula is

$$\sigma_{loc}^2(x,t) = \frac{\partial w/\partial T}{1 - \frac{x}{w}\frac{\partial w}{\partial x} + \frac{1}{4}(-\frac{1}{4} - \frac{1}{w} + \frac{x^2}{w^2})(\frac{\partial w}{\partial x})^2 + \frac{1}{2}\frac{\partial^2 w}{\partial x^2}}. \quad (3.2.4)$$

The condition that the implied volatility surface must be arbitrage free corresponds to both the denominator and numerator of 3.2.4 being positive. The volatility surface data is stored as an array of splines (one per maturity of calibrating options) which perform the interpolation in time/strike and the calculation of the derivatives $\partial w/\partial x$ and $\partial^2 w/\partial x^2$.

In the case where only a single calibrating option volatility is given, the local volatility function is constant and the Dupire model tends to the limit of the Black-Scholes model.

3.3 Advanced Fit Local Volatility Model

The general one-factor dynamics of the stock is described by the SDE

$$S^{-1}dS = (r(t) - q(t))dt + \sigma(t,S)dW,$$

where r is the discounting rate, u is the continuous relative dividend yield (or foreign interest rate for FX models) and σ is the volatility.

The usual approach (after Dupire et al.) is the semi-analytical setup of the local volatility surface from the derivatives of the implied volatility surface. This approach essentially assumes perfect calibration data given an arbitrary strike and date. It is formally not connected to any particular numerical method of pricing. In practice, even after all filtering and smoothing, it often leads to numerical instabilities and rarely allows reproducing the market with a good accuracy.

We adopt another way, based on the direct modeling within a specific parameterization of the local volatility surface. We apply the standard calibration procedure, adjusting the model to the best fit of the calibration options. This can be done globally, by fitting all the options at once, or by horizons, every time fitting the options with same exercise date. The latter procedure is faster, although may sometimes lead to a slightly worse fit at the late horizons. The idea is not new: a similar approach was suggested by Guillaume Blacher of Reech Capital PLC (see http://www.scribd.com/doc/19606988/Reech-Capital-PLC-Blacher-Pricing-With-a-Volatility-Smile).

In our parameterization, the local volatility surface $\sigma(t,S)$ is piecewise constant in time. In the beginning of every horizon t_i, the volatility surface is a linear spline $\sigma_i(S)$ as function of $\log(S/S_0)$, with flat extrapolation. The "space" nodes S_{ik} are taken from the strikes of calibration options (all or a part, as specified). Using few strikes (three or five) effectively provides a surface smoothing, thus leading to a more stable calibration. The best result is obtained if all the strikes are used for nodes and the explicit smoothing is applied by specifying the "tension" parameter λ. The calibrator then adds the penalty

$$\lambda^2 \sum (\sigma_{i+1} - \sigma_i)^2/(x_{i+1} - x_i)^2$$

to the overall calibration defect to be minimized.

The time discretization is important for the quality of the LV pricing. The dates nodes are automatically placed in the beginning of every horizon, but this not enough and a high "timesteps"

quality parameter is usually needed. In addition, as the starting horizons are shorter, we recommend to specify "dates discretization" as "log-uniform" in the method table.

The calibrator minimizes the overall χ^2 defect which, per the user's choice, can be derived from prices, normal volatilities, implied BS volatilities or their relative values. As usual in calibration, the irrelevant (insensitive) instruments should not be used. Specifically to the LV calibration, it is highly recommended to exclude the options of low exercise probability. A ready to use filter for that is available in our tools.

3.4 Hagan's SABR Model

Hagan's SABR model [32] is given by

$$dS_t = (r_t - u_t)S_t dt + \hat{\alpha}_t S_t^{\beta} dW, \tag{3.4.1}$$

$$d\hat{\alpha}_t = \xi \hat{\alpha}_t dZ, \tag{3.4.2}$$

$$\langle dW dZ \rangle = \rho dt, \tag{3.4.3}$$

where r_t is the domestic short rate and u_t is the foreign short rate. Note that the SABR model reduces to the Black-Scholes model in the limit as $\beta \to 1$, $\xi \to 0$.

This model is very attractive because there exists an analytical approximation for the implied volatility of European options that is both very accurate and extremely fast to evaluate. These can therefore be used for fast calibration of the model. The Numerix implementation uses a version of Hagan's SABR model that has mean reversion added into the volatility process. In order for the second process to have the same dimensions as the Black-Scholes volatility, we have chosen instead to simulate the process $v_t = \hat{\alpha}_t F_0^{\beta-1}$. The model is then given by

$$dS_t = (r_t - u_t)S_t dt + v_t F_0^{1-\beta} S_t^{\beta} dW, \tag{3.4.4}$$

$$dv_t = \lambda(v_0 - v_t)dt + \xi v_t dZ, \tag{3.4.5}$$

$$\langle dW dZ \rangle = \rho dt. \tag{3.4.6}$$

Using Ito's Lemma, if the price of a derivative is $B(S_t, v_t, t)$, then

$$dB = \left(\frac{\partial B}{\partial t} dt + \frac{\partial B}{\partial S_t} \left((r_t - u_t)S_t dt + v_t F_0^{1-\beta} S_t^{\beta} dW \right) + \frac{\partial B}{\partial v_t} \left(\lambda(v_0 - v_t)dt + \xi v_t dV \right) \right) + \frac{1}{2} \left(v_t^2 F_0^{2-2\beta} S_t^{2\beta} \frac{\partial^2 B}{\partial S_t^2} + 2\rho\xi v_t^2 F_0^{1-\beta} S_t^{\beta} \frac{\partial^2 B}{\partial S_t \partial v_t} + \xi^2 v_t^2 \frac{\partial^2 B}{\partial v_t^2} \right) dt. \tag{3.4.7}$$

The discounted value of the derivative is a martingale, so its drift value must be zero. Hence, the PDE used to price options under this system is

$$rB - (r-u)S\frac{\partial B}{\partial S} - \lambda(v_0 - v)\frac{\partial B}{\partial v} - \frac{v^2 F_0^{2-2\beta} S^{2\beta}}{2}\frac{\partial^2 B}{\partial S^2} - \rho\xi v^2 F_0^{1-\beta} S^{\beta} \frac{\partial^2 B}{\partial S \partial v} - \frac{\xi^2 v^2}{2}\frac{\partial^2 B}{\partial v^2} = \frac{\partial B}{\partial t}. \tag{3.4.8}$$

We make the following implementation decisions for numerical stability:

- Price the PDE in the forward measure; in other words, use $F_t = S_t \exp\left(-\int_0^t (r(s) - u(s))ds\right)$ not S_t as the dynamic variable.
- Price non-discounted variables, and do any discounting as a post-processing step.
- Solve in x-space, where $F_t = F_0 \exp(x_t)$.
- Solve in y-space, where $v_t = v_0 y_t$.

With these changes, there is less probability for leakage out of the solution domain, and calculation of forwards is guaranteed to be exact. Therefore, the SDEs to be solved are

$$dx_t = -\frac{1}{2} v_0^2 e^{(2\beta-2)\int_0^t (r-u)ds} e^{(2\beta-2)x_t} y_t^2 dt + v_0 e^{(\beta-1)\int_0^t (r-u)ds} e^{(\beta-1)x_t} y_t dW, \tag{3.4.9}$$

$$dy_t = \lambda(1 - y_t)dt + \xi y_t dZ, \tag{3.4.10}$$

$$\langle dW dZ \rangle = \rho dt, \tag{3.4.11}$$

and the PDE is

$$\frac{\partial B}{\partial t} - \frac{1}{2} v_0^2 e^{(2\beta-2)\int_0^t (r-u)ds} e^{(2\beta-2)x} y^2 \frac{\partial B}{\partial x} + \lambda(1-y)\frac{\partial B}{\partial y} + \frac{1}{2} v_0^2 e^{(2\beta-2)\int_0^t (r-u)ds} e^{(2\beta-2)x} y^2 \frac{\partial^2 B}{\partial x^2} + \rho\xi v_0 e^{(\beta-1)\int_0^t (r-u)ds} e^{(\beta-1)x} \frac{\partial^2 B}{\partial x \partial y} y^2 + \frac{\xi^2 y^2}{2}\frac{\partial^2 B}{\partial y^2} = 0.$$

A semi-analytic approximation for the implied volatility of a European option is given by Hagan. Using our notation, this approximation is:

$$\sigma_{imp} = \frac{v' z}{x(z)} \frac{1 + \left(\frac{(1-\beta)^2}{24} v'^2 + \frac{\rho\beta\xi v'}{4} + \frac{2-3\rho^2}{24}\xi^2\right) T + \cdots}{1 + \frac{(1-\beta)^2}{24} \ln^2 F/K + \frac{(1-\beta)^4}{1920} \ln^4 F/K + \cdots}, \tag{3.4.12}$$

$$v' = v_0 (F/K)^{(1-\beta)/2}, \tag{3.4.13}$$

$$z = \frac{\xi}{v'} \ln \frac{F}{K}, \tag{3.4.14}$$

$$x(z) = \ln\left(\frac{\sqrt{1 - 2\rho z + z^2} + z - \rho}{1 - \rho}\right). \tag{3.4.15}$$

In the case of the model with mean reversion, we simply make the simple substitution

$$T \to \frac{1 - \exp(-\lambda T)}{\lambda}, \tag{3.4.16}$$

$$z \to \exp(-\lambda T/2) z, \tag{3.4.17}$$

which gives sufficiently accurate results provided $0.0 < \lambda < 2.0$.

Note that the SABR analytics is obtained using the distinguished limit in singular perturbation theory [32] which corresponds to small v_0 and ξ. As we would expect, it breaks down when v_0 and ξ are simultaneously large.

In the Numerix library, there are five virtual constructors corresponding to different calibration methods:

1. Construct with externally supplied values of the model parameters $\lambda, \xi, \rho, \beta, v_0$.

2. Use a genetic algorithm to perform a least-squares fit of the parameters $\lambda, \xi, \rho, \beta, v_0$ to given market data in the form of European option implied volatilities.

3. The model parameter β is specified externally; λ is set to zero, ξ, ρ, v_0 are found in a least-squares fit to market data.

4. The parameters $\lambda, \xi, \rho, \beta$ are specified externally; v_0 is found by exactly fitting a single ATM option.

5. Least-squares fit of the parameters $\lambda, \xi, \rho, \beta$ to a list of European options; v_0 is found by exactly fitting one further ATM option.

3.5 Heston Model

The following stochastic differential equation defines Heston's model [34] for equity or FX markets:

$$dS_t = (r(t) - u(t))S_t dt + S_t\sqrt{v_t}dW, \tag{3.5.1}$$

$$dv_t = \kappa(\theta - v_t)dt + \xi\sqrt{v_t}dV, \tag{3.5.2}$$

$$\langle dW dV \rangle = \rho dt. \tag{3.5.3}$$

Here r is the short-term rate, u is the dividend curve, κ is the rate of mean reversion, θ is the long-term variance and ξ is the volatility of volatility. Note that Heston's model reduces to the Black-Scholes model in the limit as $\kappa \to 0$, $\xi \to 0$. As before, we use the forward measure and treat the non-discounted option price as the dependent variable.

Writing $F_t = F_0 e^{x_t}$ and $v_t = v_0 y_t$, Heston's model becomes

$$dx_t = -\frac{1}{2}v_0 y_t dt + \sqrt{v_0 y_t}dW, \tag{3.5.4}$$

$$dy_t = \kappa\left(\frac{\theta}{v_0} - y_t\right)dt + \xi\sqrt{\frac{y_t}{v_0}}dV, \tag{3.5.5}$$

$$\langle dW dV \rangle = \rho dt, \tag{3.5.6}$$

and the PDE is

$$\frac{\partial B}{\partial t} - \frac{v_0 y}{2}\frac{\partial B}{\partial x} + \kappa\left(\frac{\theta}{v_0} - y\right)\frac{\partial B}{\partial y} + \frac{v_0 y}{2}\frac{\partial^2 B}{\partial x^2} + \rho\xi y\frac{\partial^2 B}{\partial x \partial y} + \frac{\xi^2 y}{2v_0}\frac{\partial^2 B}{\partial y^2} = 0.$$

The Numerix implementation further allows the local volatility to be a function of the spot price. Two additional parameters α and β are included and the initial variance v_0 is replaced with

$$\widetilde{v_0}(x) = v_0(1 + \alpha x + \beta x^2), \tag{3.5.7}$$

where $x = \log(S/S_0) - \int(r-u)$. This gives rise to the implied volatility smile as a function of strike in the usual way. Thus, this model has enough degrees of freedom to interpolate between pure local volatility models (where $\xi = 0$ and the smile is entirely due to α and β) and pure stochastic volatility models (where $\alpha = \beta = 0$).

Heston's model is unique among stochastic volatility models in that there exists an exact analytic expression for the price of a European option. This is derived using a Fourier transform method which we sketch here. The solution for the European option is sought in the form

$$B(x,v,t) = e^x P_1(x,v,t) + P_0(x,v,t). \tag{3.5.8}$$

Solutions for $P_j(x,v,t)$, $j=0,1$, are sought in Fourier space:

$$P_j(x,v,t) = \int_{-\infty}^{\infty} \hat{P}_j(k,v,t) e^{ikx} dk. \tag{3.5.9}$$

By construction, $\hat{P}_j(k,v,t)$ can be written in the affine form

$$\hat{P}_j(k,v,t) = \exp(\theta C_j(k,t) + v D_j(k,t)) P_j(k,v,0), \tag{3.5.10}$$

where the initial condition for both P_j and $\hat{P}_j$ is the pseudo-probability of the option being in-the-money:

$$P_j(x,v,0) = \Theta(x), \tag{3.5.11}$$

$$\hat{P}_j(k,v,0) = \frac{1}{ik}, \tag{3.5.12}$$

where Θ is a step function.

Finally, the expressions $C_j(k,t)$ and $D_j(k,t)$ are expressed using complex variables derived from j,k,κ,ρ,ξ (see, e.g., [30]). The calibration of the Heston model uses a genetic algorithm optimization method to perform a least-squares fit of the five parameters $v_0,\xi,\kappa,\rho,\theta$ to the list of European options supplied. The analytics for Heston's model is only exact if $\alpha=0,\beta=0$, so only this case is currently calibrated.

Note that if

$$\kappa < \frac{\xi^2}{2\theta}, \tag{3.5.13}$$

then the process for v_t can go negative, which is clearly unrealistic. If this occurs, there can be minor differences (implied volatility difference in the range 10^{-4} to 10^{-3}) between the converged prices of European options priced using the simulation versus the PDE.

3.6 Time-Dependent Heston Model (Andersen-Andreasen Version)

3.6.1 Definition

SDE

Numerix implements a transformed version of the Heston model that allows for time-dependent coefficients and is motivated by Andersen and Andreasen [2]. The SDE is:

$$dS_t = (r(t) - u(t))S_t dt + S_t \sigma(t)\sqrt{z_t} dW, \tag{3.6.1}$$

$$dz_t = \overline{\kappa}(t)(1 - z_t)dt + \overline{\xi}(t)\sqrt{z_t} dV, \tag{3.6.2}$$

$$\langle dW dV \rangle = \rho(t)dt, \tag{3.6.3}$$

$$z_t\big|_{t=0} = z_0, \qquad S_t\big|_{t=0} = S_0. \tag{3.6.4}$$

The coefficients are positive deterministic functions of time, though the explicit t-dependence may be dropped to ease notation. The change of variables $v_t = \sigma^2(t) z_t$ makes (3.6.1)–(3.6.4) equivalent to the classical version (3.5.1), where the parameters κ, θ and ξ are allowed to depend on time,[2] if

$$\overline{\kappa}(t) = \frac{\kappa\theta}{\sigma(t)^2}, \tag{3.6.5}$$

$$\overline{\xi}(t) = \frac{\xi}{\sigma(t)}, \tag{3.6.6}$$

and σ satisfies

$$\frac{\kappa\theta}{\sigma^2} = \kappa + 2\frac{\sigma'}{\sigma}.$$

We may rewrite this last constraint as

$$\left(\sigma^2\right)' = \kappa(\theta - \sigma^2), \tag{3.6.7}$$

where $\sigma^2(0) = v_0/z_0$ and $\sigma > 0$. The correlation term $\rho(t)$ is the same in both parameterizations.

Note that when κ and θ are constant, the solution to (3.6.7) is

$$\sigma^2(t) = \theta + \left(\frac{v_0}{z_0} - \theta\right) e^{-\kappa t}.$$

For general κ and θ, the solution to (3.6.7) is

$$\sigma^2(t) = e^{-\int_0^t \kappa(s)ds}\left[\sigma^2(0) + \int_0^t \kappa(s)\theta(s) e^{\int_0^s \kappa(u)du} ds\right].$$

Note also that this change of variables is time dependent, so the classical model with constant coefficients will give rise to time-dependent coefficients in this parameterization.

[2] That is, given parameters κ, θ, ξ, we may find the AA parameters $\overline{\kappa}$, $\overline{\xi}$, σ by (3.6.5)–(3.6.7). Given AA parameters $\overline{\kappa}$, $\overline{\xi}$, and σ, we can invert these relationships to find classical parameters κ, θ, and ξ. However, for arbitrary AA parameters, the corresponding classical parameters will depend on time, even though classical constant parameters give rise to time-dependent AA parameters.

Analysis of the Model and Known Results

Stochastic Variance One can think of z as a normalized variance, where the normalization is done so that the mean reversion is 1. One can also think of z as a rescaling of v by pulling out a deterministic volatility factor σ, where σ^2 is exactly the deterministic component of v_t. (Note that σ^2 follows the ordinary differential equation (3.6.7), which is the same as the pure drift ordinary differential equation for the Heston model ($dv_t = \kappa(\theta - v_t)dt$).)

The variance starts at a value $z_0 > 0$ and generally will tend towards the long-term mean of 1. The time scale on which the variance mean reverts back to 1 is controlled by the rate of mean reversion $\overline{\kappa} \geq 0$. Larger values of $\overline{\kappa}$ correspond to a stronger force of reversion back to 1. The function $\overline{\xi}$ is sometimes known as the volatility of volatility, and, together with $\overline{\kappa}$, it controls how far the random variable z_t can deviate from its center of probability: the variance of z_t at some time t is a function of $\overline{\xi}(t)$ and $\overline{\kappa}(t)$. Note that it would be possible to choose functions $\overline{\kappa}$ and $\overline{\xi}$ so that z_t fails to tend towards 1, but in general this is not realistic, and $\overline{\kappa}$ and $\overline{\xi}$ are usually piecewise-constant functions.

The distribution for the volatility process is a shifted noncentral chi-squared distribution (see [1] and [47]). Write the probability density function of the noncentral chi-squared distribution with ν degrees of freedom and noncentrality parameter λ as

$$f(x,\nu,\lambda) = e^{-\lambda/2}\sum_{j=0}^{\infty}\frac{(\lambda/2)^j}{j!2^{\frac{\nu}{2}+j}\Gamma\left(\frac{\nu}{2}+j\right)}e^{-x/2}x^{\frac{\nu}{2}+j-1} \tag{3.6.8}$$

$$= \frac{1}{2}e^{-\frac{x+\lambda}{2}}\left(\frac{x}{\lambda}\right)^{\frac{\nu-2}{4}} I_{\frac{\nu}{2}-1}\left(\sqrt{\lambda x}\right), \tag{3.6.9}$$

where

$$I_\beta(y) = \left(\frac{y}{2}\right)^\beta \sum_{j=0}^{\infty}\frac{\left(y^2/4\right)^j}{j!\Gamma(\beta+j+1)}$$

is the modified Bessel function of the first kind. Define

$$d(t) = 4\frac{\overline{\kappa}(t)}{\overline{\xi}(t)^2}, \tag{3.6.10}$$

$$K(t,s) = e^{\frac{1}{2}\int_t^s \overline{\kappa}(u)du}, \tag{3.6.11}$$

$$V^2(t) = \int_0^t \overline{\xi}(u)^2 K^2(0,u)du, \tag{3.6.12}$$

$$N(t) = 4\left(\int_0^t \overline{\xi}(u)^2 K^{-2}(u,t)du\right)^{-1}. \tag{3.6.13}$$

Then for $t \in [0,T]$, the random variable $N(t)z_t$ follows a noncentral chi-squared distribution with $d(t)$ degrees of freedom and noncentrality parameter $\lambda(t) = 4z_0/V^2(t)$ [47].[3] The probability

[3]There is an apparently typo in Equation (3.5) of [47], where the author has written the order of the Bessel function as $-(d(t)-2)/2$ instead of $(d(t)-2)/2$. The modified Bessel function is symmetric in its order only when the order is an integer (so $d(t)$ would have to be an even integer). We believe that (3.6.15) is the corrected form.

density function for $z(t)$ (conditional on z_0) is

$$\begin{aligned} p(t,z_t) &= f\left(N(t)z_t, d(t), \lambda(t)\right) \\ &= e^{-\lambda(t)/2}\sum_{j=0}^{\infty}\frac{(\lambda(t)/2)^j}{j!2^{\frac{d(t)}{2}+j}\Gamma\left(\frac{d(t)}{2}+j\right)}e^{-N(t)z_t/2}\left(N(t)z_t\right)^{\frac{d(t)}{2}+j-1} \qquad (3.6.14) \\ &= \frac{1}{2}N(t)\left(\frac{N(t)z_t}{\lambda}\right)^{\frac{d(t)-2}{4}}e^{-\frac{1}{2}\left(\lambda(t)+N(t)z_t\right)}I_{\frac{d(t)}{2}-1}\left(\sqrt{\lambda(t)N(t)z_t}\right), \qquad (3.6.15) \end{aligned}$$

where f is the noncentral chi-squared distribution function and I_α is the modified Bessel function of the first kind.

The expectation and variance of z_t conditional on z_0 are

$$\mathbb{E}\left[z_t \,|\, z_0\right] = \frac{1}{N(t)}\left(d(t)+\lambda(t)\right) \qquad (3.6.16)$$

and

$$\mathsf{Var}\left[z_t \,|\, z_0\right] = \frac{2}{N(t)^2}\left(d(t)+2\lambda(t)\right). \qquad (3.6.17)$$

If the parameters are piecewise constant with $\overline{\kappa}(t)=\kappa$, $\overline{\xi}(t)=\xi$ for all $t\in[t_i,t_{i+1})$ and some constants κ and ξ, then the expectation and variance at t_{i+1} conditional on t_i are

$$\mathbb{E}\left[z_{t_{i+1}} \,|\, z_{t_i}\right] = 1+\left(z_{t_i}-1\right)e^{-\kappa\left(t_{i+1}-t_i\right)}, \qquad (3.6.18)$$

$$\mathsf{Var}\left[z_{t_{i+1}} \,|\, z_{t_i}\right] = \frac{z_{t_i}\xi^2 e^{-\kappa\left(t_{i+1}-t_i\right)}}{\kappa}\left(1-e^{-\kappa\left(t_{i+1}-t_i\right)}\right)+\frac{\xi^2}{2\kappa}\left(1-e^{-\kappa\left(t_{i+1}-t_i\right)}\right)^2. \qquad (3.6.19)$$

Consider now the situation where the parameters may not be constant, but have limits $\lim_{t\to\infty}\overline{\kappa}(t)=\kappa$ and $\lim_{t\to\infty}\overline{\xi}(t)=\xi$. Then under sufficient conditions on the coefficients, the stationary distribution for the variance is the gamma distribution

$$p^*(z) = \frac{\alpha^\alpha}{\Gamma(\alpha)}z^{\alpha-1}e^{-\alpha z},$$

where $\alpha=\frac{2\kappa}{\xi^2}$. The mean and variance of z conditional on z_0 under the stationary distribution are

$$\mathbb{E}\left[z \,|\, z_0\right]=1, \qquad \mathsf{Var}\left[z \,|\, z_0\right]=\frac{1}{\alpha}.$$

This is again consistent with the constant-parameter case.

The Feller Condition The Feller condition $\frac{2\kappa\theta}{\xi^2}>1$ is now replaced with a time-dependent version

$$\inf_t \frac{2\overline{\kappa}(t)}{\overline{\xi}(t)^2} > 1. \qquad (3.6.20)$$

(Note that the corresponding θ term is simply 1 by (3.6.2).) Set

$$\alpha(t) = \frac{2\overline{\kappa}(t)}{\overline{\xi}(t)^2}$$

and $\alpha=\inf_t\alpha(t)$, and assume $\alpha>0$. We will refer to $\alpha(t)$ as the Feller ratio.

The primary results on this topic are due to Feller [29], who studied a class of parabolic PDEs to which the Fokker-Planck equation for the Heston model belongs. Feller considered the case of constant parameters, though the analysis was extended to the time-dependent case by Maghsoodi [47]. For simplicity, we will state the results in the case of constant parameters that Feller studied, as the parameters would be either constant or piecewise constant upon implementation. In this case, the Feller ratio $\alpha(t)$ is simply a constant $\alpha = \frac{2\overline{\kappa}}{\overline{\xi}^2}$.

The Fokker-Planck equation for the variance distribution for the Heston model is

$$\frac{\partial p}{\partial t} = \frac{\overline{\xi}^2}{2}\frac{\partial^2}{\partial z^2}(zp) - \overline{\kappa}\frac{\partial}{\partial z}\left((1-z)p\right) \tag{3.6.21}$$

with the initial condition $p(z,0) = \delta(z - z_0)$. The nature of the solutions to (3.6.21) depends on the value of α.

Feller considers (3.6.21) under three regions: $\alpha \leq 0$, $0 < \alpha < 1$, and $\alpha > 1$. The case when $\alpha \leq 0$ is not considered for the Heston model, and in this case Feller finds that the solution to the Fokker-Planck equation is norm-decreasing, so that the total probability is not always unity. The regions of interest are then $0 < \alpha < 1$ and $\alpha \geq 1$.

When $0 < \alpha < 1$, the process z_t can reach the boundary $z = 0$ in finite time. The solution to (3.6.21) depends on what boundary conditions are prescribed for the PDE at $z = 0$.

If the boundary condition $p(t,0) < \infty$ is prescribed for the PDE, then this corresponds to an absorbing boundary, where z_t can reach 0 and become trapped there. This would require a modification of our SDE (3.6.2), where we set $z_t = 0$ for all $t > \tau$, where $\tau = \inf\{t : z_t = 0\}$. In this case, the density function that solves the Fokker-Planck equation is norm-decreasing, because once the variance hits zero, it is stuck there. This case is not considered for the Heston model.

If the boundary condition

$$\lim_{z\to 0}\left[\frac{\overline{\xi}^2}{2}\frac{\partial}{\partial z}(zp) - \overline{\kappa}(z-1)p\right] = 0 \tag{3.6.22}$$

is prescribed for the PDE, then this corresponds to a reflecting boundary, where z_t can reach 0, but will immediately become positive due to the positive drift. This condition is known as a zero-flux condition, because the term on the left-hand side of (3.6.22) represents the flux through the boundary. Requiring this quantity to be zero is to require that no probability density can leak through the boundary. In this case, the solution to (3.6.21) is given by (3.6.15) (with constant $\overline{\kappa}$ and $\overline{\xi}$), and this solution is positive and integrates to 1 for all time. No additional assumptions need to be made for the stochastic process. Therefore, condition (3.6.22) is the natural boundary condition to apply to the Fokker-Planck equation in the region $0 < \alpha < 1$.

When $\alpha > 1$, the process z_t cannot reach zero. No boundary condition can be prescribed for the PDE at $z = 0$. In this case, the solution is again given by (3.6.15) with constant $\overline{\kappa}$ and $\overline{\xi}$.

The parameters in (3.6.15) take on a simpler form as $K(t,s) = e^{\frac{1}{2}\overline{\kappa}(s-t)}$:

$$d(t) = d = 4\frac{\overline{\kappa}}{\overline{\xi}^2} = 2\alpha, \qquad N(t) = \frac{2\alpha}{1-e^{-\overline{\kappa}t}}, \qquad \lambda(t) = \frac{2\alpha z_0}{e^{\overline{\kappa}t}-1}.$$

The probability distribution is still the shifted noncentral chi-squared distribution

$$p(t,z) = \frac{1}{2}N(t)\left(\frac{N(t)z}{\lambda}\right)^{\frac{\alpha-1}{2}} e^{-\frac{1}{2}\left(\lambda(t)+N(t)z\right)} I_{\alpha-1}\left(\sqrt{\lambda(t)N(t)z}\right).$$

The stationary Fokker-Planck equation, which can be written as

$$0 = \left[\frac{\partial^2}{\partial z^2} (zp) - \alpha \frac{\partial}{\partial z} \left((1-z)p \right) \right],$$

has the solution

$$p^*(z) = \frac{\alpha^\alpha}{\Gamma(\alpha)} z^{\alpha-1} e^{-\alpha z}$$

for all $\alpha > 0$, which can be directly verified. Moreover, we can see that $\lim_{t\to\infty} p(t,z) = p^*(z)$ pointwise in z when $\overline{\kappa}$ and $\overline{\xi}$ are constant. A similar statement could be made for general $\overline{\kappa}$ and $\overline{\xi}$, though assumptions would have to be made about their limits and regularity.

Note that the solutions to the Fokker-Planck equation exhibit different behavior for $0 < \alpha < 1$ and $\alpha > 1$. Indeed, when $0 < \alpha < 1$, the probability density approaches infinity as $z \to 0$, even though the total area is still 1. This is clear in the stationary distribution p^*, though it is also evident by examining the leading term of $I_{\alpha-1}$ in $p(t,z)$.

We see that the variance z_t is always nonnegative. The process can only reach zero when the Feller condition is violated. However, when implementing a discretization method for numerical calculation, it is possible, depending on the choice of method, for the process z to take on negative values, even though this is not possible in the continuous version. This is true whether or not the Feller condition holds. To handle this, one can attempt to construct a discretization scheme that does not allow negative values for the variance, or one can can impose reflecting ($z_t = -z_t$) or absorbing ($z_t = 0$) boundary conditions. (Absorbing boundary conditions should generally be avoided in light of the above discussion.) Numerix adopts the former strategy of choosing a discretization method that guarantees a positive value for the variance at all times. Whatever method is chosen, the Feller ratio α is an important quantity in the analysis. See Section 3.6.2 for further discussion.

The Mixing Theorem The mixing theorem of Romano and Touzi expresses the value of a call option under (3.6.1)–(3.6.4) as a weighted sum of Black-Scholes values of that call option with an effective strike and effective volatility, where the effective variables depend only on the variance process.

Theorem 3.6.1 (Romano and Touzi, 1997)**.** *Let $C(S_0, z_0, T)$ be the call option price under (3.6.1)–(3.6.4), and let c be the Black-Scholes price for the same option. Then*

$$C(S_0, V_0, T) = \mathbb{E}\left[c\left(S_{\text{eff}}, z_{\text{eff}}, T \right) \right], \tag{3.6.23}$$

where

$$S_{\text{eff}} = S_0 \exp\left(-\frac{1}{2} \int_0^T \rho^2 \sigma^2 z dt + \int_0^T \rho\sigma\sqrt{z} dW \right), \tag{3.6.24}$$

$$z_{\text{eff}} = \frac{1}{T} \int_0^T \left(1 - \rho^2 \right) z dt. \tag{3.6.25}$$

In the special case that $\rho = 0$, we have $S_{\text{eff}} = S_0$ and $z_{\text{eff}} = \frac{1}{T}\int_0^T z dt$. Then the expectation in

(3.6.23) simplifies and we have, writing $U_T = \int_0^T z dt$,

$$C(S_0, V_0, T) = \int_0^\infty c\left(S_0, \frac{U_T}{T}, T\right) p(T, U_T) dU_T,$$

where p is the probability density function for the variance conditional on z_0 given by (3.6.15).

In [42], Lewis shows that the zero-correlation case implies that the implied volatility curve is symmetric in the strike. This and other implications of the mixing theorem are discussed in [42]. Lewis also establishes a version of Theorem 3.6.1 for arbitrary payoff functions.

Distribution for the Underlying Process The full Fokker-Planck equation for the Heston model is

$$\frac{\partial P}{\partial t} = \overline{\kappa}\frac{\partial}{\partial z}\left[(z-1)P\right] + \frac{1}{2}\frac{\partial}{\partial x}(zP) + \rho\overline{\xi}\frac{\partial^2}{\partial x \partial z}(zP) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(zP) + \frac{\overline{\xi}^2}{2}\frac{\partial^2}{\partial z^2}(zP), \tag{3.6.26}$$

where we have set $x = \log(S/S_0) - (r-u)t$. This equation can be solved by taking Fourier a transform in the x-variable and Laplace transform in the z-variable. The details are given in [26]. The authors express the probability distribution P (conditional on z_0) as a Fourier transform

$$P\left(x, v \,|\, z_0\right) = \int_{-\infty}^{\infty} e^{i\omega x}\overline{P}\left(\omega, z \,|\, z_0\right) d\omega, \tag{3.6.27}$$

where the time variable t is treated as a parameter and the dependence is omitted. The function $\overline{P}$ is found by inverting the following Laplace transform

$$\widetilde{P}(\omega, \eta) = \int_0^\infty e^{-\eta z}\overline{P}(\omega, z, \,|\, z_0) dz. \tag{3.6.28}$$

The transform $\widetilde{P}$ can be solved for, resulting in

$$\widetilde{P}(\omega, \eta) = \exp\left[-\psi z_0 + \frac{\overline{\kappa}(\Gamma - 1)t}{\xi^2} - \frac{2\overline{\kappa}}{\xi^2}\log\left[\frac{\Upsilon - e^{-\Omega t}}{\Upsilon - 1}\right]\right],$$

where

$$\Gamma = \kappa + i\rho\overline{\xi}\omega,$$

$$\Omega = \sqrt{\Gamma^2 + \overline{\xi}^2(\omega^2 - i\omega)},$$

$$\Upsilon = 1 + \frac{2\Omega}{\overline{\xi}^2\eta + (\Gamma - \Omega)},$$

$$\psi = \frac{2\Omega}{\overline{\xi}^2}\frac{1}{\Upsilon e^{\Omega t} - 1} - \frac{\Gamma - \Omega}{\overline{\xi}^2}.$$

This formulation is cumbersome, but some limiting cases are discussed in [26].

Pricing PDE

Let $C = C(S, z, t)$ be the price of an option, where t is time and C and z satisfy (3.6.1)–(3.6.4). Then C satisfies

$$-\frac{\partial C}{\partial t} = rC + \frac{1}{2}\sigma^2 z S^2 \frac{\partial C^2}{\partial S^2} + \sigma\bar{\xi}\rho z S \frac{\partial^2 C}{\partial S \partial z} + \frac{1}{2} z \bar{\xi}^2 \frac{\partial^2 C}{\partial z^2} + (r-u)S\frac{\partial C}{\partial S} + \bar{\kappa}(1-z)\frac{\partial C}{\partial z}, \tag{3.6.29}$$

where the arguments have been suppressed, and the boundary conditions are determined by the type of option.

3.6.2 Implementation

In the Numerix implementation of the time-dependent Heston model, the SDE (3.6.1)–(3.6.4) is used and σ, $\bar{\kappa}$, $\bar{\xi}$, ρ and z_0 are input parameters. These parameters can either be specified directly or found by calibrating to instruments. The CrossAsset XL interface `Heston: Time-dependent Coefficients, Calibrated to Options` uses the following headers:

- CORRELATION: ρ
- SIGMA1: σ
- SV SPOT: z_0
- SV REVERSION: $\bar{\kappa}$
- SV VOL: $\bar{\xi}$

Note that users are allowed to calibrate z_0, but z_0 can actually scale out of the SDE. For each model of the form (3.6.1)–(3.6.4) with z_0 free, an equivalent model of the same form can be constructed with $z_0 = 1$ and different (scaled) choices of σ_t, $\bar{\kappa}_t$ and $\bar{\xi}_t$. However, the scaling is a time-dependent scaling, and constant (or piecewise constant) coefficients can map to time-dependent (or non-piecewise-constant) coefficients. Because the coefficients will be represented by piecewise-constant functions in the software, allowing z_0 to be different from 1 does allow for different behavior. Note also that the process mean reverts to 1, so generally taking z_0 away from 1 will change the beginning dynamics more.

Monte Carlo Simulation

Discretization of the Variance Process While the volatility process z in (3.6.1)–(3.6.4) cannot be negative, it is possible for a discretization of the process to reach negative values of z. Therefore, when simulating the volatility process for Monte Carlo, care must be taken when applying a discretization method to (3.6.2) for sampling. It is easy to see that a simple Euler discretization of (3.6.2) runs the risk of calculating a negative value for some $z(t_i)$.

Numerix uses a strategy outlined by Anderson [1] for preventing negative stochastic volatility. The idea is to find a new process $\hat{z} \geq 0$ which matches the moments of z and is computationally efficient to sample from. Because the coefficients are piecewise constant, we may assume that they are in fact constants κ and ξ on a particular interval $[t_i, t_{i+1})$ of the discretization grid. We have the explicit expressions for the mean and variance of $z_{t_{i+1}}$ conditional on z_{t_i} in (3.6.18)

and (3.6.19), so we take $\hat{z}$ to be of a form that must be nonnegative and has two degrees of freedom to match that mean and variance.

Following [1], $\hat{z}$ is of the form

$$\hat{z}(t_{i+1}) = a(b+Z)^2, \tag{3.6.30}$$

where Z is a draw from a standard Gaussian distribution, and a and b will be chosen to match the mean and variance of $z_{t_{i+1}}$. Set

$$m = \mathbb{E}\left[z_{t_{i+1}} \mid z_{t_i}\right] = 1 + (z_{t_i} - 1)\, e^{-\kappa\left(t_{i+1}-t_i\right)},$$

$$s^2 = \mathrm{Var}\left[z_{t_{i+1}} \mid z_{t_i}\right] = \frac{z_{t_i}\xi^2 e^{-\kappa\left(t_{i+1}-t_i\right)}}{\kappa}\left(1 - e^{-\kappa\left(t_{i+1}-t_i\right)}\right) + \frac{\xi^2}{2\kappa}\left(1 - e^{-\kappa\left(t_{i+1}-t_i\right)}\right)^2,$$

and

$$\psi = \frac{s^2}{m^2}.$$

Note that we are only considering the interval $[t_i, t_{i+1})$ and all parameters are constant on this interval (the dependence on i is suppressed). It can be shown that $0 \le \psi \le \frac{\xi^2}{2\kappa}$. (Note the upper bound is the reciprocal of the Feller ratio.)

Andersen shows in [1] that if $\psi \le 2$, then we may set

$$b^2 = \frac{2}{\psi} - 1 + \sqrt{\frac{2}{\psi}}\sqrt{\frac{2}{\psi} - 1}$$

and

$$a = \frac{m}{1+b^2}$$

in order to guarantee $\mathbb{E}\left[\hat{z}(t_{i+1})\right] = m$, $\mathrm{Var}\left[\hat{z}(t_{i+1})\right] = s^2$. Note that $\psi \le 2$ must hold in order to find a value for b. If $\psi > 2$, then we cannot match the moments using this form, and we must find a different technique. Note that if the Feller condition $\frac{2\kappa}{\xi^2} \ge 1$ is satisfied, then we will have $\psi \le 2$. If the Feller condition is not satisfied, ψ may or may not be bounded by 2.

The idea when $\psi > 2$ is to approximate the distribution 3.6.15 by one with a simpler cumulative distribution function that allows for a fast sampling scheme. We use an approximate density of the form

$$f(x) = p\delta(0) + \beta(1-p)e^{-\beta x}$$

for $x \ge 0$, where δ is the Dirac delta function and $p \in [0,1]$, $\beta \ge 0$ are to be determined by moment matching. Anderson shows that as long as $\psi \ge 1$, we may set

$$p = \frac{\psi - 1}{\psi + 1}$$

and

$$\beta = \frac{2}{m(\psi+1)}$$

to give $\mathbb{E}\left[\hat{z}(t_{i+1})\right] = m$, $\mathrm{Var}\left[\hat{z}(t_{i+1})\right] = s^2$. Given these values for p and β, we select a draw for $\hat{z}(t_{i+1})$ by finding the cumulative distribution function

$$F(x) = p + (1-p)\left(1 - e^{-\beta x}\right)$$

and choosing

$$\hat{z}(t_{i+1}) = F^{-1}(U), \tag{3.6.31}$$

where U is a draw from a uniform distribution. The inverse F^{-1} is readily computable, so this scheme is computationally fast. Note that all calculations depend on the previous value z_{t_i}.

Then we have two different moment-matching methods: one that is valid for $\psi \leq 2$ and another that is valid when $\psi \geq 1$. These two regions overlap, so we merely need a technique of switching between them. The idea is to pick $\psi_c \in [1, 2]$ that will serve as the cutoff level. When $\psi \geq \psi_c$, we choose $\hat{z}(t_{i+1})$ by (3.6.30) and when $\psi < \psi_c$, we use (3.6.31).

Using this method, we guarantee that the variance is always nonnegative, whether or not the Feller condition is met. In [1], Anderson shows some graphs of the distribution functions for this method compared to the exact distribution.

3.6.3 Calibration

For calibration, we use the following technique for analytical calculation of option prices. We calculate the option price

$$e^{-\int_0^t r_s ds}\mathbb{E}\left[(S_t - K)^+\right]$$

by using a Laplace transform of the characteristic function, which is known analytically. Consider the forward price

$$F_t = S_t e^{-\int_0^t (r-u)ds},$$

which satisfies

$$dF_t = F_t \sigma(t)\sqrt{z_t} dW.$$

We write the option price as

$$e^{-\int_0^t r ds}\mathbb{E}\left[(S_t - K)^+\right] = e^{-\int_0^t u ds}\mathbb{E}\left[\left(F_t - Ke^{-\int_0^t (r-u)ds}\right)^+\right] = e^{-\int_0^t u ds}\mathbb{E}\left[(F_t - K_{\text{eff}})^+\right], \tag{3.6.32}$$

where

$$K_{\text{eff}} = Ke^{-\int_0^t (r-u)ds}$$

is the effective strike. Then the option price, written in terms of the forward, looks like the initial price where K_{eff} is used as the strike.

Now consider the process

$$y_t = \ln\left(\frac{F_t}{F_0}\right),$$

which satisfies

$$dy_t = -\frac{1}{2}\sigma(t)^2 z_t dt + \sigma(t)\sqrt{z_t} dW.$$

Then, due to affine properties of the model, one can write the ODE for the characteristic function

$$\phi(T, \zeta) = \mathbb{E}\left[e^{\zeta y_T}\right]$$

of the process y_t. The characteristic function can be used for the price calculation via a Laplace

transform.[4] Indeed,

$$(F_0 e^{y_t} - K)^+ = \frac{K}{2\pi i} \int_{C^+} \frac{1}{\zeta(\zeta-1)} \exp\left(\zeta y_t - \zeta \log\left(\frac{K}{F_0}\right)\right) d\zeta, \tag{3.6.33}$$

where the contour C^+ is parallel to the imaginary axis and passes to the right of the integrand poles (0 and 1). Then the option price can be represented in terms of a moment generating function (MGF)

$$\mathcal{K}(\zeta) \equiv \log \phi(T, \zeta),$$

i.e.,

$$\mathbb{E}\left[(F_0 e^{y_t} - K)^+\right] = \frac{K}{2\pi i} \int_{C^+} \frac{1}{\zeta(\zeta-1)} \exp\left[\mathcal{K}(\zeta) - \zeta \log\left(\frac{K}{F_0}\right)\right] d\zeta. \tag{3.6.34}$$

The standard approach is to compute this integral numerically. As we will see below, it is possible for certain cases to reduce a complex integral to a real one. We take the integral along the line $\zeta = \frac{1}{2} + i\omega$ with real ω, taking into account the contribution of the pole at $\zeta = 1$. This gives

$$\mathbb{E}\left[(F_0 e^{y_t} - K)^+\right] = F_0 - \frac{K}{2\pi} \int_{-\infty}^{\infty} \frac{1}{\omega^2 + \frac{1}{4}} \exp\left[\mathcal{K}\left(\frac{1}{2} + i\omega\right) - \left(i\omega + \frac{1}{2}\right) \log\left(\frac{K}{F_0}\right)\right] d\omega. \tag{3.6.35}$$

If the values of $\mathcal{K}\left(\frac{1}{2} + i\omega\right)$ are real, one can make the replacement

$$\exp\left[-\left(i\omega + \frac{1}{2}\right) \log\left(\frac{K}{F_0}\right)\right] \to \left(\frac{K}{F_0}\right)^{-\frac{1}{2}} \cos\left(\omega \log\left(\frac{K}{F_0}\right)\right).$$

In order to calculate the characteristic function $\phi(T, \zeta)$, we consider the conditional expectation

$$u(t; z, y) = \mathbb{E}\left[e^{\zeta y_T} \,\middle|\, z_t = z, y_t = y\right],$$

which satisfies the PDE

$$\frac{\partial u}{\partial t} - \frac{1}{2}\sigma_t^2 z \frac{\partial u}{\partial y} + \frac{1}{2}\sigma_t^2 z \frac{\partial^2 u}{\partial y^2} + \overline{\kappa}_t(1-z)\frac{\partial u}{\partial z} + \frac{1}{2}\overline{\xi}_t^2 z \frac{\partial^2 u}{\partial z^2} + \sigma_t \overline{\xi}_t \rho_t z \frac{\partial^2 u}{\partial y \partial z} = 0,$$

which we obtain by imposing a zero-drift condition. Looking for the solution in affine form,

$$u(t; z, y) = e^{A_t + B_t z + C_t y},$$

and imposing the conditions $A(T) = B(T) = 0$, $C(T) = \zeta$, we have

$$\begin{cases} C_t = \zeta, \\ B_t' = -\frac{1}{2}\overline{\zeta}_t^2 B_t^2 + \left(\overline{\kappa}_t - \sigma_t \overline{\xi}_t \rho_t \zeta\right) B_t - \sigma_t^2 \frac{\zeta^2 - \zeta}{2}, \\ A_t' = -\overline{\kappa}_t B_t. \end{cases} \tag{3.6.36}$$

The system (3.6.36) can be solved numerically or, for piecewise-constant parameters, iter-

[4] To simplify formulae, we replace K_{eff} with K and remove the multiplier $e^{-\int_0^t u ds}$ in the option price formula (3.6.32).

atively using a Riccati solution.

Coming back to the above remark about the possibility of real instead of complex integration, we should mention that, for zero correlation, the moment generating function $\mathcal{K}\left(\frac{1}{2}+i\omega\right)$ is real and we can perform real integration. However, complex integration is necessary for the case of nonzero correlation.

Finally, to calculate the European option price, one performs numerical integration on (3.6.35), wherein one finds $\mathcal{K}\left(\frac{1}{2}+i\omega\right)$ i for each value of ω that solves the system (3.6.36).

3.7 Bates Model

The Bates model extends the time-dependent Heston stochastic volatility model (3.6.1)–(3.6.4) by adding a jump component to the equity price or foreign exchange rate process. This is useful because adding jumps allows the user to fit short-dated options with realistic parameters, since market prices for short-dated options are higher in the tails than diffusion models predict. The model can be used for pricing options on either equity or foreign exchange rate. In the equity case, S_t and u_t denote the stock price and the dividend yield, respectively, and in the FX case, S_t and u_t denote the exchange rate and the foreign interest rate, respectively. In both cases, r_t denotes the domestic interest rate. The SDE is simply the time-dependent (AA) Heston PDE with an additional (compensated) jump term:

$$\frac{dS_t}{S_t} = (r_t - u_t - \lambda_t m_J)\,dt + \sigma_t\sqrt{z_t}dW + JdN_t, \tag{3.7.1}$$

$$dz_t = \overline{\kappa}_t(1-z_t)dt + \overline{\xi}_t\sqrt{z_t}dV, \tag{3.7.2}$$

$$\langle dW\,dV\rangle = \rho_t dt, \tag{3.7.3}$$

$$z_t\big|_{t=0} = z_0, \qquad S_t\big|_{t=0} = S_0. \tag{3.7.4}$$

Here N_t is a Poisson process with intensity λ_t and J is a random jump size which can be normally distributed (as with Merton jumps) or have a double exponential distribution (as with Kou jumps). The process N_t and J are independent of each other and of the Brownian motions W and V. The coefficients σ_t, $\overline{\kappa}_t$, $\overline{\xi}_t$ and ρ_t are the volatility level function, the rate of mean reversion, the volatility of volatility and the correlation as in the time-dependent Heston model. All the parameters, including λ_t, r_t and q_t are allowed to depend deterministically on time.

Note that there are two sources of randomness in the jumps. First is whether or not a jump occurs in a dt interval, which is represented by dN_t. Second is, given that a jump does occur, what the size of the jump is, which is represented by J. There are two choices for the jump sizes in the Numerix implementation of the Bates model [18]: Merton jumps [51] and Kou jumps [41].

For Merton jumps, which are recommended, the jump size J follows a normal distribution

$\mathcal{N}(\mu_J, \sigma_J)$ with

$$d\mu_J(y) = \frac{1}{\sqrt{2\pi}\sigma_J} e^{-\frac{(y-\mu_J)^2}{2\sigma_J^2}} dy, \tag{3.7.5}$$

$$\phi_J(u) = e^{\mu_J u + \frac{\sigma_J^2 u^2}{2}}, \tag{3.7.6}$$

$$m_J = \phi_J(1) - 1 = e^{\mu_J + \frac{\sigma_J^2}{2}} - 1. \tag{3.7.7}$$

For Kou jumps, which can be used for research purposes, the jumps follows a double-exponential distribution with

$$d\mu_J(y) = \left(p\lambda_+ e^{-\lambda_+ y} 1_{y\geq 0} + (1-p)\lambda_- e^{-\lambda_- |y|} 1_{y<0}\right) dy, \tag{3.7.8}$$

$$\phi_J(u) = \frac{p\lambda_+}{\lambda_+ - u} + \frac{(1-p)\lambda_-}{\lambda_- + u}, \tag{3.7.9}$$

$$m_J = \phi_J(1) - 1 = \frac{p\lambda_+}{\lambda_+ - 1} + \frac{(1-p)\lambda_-}{\lambda_- + 1} - 1, \tag{3.7.10}$$

where u must satisfy $-\lambda_- < Re(u) < \lambda_+$.

Note that the time-dependent interface for the Bates model, in which the parameters can either be specified or are found by calibrating to options, uses the SDE (3.7.1)—(3.7.4). There are also interfaces for both Merton and Kou jumps where users input the parameters directly (with no calibration), and this interfaces extends the classical Heston model (3.5). The SDE in this case is

$$\frac{dS_t}{S_t} = (r_t - u_t - \lambda m_J)\, dt + \sqrt{v_t} dW + J dN_t, \tag{3.7.11}$$

$$dv_t = \kappa(\theta - v_t)dt + \xi\sqrt{v_t} dV, \tag{3.7.12}$$

$$\langle dW dV \rangle = \rho dt. \tag{3.7.13}$$

The relationship between the volatility processes is $v_t = \sigma_t^2 z_t$, and the mapping of the coefficients is the same as in the Heston model (see Section 3.7). Note also that because the CAXL interface extends the corresponding Heston interface, the local volatility terms α and β are included, but in this case they should be set to 0.

The PDE that is used to price options is similar to that of the Heston model, with additional terms for the jump part, which gives rise to a partial integro-differential equation. If we set $x = \log(S/S_0)$ and consider an option value $C(x, z, t)$, then, dropping the functional dependence for ease of notation, we have

$$\frac{\partial C}{\partial t} + \left(r - u - \frac{\sigma^2 z}{2} - \lambda m_J\right)\frac{\partial C}{\partial x} + \frac{\sigma^2 z}{2}\frac{\partial^2 C}{\partial x^2} + \overline{\kappa}(1-z)\frac{\partial C}{\partial z} + \frac{\overline{\xi}^2 z}{2}\frac{\partial^2 C}{\partial z^2}$$
$$+ \rho\sigma\overline{\xi} z \frac{\partial^2 C}{\partial x \partial z} + \lambda \int_{\mathbb{R}} \left[C(x+y, z, t) - C(x, z, t)\right] d\mu_J(y) = 0. \tag{3.7.14}$$

3.8 Local Stochastic Volatility Model

3.8.1 Definition

The implied volatility surface of a stock or foreign exchange rate is usually V-shaped at early horizons, with a distinct minimum at the ATM strike (see, e.g., [30]). Such a structure is commonly attributed to the effect of market motions that are not captured by Brownian (Black-Scholes) dynamics. Models with stochastic volatility (SV) and/or jumps (e.g., Heston or Bates) are able to reproduce the implied volatility surface properties at early horizons and close to ATM, but not necessarily elsewhere. To allow the model to fit the market at all horizons and at all values of moneyness, these dynamics are augmented with local volatility (LV). The resulting local stochastic volatility (LSV) model has the advantages of both types of models. It inherits the realistic dynamics of a stochastic volatility model and has the flexibility of a local volatility model when it comes to achieving a close fit to European quotes. The idea is to achieve a coarse fit of the market data with a Heston-like model, and then finalize it with a relatively gentle LV structure.

Two basic types of stochastic volatility are supported by the LSV model: an ordinary CIR process and a lognormal process.

SDE

The default flavor for the LSV model has a CIR process for the stochastic volatility. In this case, the model is essentially a Heston model (with the AA parameterization) augmented by a local volatility function $\sigma(t,S)$. The SDE for the LSV model with CIR volatility is

$$\frac{1}{S}dS = \left(r_d(t) - r_f(t)\right)dt + \lambda(t)\sigma(t,S)\sqrt{z}dW_S \tag{3.8.1}$$

$$dz = a(t)(1-z)dt + \gamma(t)\sqrt{z}dW_z, \tag{3.8.2}$$

$$dW_S dW_z = \rho(t)dt \tag{3.8.3}$$

Here r_d and r_f are the domestic and foreign interest rates in the case of an FX model. For equity, these are replaced with the interest rate and dividend yield, respectively. The function λ is a volatility level function, σ is the local volatility, a is the reversion speed, γ is the volatility of volatility and ρ is the correlation.

Numerix has also implemented lognormal volatility for the stochastic volatility. In this case, the SDEs are

$$\frac{1}{S}dS = \left(r_d(t) - r_f(t)\right)dt + \lambda(t)\sigma(t,S)e^z dW_S \tag{3.8.4}$$

$$dz = -a(t)zdt + \gamma(t)dW_z, \tag{3.8.5}$$

$$dW_S dW_z = \rho(t)dt \tag{3.8.6}$$

In either case, we simulate the underlying process with

$$X = \log\frac{S}{S_0}$$

instead of S. The process X follows

$$dX = \left(r_d(t) - r_f(t)\right) dt - \frac{1}{2}\lambda^2(t)\sigma^2\left(t, S_0 e^X\right) z dt + \lambda(t)\sigma\left(t, S_0 e^X\right)\sqrt{z} dW_S \tag{3.8.7}$$

in the case of CIR volatility and

$$dX = \left(r_d(t) - r_f(t)\right) dt - \frac{1}{2}\lambda^2(t)\sigma^2\left(t, S_0 e^X\right) e^{2z} dt + \lambda(t)\sigma\left(t, S_0 e^X\right) e^z dW_S, \tag{3.8.8}$$

in the case of lognormal volatility, where $X_0 = 0$.

Pricing PDE

Let $C = C(S, z, t)$ be the price of an option at time t and with S and z satisfying (3.8.1)–(3.8.2). Then C satisfies

$$-\frac{\partial C}{\partial t} = \frac{1}{2}\lambda^2\sigma^2 z S^2\frac{\partial^2 C}{\partial S^2} + \lambda\sigma\gamma\rho z S\frac{\partial^2 C}{\partial S \partial z} + \frac{1}{2}\gamma^2 z\frac{\partial^2 C}{\partial z} + (r_d - r_f)S\frac{\partial C}{\partial S} + a(1-z)\frac{\partial C}{\partial z} + r_d C. \tag{3.8.9}$$

For the lognormal flavor (3.8.4)–(3.8.5), we have

$$-\frac{\partial C}{\partial t} = \frac{1}{2}\lambda^2\sigma^2 e^{2z} S^2\frac{\partial^2 C}{\partial S^2} + \lambda\sigma\gamma\rho e^z S\frac{\partial^2 C}{\partial S \partial z} + \frac{1}{2}\gamma^2\frac{\partial^2 C}{\partial z} + (r_d - r_f)S\frac{\partial C}{\partial S} + -az\frac{\partial C}{\partial z} + r_d C. \tag{3.8.10}$$

Calibration

The calibration of the LSV model is done in two pases. First, the $\sigma(t, S)$ term is omitted, and the resulting model is calibrated to the calibration instruments. The calibration algorithm then corrects the volatility level function $\lambda(t)$ with the local volatility factor $\sigma(t, S)$. The implied volatility surface of the calibrating instruments is used an an initial guess for $\sigma(t, S)$. During this pass, $a(t)$ is not recalibrated, while $\rho(t)$ and $\gamma(t)$ have the option to be recalibrated. The calibration of ρ and γ can be done through a mixing fraction.

The LSV model can be calibrated to European options and barrier instruments. The calibration instruments are automatically filtered to include only those instruments whose exercise probability exceeds a certain threshold (which can be selected).

3.9 Equity Convertible Model

This one-factor equity convertible model can be viewed as the extension of a local volatility model to take into account the possibility of an asset default.[5] The state variable of the model—the stock value S_t—obeys the stochastic differential equation

$$\frac{dS_t}{S_t} = (r(t) - q(t) + \lambda(t, S_{t-}))dt + \sigma(t, S_{t-})dW_t - dN_t, \tag{3.9.1}$$

where $r(t)$ and $q(t)$ are the interest and dividend rates, respectively, $\sigma(t, S)$ is the local volatility, dW corresponds to the standard Brownian motion and $\lambda(t, S)$ is the intensity of the default

[5] For a discussion of this approach, see for example [3].

process N_t. The stock price jumps to zero whenever default occurs, wherein $N_t = 1$. We assume the following form of the default intensity dependence on the stock:

$$\lambda(t, S) = a(t)(S_0/S)^{\gamma}, \tag{3.9.2}$$

where S_0 is the initial value of the stock and the exponents γ serves to regulate the strength of correlations between the stock price and default intensity. The function $a(t)$ is determined from the requirement of matching the survival curve by using forward induction.

3.10 Universal Local Stochastic Volatility with Jump Model

The universal local stochastic volatility with jump (ULSVJ) model is a framework for general jump diffusion models. Users are able to combine a number of desired features and build a range of models, including the following models:

- Black-Scholes with and without jumps
- Local volatility (Dupire) with and without jumps
- Heston with and without jumps
- SV CEV with and without jumps
- Local stochastic volatility (LSV) with and without jumps

Currently, only Merton jumps in the underlying spot process are implemented.

The ULSVJ framework is described by the system of SDEs

$$\frac{dS_t}{S_t} = (r - q - \lambda m_J)\, dt + \sigma(t, S_t)\sqrt{v_t} dW_t + J dN_t \tag{3.10.1}$$

$$dv_t = \kappa(\theta - v_t)dt + \xi v_t^{\beta} dZ_t \tag{3.10.2}$$

$$dW_t dZ_t = \rho dt \tag{3.10.3}$$

subject to the initial conditions $v_{t=0} = v_0$, $S_{t=0} = S_0$. Here:

- r is the risk-neutral short rate of the numeraire currency.
- q is the forward implied short rate of the foreign currency for FX instruments, or the dividend for EQ instruments.
- κ is the rate of mean reversion.
- θ is the long-term variance.
- ξ is the volatility of volatility.
- σ is a deterministic local volatility function.
- β controls the type of distribution of the volatility; special cases are $\beta = 1$ for lognormal, $\beta = 0$ for normal and $\beta = 1/2$ for CIR.
- ρ is the correlation between the Brownian motions for the price and volatility processes.

J is the jump size, which (for Merton jumps) follows a normal distribution with mean μ_J and volatility σ_J.

m_J is the average of the relative price jump size (not the same as μ_J).

N_t is the Poisson process with intensity λ (assumed to be constant) for the jumps.

Note also that r, q, κ, θ, ξ and ρ can be deterministic functions of time (though the time-dependence will be suppressed).

In the equity case, the dividend can be defined in different formats, including continuous or discontinuous dividend yield or cash dividends. In all these cases, the REPO (cost of carry) curve can be specified by the user.

Merton jumps are normally distributed with Lévy measure

$$\nu(dx) = \lambda \frac{1}{\sqrt{2\pi}\sigma_J} e^{-\frac{(x-\mu_J)^2}{2\sigma_J^2}} dx. \tag{3.10.4}$$

Option Pricing in the ULSVJ Framework

The ULSVJ framework supports the following types of options:

- By payoff:
  - European
  - American
  - Bermudan
  - Digital cash-or-nothing, asset-or-nothing
- By boundary conditions:
  - Vanilla
  - Single and double barrier with or without rebates
  - No-touch, one-touch, double-no-touch

Any combination of payoff and boundary-condition options is also supported. The computation of Greeks is supported as well.

Let S_t be the underlying process. Options written on S_t are priced using a standard martingale approach. Set $x = \log(S_t/S_0)$ and $\tau = T - t$, where T is the maturity of the option. The value of the option, C, satisfies

$$rC = -\frac{\partial C}{\partial \tau} + \left(r - q - \frac{\sigma^2 v}{2}\right)\frac{\partial C}{\partial x} + \kappa(\theta - v)\frac{\partial C}{\partial v} + \frac{\sigma^2 v}{2}\frac{\partial^2 C}{\partial x^2} + \xi\sigma\rho v^{\beta+1/2}\frac{\partial^2 C}{\partial x \partial v} + \frac{\xi^2 v^{2\beta}}{2}\frac{\partial^2 C}{\partial v^2} + \int_{\mathbb{R}}\left[C(x+y,v,t) - C(x,v,t) - (e^y - 1)\frac{\partial C}{\partial x}(x,v,t)\right]\nu(dy) \tag{3.10.5}$$

where $C = C(x, v, \tau)$ and $\sigma = \sigma(x, \tau)$ when the arguments are suppressed. The boundary condition is $C(x, v, T) = h(x)$ for some option payoff h, in addition to any other boundary conditions that depend on the type of option. Note that the jump term in the drift term of (3.10.1) has been

absorbed into the integral. The solutions of (3.10.5) usually belong to the class of viscosity solution.

Note that the shift operator $T_a : L^2 \to L^2$ given by

$$T_a f(x) = f(x+a)$$

can be represented by

$$T_a = \exp\left(a\frac{\partial}{\partial x}\right).$$

We may then write

$$\int_{\mathbb{R}} \left[C(x+y,v,t) - C(x,v,t) - (e^y - 1)\frac{\partial C(x,v,t)}{\partial x}\right] \nu(dy) = \mathcal{J}C(x,v,t),$$

where

$$\mathcal{J} \equiv \int_{\mathbb{R}} \left[e^{y\frac{\partial}{\partial x}} - 1 - (e^y - 1)\frac{\partial}{\partial x}\right] \nu(dy). \tag{3.10.6}$$

Formally, one can compute the integral defining $\mathcal{J}$ if one treats $\frac{\partial}{\partial x}$ as a constant. This requires some assumptions about existence and convergence which we will not discuss here. Write (3.10.5) as

$$\partial_\tau C(x,v,\tau) = [\mathcal{D} + \mathcal{J}]C(x,v,\tau), \tag{3.10.7}$$

where $\tau = T - t$ and $\mathcal{D}$ is given by

$$\mathcal{D} \equiv -r + \left(r - q - \frac{1}{2}\sigma^2 v\right)\frac{\partial}{\partial x} + \kappa(\theta - v)\frac{\partial}{\partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2}{\partial x} + \xi\sigma\rho v^{\beta+1/2}\frac{\partial^2}{\partial x \partial v} + \frac{1}{2}\xi^2 v^{2\beta}\frac{\partial^2}{\partial v^2}. \tag{3.10.8}$$

The parabolic operator $\mathcal{D}$ is an infinitesimal generator of diffusion, while $\mathcal{J}$ is an infinitesimal generator of the jump process.

For jumps with finite variation and finite activity, the last two terms in $\mathcal{J}$ could be integrated out and added to $\mathcal{D}$. In the case of jumps with finite variation and infinite activity, the last term of $\mathcal{J}$ could be integrated out and added to $\mathcal{D}$. However, we do not do this because adding these two terms to $\mathcal{D}$ negatively influence the stability of the finite-difference scheme that we use to solve $\mathcal{D}C(x,v,t) = 0$.

We now use a splitting technique of Strang, along with a result from Thalhammer & Koch (2010). We apply the splitting technique to the jump and diffusion parts. The second step of the splitting method gives rise to the equation

$$\partial_\tau C(x,v,\tau) = \mathcal{J}C(x,v,\tau). \tag{3.10.9}$$

The formal solution of (3.10.9) is given by

$$C(x,v,\tau+\Delta\tau) = e^{\Delta\tau\mathcal{J}}C(x,v,\tau) \tag{3.10.10}$$

The problem is then reduced to computing the exponential, which is done numerically. We must build a grid $\mathbf{G}(x,v)$ to discretize the operator ∂_x (and, therefore, the operator $\mathcal{J}$). This discretization should be done such that:

- **G** is a superset of the PDE grid that is used to solve $\partial_\tau C(x,v,\tau) = \mathcal{D}C(x,v,\tau)$ in the first and

third steps of the splitting method. This avoids having to interpolate.

- The discretization provides the necessary order of approximation in space of the operator $\mathcal{J}$.
- The discretization provides unconditional stability of the solution (3.10.10).
- The discretization guarantees positivity of (3.10.10).

Example: Merton jumps We illustrate the splitting process for Merton jumps. Using (3.10.4) for $\nu(dy)$ in (3.10.6) gives

$$\mathcal{J} = \lambda \left(e^{\mu_J d_x + \frac{1}{2}\sigma_J^2 d_x^2} - m d_x - 1 \right),$$

where

$$m = e^{\mu_J + \frac{1}{2}\sigma_J^2 - 1}$$

and $d_x = \frac{\partial}{\partial x}$, $d_x^2 = \frac{\partial^2}{\partial x^2}$ were treated as constants in the integration.

Putting this into (3.10.10), the problem reduces to solving the evolutionary pure jump equation

$$C(x, v, \tau + \Delta\tau) = \mathcal{A} C(x, v, \tau), \tag{3.10.11}$$

where

$$\mathcal{A} \equiv \exp\left[\lambda \Delta\tau \left(e^{\mu_J d_x + \frac{1}{2}\sigma_J^2 d_x^2} - m d_x - 1 \right) \right].$$

A matrix exponential method for dealing with the double exponential operator in (3.10.11) is considered in Tangman (2011).

Numerical Methods for ULSVJ

The jump diffusion framework requires an accurate and fast computation algorithm for both model calibration and valuation. Because Monte-Carlo methods possess severe drawbacks in efficiency and performance, we use a finite difference method for calibration and valuation. The finite difference method has the following properties:

- The discretization is second order in both space and time variables.
- The mesh is adaptive to allow for a higher concentration of mesh points near strike and/or barrier levels.
- The N-dimensional PDE is reduced to N one-dimensional PDEs through operator splitting. This includes:
  - Orthogonalization of the second diffusion process
  - A pseudo-parabolic PDE approach for the jump process

USLV Calibration

The finite-difference-based model calibration has the following features:

- Calibration to FX barrier options, besides the vanilla implied-volatility surface. Currently, OT, NT and DNT barriers are allowed in the model calibration, as well as KO and KI options.

- The calibration target is the option price. Therefore, out-of-the-money vanilla options are preferred in model calibration.
- Users are able to choose either type of options (vanilla, exotic or both) for either part of the model (SV, LV, Jumps) in the LV, ISV and SV calibration.
- Users are allowed to use a local or global solver in model calibration. Global solvers can prevent the solver terminating at a local minimum.
- Each parameter in the model can be included or excluded in the calibration.
- A vector of models is created for the given set of options, one model per instrument. This allows a highly adaptive scheme to be constructed for pricing each particular instrument. However, this is not as universal as kernel pricing.
- The Heston and Bates model have a smart initial guess technique implemented.

Chapter 4

Commodity models

4.1 Black Model

The stochastic differential equation governing the risk-neutral evolution of the underlying commodity price $S(t)$ is

$$dS = [r(t) - \delta(t)]Sdt + \sigma(t)SdW(t). \tag{4.1.1}$$

The model depends on two parameters:

- the instantaneous volatility $\sigma(t)$
- the convenience yield $\delta(t)$.

The convenience yield parameter is adjusted internally to match the specified initial forward price curve. The drift term of Eq. (4.1.1) depends on the short rate $r(t)$ corresponding to the selected interest rate model.

Model (4.1.1) is a generalization of the Black model. The original Black model assumed constant instantaneous volatility and constant interest rate.

4.2 Gabillon CMDTY Model

The Gabillon model is a two-factor lognormal model in which the first short-term factor mean reverts to the second lower long-term factor. The Gabillon model dynamics are given by:

$$d\log S = \left(\beta(\log L - \log S) + \mu\right) dt + \sigma_s dW_s \tag{4.2.1}$$

$$d\log L = -\frac{1}{2}\sigma_\ell^2 dt + \sigma_\ell dW_\ell. \tag{4.2.2}$$

This model is Gaussian and may be solved directly by expressing the model as a vector of Ornstein-Uhlenbeck (OU) processes. Define

$$\ell = \log L, \qquad s = \log S,$$

$$\vec{s} = (s, \ell)^T, \qquad \boldsymbol{\sigma} = \text{diag}(\sigma_s, \sigma_\ell), \qquad \boldsymbol{\kappa} = \beta \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix},$$

$$d\vec{W} = (dW_s, dW_\ell)^T, \qquad \boldsymbol{\mu} = \left(\mu, -\frac{1}{2}\sigma_\ell^2\right)^T.$$

We may then rewrite (4.2.1) and (4.2.2) as

$$d\vec{s} = \boldsymbol{\kappa}\vec{s}dt + \boldsymbol{\mu}dt + \boldsymbol{\sigma}d\vec{W}. \tag{4.2.3}$$

The solution of (4.2.3) is given by

$$\vec{s}_t = e^{\boldsymbol{\kappa}t}\vec{s}_0 + \int_0^t e^{\boldsymbol{\kappa}(t-t')}\vec{\mu}_{t'}dt' + \int_0^t e^{\boldsymbol{\kappa}(t-t')}\boldsymbol{\sigma}_{t'}d\vec{W}_{t'}. \tag{4.2.4}$$

We may eliminate S and L by introducing the forward prices F, where $F_{t,T} = \mathbb{E}_t S_T = \mathbb{E}_t(e^{s_T})$. The processes in F are martingales and are described by

$$dF_{t,T} = F_{t,T}\left(B\sigma_s dW_s + (1-B)\sigma_\ell dW_\ell\right), \tag{4.2.5}$$

where

$$B = e^{-\beta(T-t)}.$$

We may rewrite (4.2.5) as

$$d\log F_{t,T} = B\sigma_s dW_s + (1-B)\sigma_\ell dW_\ell - \frac{1}{2}\left(B^2\sigma_s^2 + (1-B)^2\sigma_\ell^2 + 2B(1-B)\rho\sigma_s\sigma_\ell\right)dt, \tag{4.2.6}$$

where ρ is the correlation between the Brownian motions. The solution to (4.2.6) is found in a manner similar to (4.2.4) and is given by

$$F_{t,T} = F_{t_0,T}\exp\left[\int_0^t B\sigma_s dW_{s,t'}dt' + \int_0^t (1-B)\sigma_\ell dW_{\ell,t'} - \frac{1}{2}\int_0^t \left(B^2\sigma_s^2 + (1-B)^2\sigma_\ell^2 + 2B(1-B)\rho\sigma_s\sigma_\ell dt\right)\right]. \tag{4.2.7}$$

Here, the drift terms have all been absorbed into the definition of the forward curve, leaving only an exponential martingale containing Brownian motion terms and their compensator.

The calibration parameters of this model are:

σ_s: short-factor volatility

σ_ℓ: long-factor volatility

ρ: correlation

β: mean reversion

Any of these parameters can be made functions of time. Constant parameters lead to a term structure of volatilities that is stationary in time—the curve shape will be the same throughout the simulation of the price process, as expected for features of the term structure that are persistent. Nonconstant parameters lead to a volatility curve whose features "roll off" the forward curve as it evolves, as occurs when such features are tied to specific dates or seasons (such as winter/summer seasonality) or to an anticipated economic or political event.

The expression for the Gabillon implied volatility at maturity T is

$$\bar{\sigma}_T^2 = \left(\sigma_s^2 + \sigma_\ell^2 - 2\rho\sigma_s\sigma_\ell\right)\left(\frac{e^{-2\beta\delta T} - e^{-2\beta T}}{2\beta(T-\delta T)}\right) + 2\sigma_\ell(\rho\sigma_s - \rho_\ell)\left(\frac{e^{-\beta\delta T} - e^{-\beta T}}{\beta(T-\delta T)}\right) + \sigma_\ell^2, \tag{4.2.8}$$

where $\delta T > 0$ is the time difference between the option expiry and the underlying futures expiry.

Volatility term structures in the two-factor Gabillon model are monotone increasing if $\sigma_s < \sigma_\ell$ and monotone decreasing if $\sigma_s > \sigma_\ell$. In general, the short-factor volatility is approximately equal to the volatility of the option on the front month contract, whereas the long-term volatility is approximately equal to the volatility of the option on the back-month contract. The mean reversion and correlation control the slope of the volatility term structure, as well as the correlation between contracts.

On the first few future contracts, the influence of the long-factor volatility σ_ℓ and the correlation ρ is small, and therefore the implied vol given by (4.2.8) can be approximated by

$$\bar{\sigma}_T^2 \approx \sigma_s^2\left(\frac{e^{-2\beta\delta T} - e^{-2\beta T}}{2\beta(T-\delta T)}\right) \tag{4.2.9}$$

for the first few futures. Then given an estimate of β over the first few contracts, σ_s^2 can be estimated by the approximation in (4.2.9).

For very large T, (4.2.8) behaves like

$$\bar{\sigma} \approx \sigma_\ell^2 + \frac{1}{\beta(T-\delta T)}\left[\left(\sigma_s^2 + \sigma_\ell^2 - 2\rho\sigma_s\sigma_\ell\right)\frac{e^{-2\beta\delta T}}{2} + 2\sigma_\ell(\rho\sigma_s - \sigma_\ell)e^{-\beta\delta T}\right]. \tag{4.2.10}$$

The Gabillon model[1] is a model for a spot process that is cast into the form of a model for forward prices. Although it is a spot-type model, the Gabillon model has realistic forward curve dynamics and has therefore become widely used among major financial institutions.

4.3 Gabillon 3F CMDTY Model—Beta

The Gabillion three-factor model is a three-dimensional extension of the Gabillon model in which the short factor S mean reverts to a medium factor M, which in turn mean reverts to a long

[1] J. Gabillon, "The Term Structures of Oil Futures Prices", Oxford Institute for Energy Studies, 1991.

factor L. The Gabillon 3F model is defined by

$$\begin{aligned} d\log S &= \left(\beta_{sm}(\log M - \log S) + \mu_s\right) dt + \sigma_s dW_s \\ d\log M &= \left(\beta_{ml}(\log L - \log M) + \mu_m\right) dt + \sigma_m dW_m \\ d\log L &= -\frac{1}{2}\sigma_\ell^2 dt + \sigma_\ell dW_\ell \\ dW_s dW_m &= \rho_{sm} \\ dW_m dW_\ell &= \rho_{m\ell} \\ dW_s dW_\ell &= \rho_{s\ell}. \end{aligned}$$

As with the two-factor Gabillon model, we may eliminate S, M and L and write the SDEs directly for a few stochastic process, the forward prices F. These are martingales with dynamics given by

$$\frac{dF_{t,T}}{F_{t,T}} = B_{sm;T-t}(\sigma_s dW_s - \sigma_m dW_m) + \sigma_\ell dW_\ell + (\sigma_\ell dW_\ell - \sigma_m dW_m)\,\frac{\beta_{m\ell}B_{sm;T-t} - \beta_{sm}B_{m\ell;T-t}}{\beta_{m\ell} - \beta_{sm}}, \quad (4.3.1)$$

where we have set

$$\begin{aligned} s = \log S, \qquad m = \log M, \qquad \ell = \log L \\ B_{sm;t} = e^{-\beta_{sm}t}, \qquad B_{m\ell;t} = e^{-\beta_{m\ell}t}. \end{aligned}$$

The solution of (4.3.1) is given by

$$\begin{aligned} F_{t,T} = F_{t_0,T}\,\exp\Bigg[&\int_{t_0}^{t} B_{sm;T-t'}\sigma_s dW_{s,t'} - \int_{t_0}^{t} B_{sm;T-t'}\sigma_m dW_{m,t'} + \int_{t_0}^{t}\sigma_\ell dW_{\ell,t'} \\ &+ \int_{t_0}^{t}\frac{\beta_{m\ell}B_{sm;T-t'} - \beta_{sm}B_{m\ell;T-t'}}{\beta_{sm} - \beta_{m\ell}}\sigma_\ell dW_\ell, t' - \int_{t_0}^{t}\frac{\beta_{m\ell}B_{sm;T-t'} - \beta_{sm}B_{m\ell;T-t'}}{\beta_{sm} - \beta_{m\ell}}\sigma_m dW_{m,t'} \\ &-\frac{1}{2}\int_{t_0}^{t}\hat{s}^T e^{\kappa(T-t')}\mathrm{Cov}e^{\kappa(T-t')}\hat{s}dt'\Bigg], \end{aligned} \quad (4.3.2)$$

where κ is a direct generalization of $\bar{\kappa}$ in the 2F model:

$$\kappa = \beta_{sm}\kappa_{sm} + \beta_{m\ell}\kappa_{m\ell},$$

$$\kappa_{sm} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \kappa_{m\ell} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

Cov is the factor covariance matrix and $\hat{s} = (1,0,0)^T$.

As with the two-factor model, it is possible to calculate the implied volatilities as functions of model parameters and work out some limits for long and short maturity times. Taking $T_0 \sim T$,

$T \to \infty$, the volatility is approximated by

$$\bar{\sigma}^2 \approx \sigma_\ell^2 + \frac{1}{T}\left[\frac{\sigma_s^2 + \sigma_m^2 - 2\sigma_s\sigma_m\rho_{sm} + 4\rho_{s\ell}\sigma_s\sigma_\ell - 4\rho_{m\ell}\sigma_m\sigma_\ell}{2\beta + sm}\right.$$
$$+ \frac{1}{2}\left(\sigma_\ell^2 + \sigma_m^2 - 2\rho_{m\ell}\sigma_m\sigma_\ell\right)\frac{\beta_{sm}^2 + \beta_{m\ell}^2 + 3\beta_{sm}\beta_{m\ell}}{(\beta_{sm} + \beta_{m\ell}\beta_{sm}\beta_{m\ell}}$$
$$+ \frac{(\beta_{m\ell} - \beta_{sm})(\beta_{m\ell} + 2\beta_{sm})}{(\beta_{sm} + \beta_{m\ell})\beta_{sm})}\left(\sigma_s\sigma_\ell\rho_{s\ell} + \sigma_m^2 - \sigma_m\sigma_\ell\rho_{m\ell} - \sigma_s\sigma_m\rho_{sm}\right)$$
$$\left. +2\left(\sigma_\ell^2 - \sigma_m\sigma_\ell\rho_{m\ell}\right)\left(\frac{1}{\beta_{sm}} + \frac{1}{\beta_{m\ell}}\right)\right]. \quad (4.3.3)$$

For the short end of the curve ($T_0 \to t_0$), we obtain

$$\bar{\sigma}^2 \approx \sigma_s^2 + \mathcal{O}(T_0 - t_0). \quad (4.3.4)$$

Note that $\bar{\sigma}$ is defined as $\bar{\sigma} = \left(\frac{1}{T_0 - t_0}\int_{t_0}^{T_0}\sigma^2\right)^{1/2}$.

4.4 Schwartz One-Factor Model

The stochastic differential equation governing the risk-neutral evolution of the underlying commodity price $S(t)$ is

$$dS = (r(t) - \xi(t) - k(t)\ln S)Sdt + \sigma(t)SdW(t). \quad (4.4.1)$$

The model depends on three parameters:

- the mean-reversion rate $k(t)$
- the instantaneous volatility $\sigma(t)$
- the mean-reversion level parameter $\xi(t)$.

The mean-reversion level parameter is adjusted internally to match the specified initial forward price curve. The drift term of Eq. (4.4.1) depends on the short rate $r(t)$ corresponding to the selected interest rate model.

Model (4.4.1) is a generalization of the Schwartz one-factor model. The original Schwartz one-factor model assumes constant parameters and a constant interest rate.

4.5 Gibson Schwartz Two-Factor Model

The stochastic differential equations governing the risk-neutral evolution of the underlying commodity price $S(t)$ and stochastic convenience yield $\delta(t)$ are

$$dS = (r(t) - \delta)Sdt + \sigma_1(t)SdZ_1, \quad (4.5.1)$$
$$d\delta = [\xi(t) - k(t)\delta]dt + \sigma_2(t)dZ_2. \quad (4.5.2)$$

The model depends on five parameters:[2]

- the convenience yield mean reversion rate $k(t)$
- the convenience yield mean reversion level $\xi(t)$
- the commodity price instantaneous volatility $\sigma_1(t)$
- the convenience yield instantaneous volatility $\sigma_2(t)$
- the correlation coefficient ρ, defined by $\rho dt = dZ_1(t)dZ_2(t)$.

The mean reversion level $\xi(t)$ is adjusted internally to match the initial forward price curve. The drift term of Eq. (4.5.1) depends on the short rate $r(t)$, which corresponds to the selected interest-rate model.

The model (4.5.1)–(4.5.2) is a generalization of the Gibson-Schwartz model. The original Gibson-Schwartz model assumes constant parameters. The hybrid model specification for (4.5.1)–(4.5.2) utilizes the fact that model (4.5.1)–(4.5.2) could be interpreted as a combination of the Black-Scholes foreign exchange model and the Hull-White interest rate model.

4.6 Schwartz Three-Factor Model

This model is a combination of Gibson-Schwartz two-factor model (Section 4.5) with a Hull-White one-factor model (Section 1.1.2) for the domestic short rate.

4.7 Heston Model with Time-Dependent Parameters

The stochastic differential equations governing the risk-neutral evolution of the underlying commodity price $S(t)$ and variance $z(t)$ are

$$dS = [r(t) - \xi(t)]Sdt + \lambda(t)\sqrt{z}SdW_s(t), \tag{4.7.1}$$

$$dz = \theta(t)(1 - z)dt + \gamma(t)\sqrt{z}dW_z(t). \tag{4.7.2}$$

This model depends on five parameters:

- the instantaneous commodity price volatility $\lambda(t)$
- the commodity price level parameter $\xi(t)$
- the instantaneous volatility of volatility $\gamma(t)$
- the variance mean reversion rate $\theta(t)$
- the correlation coefficient $\rho(t)$, defined by $\rho(t)dt = dW_s(t)dW_z(t)$.

The commodity price level parameter is adjusted internally to match the specified initial forward price curve. The drift term of Eq. (4.7.1) depends on the short rate $r(t)$ that corresponds to the selected interest rate model.

The variance process in model (4.7.1)–(4.7.2) is similar to the Andersen-Andreasen formulation [2].

[2] A deterministic interest rate is assumed. The specification of the stochastic-interest-rate model may add the additional parameters (i.e., the corresponding correlation coefficients).

4.8 Mean-Reverting Stochastic Volatility Model

Mean Reverting Stochastic Volatility model is relevant for commodities such as precious metals or others for which storage is not an issue.

The stochastic differential equations governing the risk-neutral evolution of the underlying commodity price $S(t)$ and variance $z(t)$ are

$$dS = [r(t) - \xi(t) - k(t)lnS)]Sdt + \lambda(t)\sqrt{z}SdW_s(t), \tag{4.8.1}$$

$$dz = \theta(t)(1 - z)dt + \gamma(t)\sqrt{z}dW_z(t). \tag{4.8.2}$$

This model depends on six parameters:

- the instantaneous commodity price volatility $\lambda(t)$
- the commodity price mean reversion rate $k(t)$
- the commodity price mean reversion level $\xi(t)$
- the instantaneous volatility of volatility $\gamma(t)$
- the instantaneous variance mean reversion rate $\theta(t)$
- the correlation coefficient $\rho(t)$, defined by $\rho(t)dt = dW_s(t)dW_z(t)$.

The commodity price mean reversion level parameter is adjusted internally to match the specified initial forward price curve. The drift term in Eq. (4.8.1) depends on the short rate $r(t)$ that corresponds to the selected interest rate model.

Model (4.8.1)–(4.8.2) is a generalization of the Geman model [31]. The original model assumes constant parameters and a (implicitly) constant interest rate. The variance process in model (4.8.1)–(4.8.2) is similar to the Andersen-Andreasen formulation [2].

Chapter 5

Inflation Models

5.1 Generalized Jarrow-Yildirim Approach

We use the approach proposed by Jarrow and Yildirim [58] for modeling inflation and nominal rates. Their methodology is based on a foreign-currency analogy. In this analogy, real rates are viewed as interest rates in the real (i.e., foreign) economy, and the CPI is the exchange rate between the nominal (i.e., domestic) and real "currencies". In this setting, the valuation of an inflation-indexed payoff becomes equivalent to that of a cross-currency interest rate derivative.

The original JY model is formulated in terms of instantaneous forward rates. One can show that it can be reformulated in terms of short rates. Details can be found in Mercurio [48].

The result is as follows: In the nominal risk-neutral measure, the equations for the nominal rate, real rate and CPI coincide with the equations for the domestic rate, foreign rate and FX rate, respectively, in the cross-currency model implemented by Numerix.

The JY model assumes that both nominal and real (instantaneous) rates are normally distributed under their respective risk-neutral measures. That is, it assumes that the IR components in the cross-currency model are HW models. Our approach also allows for the BK models as IR components of the cross-currency model.

5.2 Inflation Market Models

This section discusses the inflation market models (IMMs) that are implemented in Numerix. Market models of inflation are similar to Libor market models (LMMs) of interest rates [22]. In the LMM approach, a forward Libor rate of a particular time period in the future is modeled as a separate lognormal process. Similarly, the market model of inflation deals with the modeling of forward prices of the consumer price index (CPI) [38, 19]. Modeling each CPI forward by a separate process allows fitting the quotes on zero-coupon inflation-indexed (ZCII) swaps and inflation-indexed options. Tuning the shape of volatility time dependence and correlations of different forwards provides the necessary freedom for fitting the prices of year-on-year inflation-indexed (YOYII) swaps as well as caps/floors.

Inflation derivatives that are typically used in the calibration of inflation models can be separated into two groups. The first group includes zero-coupon inflation-indexed swaps and inflation-indexed options. Instruments in this group are direct derivatives of the inflation index

such that their prices are not sensitive to the relative dynamics of CPI forwards. The second group of instruments includes year-on-year swaps and year-on-year caps/floors. This sector could have been better represented via a model dealing with the dynamics of CPI rates, i.e., the ratios of adjacent CPI forwards. Such approach was suggested in [39]. The disadvantage of this approach is that the pricing of instruments of the first sector becomes inaccessible even within Monte Carlo (MC) techniques. An intervening compromise in between index and rate modeling was suggested in [49]. There one still deals with modeling of CPI indices, but volatility is defined in terms of volatility of forward CPI rates. This allows for a robust calibration to the instruments of the second market sector, while keeping the pricing of instruments of the first sector still available. Numerix implementations provide the flexibility in referencing the volatility to either indices or rates.

With the market of inflation derivatives becoming more and more liquid, one faces the problem of simulation calibration to sets of derivatives that differ in their strike only. In order to fit the market smile, similar to the LMM approach, one incorporates stochastic volatility into the model. At present, Numerix models include CIR and SABR stochastic processes. Following original publications (see [49, 50]), Numerix provides a Heston model for index-based volatility, and a SABR model for rate-based volatility.

The inflation market model can be used in a standalone manner or as a component of a hybrid setup along with an LMM component. In the former case, the interest rates are assumed to be deterministic. In the latter case, the dynamics of interest rate is governed by the LMM model.

The outline of the rest of the section is:

- Inflation instruments
- Role of LMM
- Inflation market model
- Hybrid IMMs
- Analytic pricing of inflation derivatives
- Stochastic volatility
- IMM calibration
- IMM SABR model beyond frozen drift approximation

Some technical details are relegated to appendices.

5.2.1 Inflation Instruments

In this section, we consider the inflation instruments that are often used in IMM calibration. Calibration procedures will be described in Section 5.2.7.

ZCII Swap

In a ZCII swap, one counterparty agrees to exchange a single, inflation-driven payment for a fixed payment at a certain time in the future. Both payments are made at the contract maturity, T_M. The floating payment is

$$N\left(\frac{I(T_M)}{I_0}-1\right), \tag{5.2.1}$$

where N is the notional, and I_0 is initial index value. The fixed payment is

$$N((1+K)^M-1), \tag{5.2.2}$$

where K is the strike and $M=T_M-T_0$ is the number of years spanned by the swap. A ZCII swap is quoted in the value of strike that makes the contract costless at the origin. Strikes are usually given in percents. Market values of swaps are supposed to be stripped such that the model gets a known inflation forward curve.

ZCII Option

A ZCII option pays the amount

$$N\left[\omega\left(\frac{I(T_M)}{I_0}-(1+K)^M\right)\right]^+ \tag{5.2.3}$$

at maturity T_M. Here N is the notional, I_0 is the initial index value, K is the strike, $M=T_M-T_0$ and $\omega=1,-1$ for calls and puts, respectively. Options are usually quoted in bp of the notional, and the strike is usually given as a percentage.

Year-on-Year Inflation-Indexed Swap

The year-on-year inflation-indexed (YOYII) swap exchanges a stream of inflation driven, floating, periodic payments per a similar stream of fixed payments. Payments are made at the ends of each swap interval. The amount of fixed payment at the end of the interval $(\mathcal{T}_n,\mathcal{T}_{n+1})$ is

$$N\phi_n K, \tag{5.2.4}$$

where N is the notional, ϕ_n is the fixed-leg year fraction of $(\mathcal{T}_n,\mathcal{T}_{n+1})$ and K is the strike. The corresponding floating payment is

$$N\psi_n\left(\frac{I(\mathcal{T}_{n+1})}{I(\mathcal{T}_n)}-1\right), \tag{5.2.5}$$

where ψ_n is the floating-leg year fraction of the interval $(\mathcal{T}_n,\mathcal{T}_{n+1})$. Payments are usually annual, so $\phi_n,\psi_n\approx 1$. Swaps are quoted in par values of the spreads, κ, in percents.

YOYII Caps/Floors

YOYII cap/floors consist of caplets/floorlets, each representing an option that pays

$$N\left(\omega\left(\frac{I(\mathcal{T}_n)}{I(\mathcal{T}_{n-1})}-1\right)\psi_n-\omega K\phi_n\right)^+ \tag{5.2.6}$$

at $\mathcal{T}_n$, where N is the notional, K is the strike and $\omega=1,-1$ for caplets and floorlets, respectively. Caps/floors are quoted in bp of the notional.

5.2.2 Role of LMM

In the IMM approach, it is standard to model the interest rates via LMM. A description of LMM is found in Section 1.2.

5.2.3 Inflation Market Model

The inflation market model sets the dynamics of CPI forward rates, which are given by

$$F_i(t)=\mathbb{E}_{\mathcal{T}_i}[I(\mathcal{T}_i),\mathcal{F}_t], \tag{5.2.7}$$

where $I(t)$ is the spot price of CPI index and the expectation is taken under the $\mathcal{T}_i$-forward measure (see [22], [38], [19]). The set of forward maturities is

$$0<\mathcal{T}_0<\mathcal{T}_1<\cdots<\mathcal{T}_{N_I}. \tag{5.2.8}$$

In general, forward maturities differ from Libor maturities. The model assumes that the dynamics of CPI forwards are lognormal:

$$dF_n(t)=F_n(t)\,\lambda_n^I(t)\cdot(\rho\,\tilde{\mu}_n(t)dt+dW^I(t)). \tag{5.2.9}$$

where $W^I(t)=\{W_0^I(t),\cdots,W_{F_I-1}^I(t)\}$ is a F_I-factor Brownian motion, λ_n^I is the volatility and $\tilde{\mu}_n$ is a drift term. The components of the vector $W^I(t)$, as in the case of the LMM model for interest rates, are assumed to be mutually independent. However, the Brownian motions of two sectors are correlated by a matrix ρ:

$$dW_f^I(t)\,dW_g^L(t)=\rho_{fg}\,dt, \tag{5.2.10}$$

where W^L is the Brownian motion from the LMM interest rate model (see Section 1.2). The correlation matrix ρ is assumed to be time independent. If CPI maturity dates belong to the set of Libor maturities, then the function $\tilde{\mu}_n(t)$ coincides with $\mu_{\eta(T_n)}(t)$ from the LMM model. To handle the general case, we interpolate linearly in between the adjacent Libor rates:

$$\tilde{\mu}_n(t)=\frac{T_{\eta(\mathcal{T}_n)}-\mathcal{T}_n}{T_{\eta(\mathcal{T}_n)}-T_{\eta(\mathcal{T}_n)-1}}\,\mu_{\eta(\mathcal{T}_n)-2}(t)+\frac{\mathcal{T}_n-T_{\eta(\mathcal{T}_n)-1}}{T_{\eta(\mathcal{T}_n)}-T_{\eta(\mathcal{T}_n)-1}}\,\mu_{\eta(\mathcal{T}_n)-1}(t). \tag{5.2.11}$$

Note that the measure has to be fixed at the time of Libor model formulation; i.e. the measure is set on Libor maturity dates. For analytic calculations, it is sometimes convenient to turn to the terminal measure associated with one of the index maturities. In this case, one may need

to implement an additional interpolation for switching to the index forward measure if the latter is off the Libor maturity dates.

We assume the volatility to have a form similar to that of IR components,

$$\lambda^I_{n,f}(t) = \gamma^I_n(t)\,\mathcal{N}^I_{n,f}(t), \tag{5.2.12}$$

where each vector $\mathcal{N}^I_n(t)$ is normalized to unity: $\sum_f \mathcal{N}^I_{nf}(t)\mathcal{N}^I_{nf}(t) = 1$. This assumes that the correlation matrix of forwards is

$$C^I_{nm}(t) = \mathcal{N}^I_n(t) \cdot \mathcal{N}^I_m(t). \tag{5.2.13}$$

Index-Based vs. Rate-Based Volatility

Model (5.2.9) defines the dynamics of index forwards through the index volatilities. This form is most convenient in terms of calibrating the model to the CPI sector of the market. (Simultaneous calibration to both sectors of the market is still possible in this approach.) Alternatively, one can primarily target the calibration to the market of YOYII swaps/caps/floors. In this case, it is more convenient to formulate dynamics as

$$dF_n(t) = F_n(t) \sum_{m=\tilde\eta(t)}^{n} \lambda^R_m(t) \cdot (\rho\,\tilde\mu_n(t)\,dt + dW^I(t)), \tag{5.2.14}$$

where $\tilde\eta(t)$ gives the index of first forward maturity strictly larger than t. In this formulation, $\lambda^R_m(t)$ can be interpreted as CPI rate volatility. Indeed, defining the nth CPI rate as

$$\mathcal{R}_n(t) = \frac{F_n(t)}{F_{n-1}(t)} \tag{5.2.15}$$

and using Ito's lemma, we get

$$d\mathcal{R}_n(t) = \mathcal{R}_n(t)\,(r_n(t)\,dt + \lambda^R_n(t) \cdot dW^I), \tag{5.2.16}$$

$$r_n(t) = \rho\,\tilde\mu_n(t) \cdot \sum_{m=\tilde\eta(t)}^{n} \lambda^R_m(t) - \rho\,\tilde\mu_{n-1}(t) \cdot \sum_{m=\tilde\eta(t)}^{n-1} \lambda^R_m(t) - \lambda^R_n(t) \cdot \sum_{m=\tilde\eta(t)}^{n-1} \lambda^R_m(t). \tag{5.2.17}$$

Comparing the models (5.2.9) and (5.2.14), we obtain that rate and index volatilities are related by

$$\lambda^I_n(t) = \sum_{m=\tilde\eta(t)}^{n} \lambda^R_m(t). \tag{5.2.18}$$

The matrix of CPI rate volatilities $\lambda^R_{n,f}(t)$ is parameterized similar to $\lambda^I_{n,f}(t)$:

$$\lambda^R_{n,f}(t) = \gamma^R_n(t)\,\mathcal{N}^R_{n,f}(t), \tag{5.2.19}$$

where each vector $\mathcal{N}^R_n(t)$ is normalized. The corresponding correlation matrix is

$$C^R_{nm}(t) = \mathcal{N}^R_n(t) \cdot \mathcal{N}^R_m(t). \tag{5.2.20}$$

5.2.4 Hybrid IMMs

It is common in the literature to consider use the LMM to model the interest rate component of the IMM. In many cases, however, using the LMM may lead to unnecessary setup complications and extra computational burden. Moreover, the analytic calculations in the case of the LMM cannot be done without making some sort of approximations, like freezing drift terms (see below). In this respect, it makes sense to consider also a more simple IR model which would allow for full analytic study. The natural candidate is the HW model. Thus, we consider two models of interest rates, HW and LMM, beginning with the former.

Hybrid IMM+HW

The Hull-White model is used for the short-rate interest rate. We consider the symmetric multi-factor version of this model where the short interest rate is given by the sum of stochastic processes $\{s_i\}$:

$$r(t) = \theta(t) + \sum_{i=0}^{F_{\text{IR}}} s_i(t). \tag{5.2.21}$$

Each process s_i is driven by the SDE

$$ds_i(t) = -a_i(t)\, s_i(t)\, dt + \sigma_i(t)\, dW_i^{\text{IR}}(t), \tag{5.2.22}$$

where $a_i(t)$ and $\sigma_i(t)$ are the reversion and volatility of ith process, respectively. The function $\theta(t)$ in Eq. (5.2.21) is chosen so that the yield curve is matched. The Brownian motions that drive the ith and jth components of $s(t)$ are correlated by

$$dW_i^{\text{IR}}(t)\, dW_j^{\text{IR}}(t) = \rho_{ij}^{\text{IR}}\, dt. \tag{5.2.23}$$

The CPI and IR sectors can be correlated as well:

$$dW_f(t)\, dW_g^{\text{IR}}(t) = \rho_{fg}\, dt. \tag{5.2.24}$$

The function $\tilde{\mu}_n(t)$ in Eq. (5.2.14) is calculated in Appendix D of Section 5.2.9 for the case of the HW model. In the risk-neutral measure, the result is

$$\tilde{\mu}_n^i(t) = \sum_{j=0}^{F_{\text{IR}}} \sigma_j(t) \int_t^{\mathcal{T}_n} e^{-\int_t^{T'} a_j(t')dt'}\, dT' \tag{5.2.25}$$

and in the $\mathcal{T}_n$-forward measure, the result is

$$\tilde{\mu}_m^i(t) = -\sum_{j=0}^{F_{\text{IR}}} \sigma_j(t) \int_{\mathcal{T}_m}^{\mathcal{T}_n} e^{-\int_t^{T'} a_j(t')dt'}\, dT'. \tag{5.2.26}$$

Hybrid IMM+LMM

The LMM defines dynamics of forward Libor rates

$$L_n(t) = \frac{1}{\delta_n}\left(\frac{P(t,T_n)}{P(t,T_{n+1})} - 1\right),$$

where $P(t,T_n)$ is the value of a zero-coupon bond that matures at T_n and $\delta_n = T_{n+1} - T_n$. The set of zero-coupon bond maturities is

$$0 < T_0 < T_1 < \cdots < T_{N^L}. \tag{5.2.27}$$

In the simplest case, the forward Libor rates satisfy the stochastic PDE

$$dL_n(t) = L_n(t)\,\lambda_n^L(t)\cdot(\mu_n^L(t)dt + dW^L(t)), \tag{5.2.28}$$

where $W^L(t) = \{W_0^L(t),\cdots,W_{F^L-1}^L(t)\}$ is an F^L-factor Brownian vector with independent components. The Brownian motions for CPI and IR are correlated by

$$dW_f(t)\,dW_g^L(t) = \rho_{fg}\,dt. \tag{5.2.29}$$

The correlation matrix ρ is assumed to be time independent. The volatility matrix can often be written in the form

$$\lambda_{n,f}^L(t) = \gamma_n^L(t)\,\mathcal{N}_{n,f}^L(t), \tag{5.2.30}$$

where each vector $\mathcal{N}_n^L(t)$ is normalized to unity: $\sum_f(\mathcal{N}_{n,f}^L)^2 = 1$. This form assumes the correlation matrix

$$C_{nm}^L(t) = \mathcal{N}_n^L(t)\cdot\mathcal{N}_m^L(t). \tag{5.2.31}$$

between different Libor rates. The function μ_n^L in Eq. (5.2.28) at a fixed Libor index is a vector in factor space. It adds the drift to the Brownian motion. Its particular form is specific to the measure choice. Below we specify the function μ_n^L in the terminal measure and the running spot measure.

Terminal measure In the terminal measure, which is associated with numerarie

$$N(t) = P(t,T_N)/P(0,T_N),$$

the function μ_n^L takes the form

$$\mu_n^L(t) = -\sum_{j=n+1}^{N-1}\frac{\delta_j\,L_j(t)}{1+\delta_j\,L_j(t)}\,\lambda_j^L(t). \tag{5.2.32}$$

Rolling spot measure The rolling spot measure is associated with the numeraire

$$N(t) = \frac{1}{P(T_0,T_1)}\cdots\frac{1}{P(T_{i-1},T_i)}\,\frac{P(t,T_{i+1})}{P(T_i,T_{i+1})},\quad T_n \le t \le T_{n+1}. \tag{5.2.33}$$

In this case, the function μ_n^L takes the form

$$\mu_n^L(t) = \sum_{j=\eta(t)}^{n}\frac{\delta_j\,L_j(t)}{1+\delta_j\,L_j(t)}\,\lambda_j^L(t), \tag{5.2.34}$$

where $\eta(t) = m+1$ for $T_m < t \le T_{m+1}$.

If the CPI maturity dates belonged to the set of Libor maturities, then the function $\mu_n(t)$

would just coincide with $\rho\,\mu_{\eta(T_n)}(t)$. To handle the general case, we make a linear interpolation in between the adjacent Libor rates:

$$\mu_n(t) = \frac{T_{\eta(\mathcal{T}_n)} - \mathcal{T}_n}{T_{\eta(\mathcal{T}_n)} - T_{\eta(\mathcal{T}_n)-1}}\,\rho\,\mu^L_{\eta(\mathcal{T}_n)-2}(t) + \frac{\mathcal{T}_n - T_{\eta(\mathcal{T}_n)-1}}{T_{\eta(\mathcal{T}_n)} - T_{\eta(\mathcal{T}_n)-1}}\,\rho\,\mu^L_{\eta(\mathcal{T}_n)-1}(t). \tag{5.2.35}$$

Note that the measure has to be fixed at the time of Libor model formulation; i.e., the measure is set on Libor maturity dates. For analytic calculations, it is sometimes convenient to turn to the terminal measure associated with one of index maturities. In this case, one may need to implement an additional interpolation for turning into that index forward measure if the latter is off the Libor maturity dates.

5.2.5 Analytic Pricing of Inflation Derivatives

It is most convenient to perform analytic pricing of inflation derivatives in the forward terminal measure associated with instrument maturity. In this measure, the drift $\tilde{\mu}_n(t)$ becomes zero. This allows for a straightforward calculation of index options. Indeed, the value of a call/put option on an underlying X that has lognormal dynamics under T-forward measure is

$$\begin{aligned}\mathbb{E}[D(T)(X(T)-\tilde{K})^+] &= B(0,T)\,\mathbb{E}_T[(X-\tilde{K})^+] \\ &= \omega\,B(0,T)\left(\mathbb{E}_T[X(T)]\,\Phi(\omega\,d_+) - \tilde{K}\Phi(\omega\,d_-)\right).\end{aligned} \tag{5.2.36}$$

Here $D(t)$ is the IR discount factor, $\mathbb{E}_T$ is the expectation under $T-$forward measure, $B(0,T)$ is the value of IR zero-coupon bond that matures at T, $\Phi(x)$ is the cumulative normal distribution function, $\omega = 1, -1$ for call and put options, respectively, and $d_\pm$ are given by

$$d_\pm = \frac{\ln(\mathbb{E}_T[X(T)]/\tilde{K}) \pm V_T[X(T)]/2}{\sqrt{V_T[X(T)]}}, \tag{5.2.37}$$

where $V_T[X(T)]$ is the variance of the underlying $X(T)$ calculated under T-forward measure.

Analytic pricing of swaps/caps/floors under IMM is difficult to achieve because of the presence of the drift function $\tilde{\mu}$, which is a complicated nonlinear function of Libor rates. To make the analytical approach feasible, it is customary to approximate the dynamics of forwards by freezing the Libor rates in the drift function at their values at the origin. We note that frozen drift approximations become exact in the case of deterministic interest rates, i.e, in the case of standalone IMM. Below we calculate the values of inflation index derivatives that were introduced in Section 5.2.1.

ZCII Option

The payment of ZCII options that mature at $\mathcal{T}_n$ (see Eq. (5.2.3)) can be put in the form of Eq. (5.2.36) by setting

$$X(\mathcal{T}_n) = N\frac{I(\mathcal{T}_n)}{I_0} = N\frac{F_n(\mathcal{T}_n)}{I_0}, \qquad \tilde{K} = (1+K)^M. \tag{5.2.38}$$

Since $F_n(t)$ is a martingale under the forward measure associated with $\mathcal{T}_n$, we have

$$\mathbb{E}_{\mathcal{T}_n}[X(\mathcal{T}_n)] = N\frac{F_n(0)}{I_0}. \tag{5.2.39}$$

Index-based model Since the drift $\tilde{\mu}_n(t)$ in Eq. (5.2.9) is zero in the $\mathcal{T}_n-$forward measure, we have

$$V_{\mathcal{T}_n}[X(\mathcal{T}_n)] = \int_0^{\mathcal{T}_n} \lambda_n^I(t)\cdot\lambda_n^I(t)dt \tag{5.2.40}$$

for the variance.

Rate-based model In the case of model (5.2.14), equations (5.2.39) and (5.2.40) are applicable as well. To calculate the variance (5.2.40), one has to express the index volatility $\lambda_n^I(t)$ through the rate volatility $\lambda_n^R(t)$ with the help of Eq. (5.2.18).

YOYII Swap

The YOYII swap is defined in Section 5.2.1. To find the value of YOYII swap, one has to calculate expected value of the ratio of two indices. Turning to the terminal measure associated with maturity time $\mathcal{T}_n$, we have to find

$$\mathbb{E}_{\mathcal{T}_n}\left[\frac{I(\mathcal{T}_n)}{I(\mathcal{T}_{n-1})}\right] = \mathbb{E}_{\mathcal{T}_n}\left[\frac{F_n(\mathcal{T}_n)}{F_{n-1}(\mathcal{T}_{n-1})}\right]. \tag{5.2.41}$$

Index-based model In the case of model (5.2.9), using Eq. (5.2.172) in Section 5.2.9 gives

$$\mathbb{E}_{\mathcal{T}_n}\left[\frac{I(\mathcal{T}_n)}{I(\mathcal{T}_{n-1})}\right] = e^{-\rho\int_0^{\mathcal{T}_{n-1}}\lambda_{n-1}^I(t)\cdot\tilde{\mu}_{n-1}(t)dt + V_{n-1,n-1}(\mathcal{T}_{n-1}) - V_{n,n-1}(\mathcal{T}_{n-1})}, \tag{5.2.42}$$

where

$$V_{nm}(t) = \int_0^t \lambda_n^I(t')\cdot\lambda_m^I(t')dt'. \tag{5.2.43}$$

Recall that $\tilde{\mu}_{n-1}(t)$ here is defined in $\mathcal{T}_n-$forward measure.

Rate-based model In the case of model (5.2.14), the expected index ratio can be found directly from Eqs. (5.2.16) and (5.2.17):

$$\mathbb{E}_{\mathcal{T}_n}\left[\frac{I(\mathcal{T}_n)}{I(\mathcal{T}_{n-1})}\middle|\mathcal{F}_t\right] = e^{-\rho\int_0^{\mathcal{T}_{n-1}}\lambda_{n-1}^I(t)\cdot\tilde{\mu}_{n-1}(t)dt - \int_0^{\mathcal{T}_{n-1}}\lambda_n^R(t)\cdot\lambda_{n-1}^I(t)dt}, \tag{5.2.44}$$

where λ_n^I is defined in Eq. (5.2.18). The two expressions in (5.2.42) and (5.2.44), agree due to Eq. (5.2.18).

YOYII Caps/Floors

YOYII caps/floors are defined in Section 5.2.1. A cap/floor is composed of caplets/floorlets. To find the value of a cap/floor, it is sufficient to calculate the values of caplets/floorlets that

compose the cap/floor. Payoff of a caplet/floorlet, according to Eq. (5.2.5), can be put to the form of Eq. (5.2.36) by setting

$$X(\mathcal{T}_n) = N\psi_n \frac{I(\mathcal{T}_n)}{I(\mathcal{T}_{n-1})}, \tag{5.2.45}$$

$$\tilde{K} = N(K\phi_n + \psi_n). \tag{5.2.46}$$

The expected value of $X(\mathcal{T}_n)$ is

$$\mathsf{E}_{\mathcal{T}_n}[X(\mathcal{T}_n)] = N\psi_n \mathsf{E}_{\mathcal{T}_n}\left[\frac{I(\mathcal{T}_n)}{I(\mathcal{T}_{n-1})}\right]. \tag{5.2.47}$$

The expected ratio of indices was calculated in the previous section and is given by Eqs. (5.2.42) and (5.2.44).

Index-based model The variance, according to Eq. (5.2.171) in Section 5.2.9 is

$$V_{\mathcal{T}_n} = V_{nn}(\mathcal{T}_n) + V_{n-1,n-1}(\mathcal{T}_{n-1}) - 2V_{n,n-1}(\mathcal{T}_{n-1}). \tag{5.2.48}$$

Rate-based model In the case of model (5.2.14), the variance of the nth forward index, according to Eq. (5.2.16), is given by

$$V_{\mathcal{T}_n} = \int_0^{\mathcal{T}_n} \lambda_n^R(t) \cdot \lambda_n^R(t) dt. \tag{5.2.49}$$

5.2.6 Stochastic Volatility

In this section, we present IMMs with stochastic volatility. Stochastic volatility is usually introduced by incorporating stochastic dynamics into the original deterministic volatility. Dynamics of the stochastic volatility are driven by a specific model. The underlying deterministic volatility can refer to either index or rate volatility. We consider two stochastic volatility models. The first model incorporates CIR dynamics into the forward index volatility [49]. The second model [50] incorporates SABR dynamics into the volatility of rates.

Heston Model

Let us first assume that the interest rates are deterministic. Following the Heston approach, we introduce a new stochastic variable, z, into Eq.(5.2.9):

$$dF_n(t) = F_n(t)\sqrt{z(t)}\,\lambda_n^I(t) \cdot dW^I(t). \tag{5.2.50}$$

The dynamics of z are governed by

$$dz(t) = a(t)(1 - z(t))\,dt + \sqrt{z(t)}\gamma(t)\,dV(t). \tag{5.2.51}$$

The Brownian noise V, in general, is correlated with W^I. Since V is a scalar, these correlations are represented by a vector:

$$\langle dV(t)\,dW_k^I(t)\rangle = \rho_k^{\mathsf{iv}}\,dt. \tag{5.2.52}$$

It is convenient to write the stochastic dynamics for the logarithm of forward CPI,

$$y_n(t) = \ln\left(\frac{F_n(t)}{F_n(0)}\right). \tag{5.2.53}$$

Equation (5.2.50) then becomes

$$dy_n(t) = \sqrt{z(t)}\,\lambda_n^I(t)\cdot dW^I(t) - \frac{1}{2}z(t)\,\lambda_n^I(t)\cdot\lambda_n^I(t)\,dt. \tag{5.2.54}$$

Evaluation of ZCII Options

The payout of the ZCII option that matures at $\mathcal{T}_n$ is given by Eq. (5.2.3). Turning to the $\mathcal{T}_n$-forward measure, we obtain

$$N\mathbb{E}\left[D(T)\left(\frac{I(\mathcal{T}_n)}{I_0}-\tilde{K}\right)^{\pm}\right] = N\,B(0,T)\,\mathbb{E}_{\mathcal{T}_n}[(X_0e^{y_n}-\tilde{K})^{\pm}] \tag{5.2.55}$$

for the option value , where $\tilde{K}=(1+K)^M$ and $X_0 = F_n(0)/I_0$. The expected value $\mathbb{E}_{\mathcal{T}_n}[(X_0e^{y}-\tilde{K})^{\pm}]$ in the above equation can be fully expressed in terms of a Fourier image of the characteristic function ϕ_n (see Section 5.2.9):

$$\mathbb{E}_{\mathcal{T}_n}[(X_0e^{y}-\tilde{K})^{\pm}] = \left\{\begin{array}{c} S_0\phi(1) \\ \tilde{K}\end{array}\right\} - \frac{\tilde{K}}{2} - \frac{\tilde{K}}{2\pi}\Re\int\frac{1}{\omega(\omega+i)}e^{-i\omega\ln(\tilde{K}/S_0)}\phi(i\omega)d\omega. \tag{5.2.56}$$

The characteristic function is defined as

$$\phi_n(\xi) = \mathbb{E}_{\mathcal{T}_n}\left[e^{\xi y(\mathcal{T}_n)}\right]. \tag{5.2.57}$$

To find this function, we generalize its definition by imposing conditioning at some earlier time t:

$$\phi_n(y,z,t;\xi) = \mathbb{E}_{\mathcal{T}_n}[\,e^{\xi y_n(\mathcal{T}_n)}\,|\,y_n(t)=y, z(t)=z\,]. \tag{5.2.58}$$

We write the backward PDE equation

$$\left(\frac{\partial}{\partial t} + a(1-z)\frac{\partial}{\partial z} - \frac{zD_1}{2}\frac{\partial}{\partial y} + \frac{z\gamma^2}{2}\frac{\partial^2}{\partial z^2} + \frac{zD_2}{2}\frac{\partial^2}{\partial y^2} + z\gamma M\frac{\partial^2}{\partial z\,\partial y}\right)\phi_n(y,z,t;\xi) = 0, \tag{5.2.59}$$

where

$$D_1 = D_2 = \lambda_n^I\cdot\lambda_n^I, \tag{5.2.60}$$

$$M = \lambda_n^I\cdot\rho^{\mathsf{iv}}, \tag{5.2.61}$$

and time dependence of coefficients is dropped for the sake of brevity. We consider a more general case where $D_1\neq D_2$ than is now needed for a reason that will become clear below.

Equation (5.2.59) can be solved with the help of the affine ansatz

$$\phi_n(y,z,t;\xi) = e^{A(t)+B(t)z+C(t)y}. \tag{5.2.62}$$

Inserting this equation into Eq. (5.2.59), we obtain

$$C = \xi, \tag{5.2.63}$$

$$A' = -aB, \tag{5.2.64}$$

$$B' = -\frac{\gamma^2}{2}B^2 + (a - \xi\,\gamma\,M)\,B + \frac{1}{2}(D_1\xi - D_2\xi^2), \tag{5.2.65}$$

where prime denotes derivatives with respect to time. The analytic solution of these equations in the case of piecewise-constant parameters is presented in Appendix C in Section 5.2.9.

Evaluation of an YOY Caplet/Floorlet

The payoff of the caplet/floorlet is given by Eq. (5.2.78). To evaluate its payoff, we turn to the forward measure associated with its maturity:

$$N\mathbb{E}\left[\left(\left(\frac{I(\mathcal{T}_n)}{I(\mathcal{T}_{n-1})} - 1\right)\psi_n - K\phi_n\right)^{\pm}\right] = N\,B(0,T)\,\mathbb{E}_{\mathcal{T}_n}[(X_0 e^{y_n - y_{n-1}} - \tilde{K})^{\pm}], \tag{5.2.66}$$

where $X_0 = F_n(0)/F_{n-1}(0)$, and $\tilde{K} = K\phi_n - \psi_n$. As in the case of index option, we use representation (5.2.56) with the characteristic function defined as

$$\phi_{nm}(\xi) = \mathbb{E}_{\mathcal{T}_n}[\,e^{\xi(y_n(\mathcal{T}_n) - y_m(\mathcal{T}_m))}\,]. \tag{5.2.67}$$

The characteristic function can be calculated with a similar technique as in the ZCII option case. First, assuming that $\mathcal{T}_n > \mathcal{T}_m$, we write Eq. (5.2.67) as

$$\phi_{nm}(\xi) = \mathbb{E}_{\mathcal{T}_n}\left[e^{-\xi\,y_m(\mathcal{T}_m)}\mathbb{E}_{\mathcal{T}_n}\left[e^{\xi\,y_n(\mathcal{T}_n)}\,|\,\mathcal{F}_{\mathcal{T}_m}\right]\right]. \tag{5.2.68}$$

The inner expectation can be calculated as in the previous case, giving

$$\phi_{nm}(\xi) = e^{A(\mathcal{T}_m)}\mathbb{E}_{\mathcal{T}_n}\left[e^{\xi(y_n(\mathcal{T}_m) - y_m(\mathcal{T}_m)) + B(\mathcal{T}_m)z}\right]. \tag{5.2.69}$$

To proceed further, one needs to construct the characteristic function of $y_{nm}(\mathcal{T}) = y_n(\mathcal{T}) - y_m(\mathcal{T})$ that obeys

$$dy_{nm}(t) = \sqrt{z(t)}\,(\lambda^I_n(t) - \lambda^I_m(t))\cdot dW^I(t) - \frac{1}{2}z(t)\,\lambda^I_n(t)\cdot\lambda^I_n(t)\,dt + \frac{1}{2}z(t)\,\lambda^I_m(t)\cdot\lambda^I_m(t)\,dt. \tag{5.2.70}$$

Deriving the backward PDE for Eqs. (5.2.51) and (5.2.70), we obtain Eq. (5.2.59) with the coefficients

$$D_1 = \lambda^I_n(t)\cdot\lambda^I_n(t) - \lambda^I_m(t)\cdot\lambda^I_m(t), \tag{5.2.71}$$

$$D_2 = (\lambda^I_n(t) - \lambda^I_m(t))\cdot(\lambda^I_n(t) - \lambda^I_m(t)), \tag{5.2.72}$$

$$M = (\lambda^I_n - \lambda^I_m)\cdot\rho^{\text{iv}}. \tag{5.2.73}$$

SABR $\beta = 1$ Model

In the case of the Heston model, the stochastic volatility was introduced as a multiplier of index volatilities. This is equivalent to making the *index* volatilities stochastic. In the case of the SABR model, we follow [50] and impose stochastic dynamics onto *rate* volatilities:

$$d\gamma_m^R(t) = \nu_m \, \gamma_m^R(t) \, dV(t). \tag{5.2.74}$$

The Brownian noise V is correlated with the index Brownian motions as defined by Eq. (5.2.52). In [50], each volatility process is driven by its own Brownian noise. We reduce the number of stochastic volatility factors to a minimum in order to make the MC simulations feasible. In terms of the model considered by Mercurio and Moreni, model (5.2.74) assumes perfect correlations of the Brownian noises that govern the stochastic dynamics of the rate volatilities. We note that it is hardly possible to get the correlation matrix of stochastic volatility factors by calibrating the model to sets of standard instruments. Thus, on the practical side, the performed minimization of factors simplifies the setup and essentially does not limit the model functionality.

According to Eq. (5.2.16), in the terminal measure associated with the maturity of the nth rate, the rate dynamics are given by

$$d\mathcal{R}_n(t) = \mathcal{R}_n(t) \left(r_n(t)dt + \gamma_n^R(t) dW_n^R \right), \tag{5.2.75}$$

where the Brownian noise dW_n^R is

$$dW_n^R = N_n^R \cdot dW^I \tag{5.2.76}$$

and the drift term is

$$r_n(t) = -\rho \, \tilde{\mu}_{n-1}(t) \cdot \sum_{m=\tilde{\eta}(t)}^{n-1} \mathcal{N}_m^R \, \gamma_m^R(t) - \gamma_n^R(t) \sum_{m=\tilde{\eta}(t)}^{n-1} C_{nm}^R \, \gamma_m^R(t). \tag{5.2.77}$$

Evaluation of YOY Caplet/Floorlet

As in the case of Heston models (see Eq. (5.2.78)), we evaluate a caplet/floorlet in the $\mathcal{T}_n$-forward measure:

$$N\mathbb{E}\left[\left(\left(\frac{I(\mathcal{T}_n)}{I(\mathcal{T}_{n-1})} - 1 \right) \psi_n - \kappa \, \phi_n \right)^{\pm} \right] = N \, B(0,T) \mathbb{E}_{\mathcal{T}_n}[(\mathcal{R}_n(\mathcal{T}_n) - K)^{\pm}]. \tag{5.2.78}$$

To proceed with analytically, following Mercurio and Moreni, we freeze the values of volatilities in the drift term

$$r_n(t) \to \bar{r}_n(t) = -\rho \, \tilde{\mu}_{n-1}(t) \cdot \sum_{m=\tilde{\eta}(t)}^{n-1} \mathcal{N}_m^R \, \gamma_m^R(0) - \gamma_n^R(0) \sum_{m=\tilde{\eta}(t)}^{n-1} C_{nm}^R \, \gamma_m^R(0), \tag{5.2.79}$$

where $\tilde{\mu}$ is taken within a frozen Libor approximation. Under the approximation (5.2.79), the drift $r_n(t)$ becomes deterministic and can be removed in Eq. (5.2.75) via transformation

$$\mathcal{R}_n(t) = \tilde{\mathcal{R}}_n(t) e^{\int_0^t \bar{r}_n(t')dt'}. \tag{5.2.80}$$

The dynamics of the transformed variable become

$$d\tilde{\mathcal{R}}_n(t) = \tilde{\mathcal{R}}_n(t)\,\gamma_n^R(t)\cdot dW_n^R. \tag{5.2.81}$$

Equations (5.2.74) and (5.2.81) correspond to the SABR model with $\beta = 1$. An option on the original rate R_n can be found from the option on the transformed rate $\tilde{R}_n$ according to

$$\mathbb{E}_{\mathcal{T}_n}\left[(\mathcal{R}_n - K)^+\right] = Z_n \mathbb{E}_{\mathcal{T}_n}\left[(\tilde{\mathcal{R}}_n - K/Z_n)^+\right], \tag{5.2.82}$$

$$Z_n = e^{\int_0^{T_n} \bar{r}_n(t)dt}. \tag{5.2.83}$$

Mercurio and Moreni use the results of the original approximate treatment by Hagan et al. In this approach, the option value is given by the standard Black-Scholes formula

$$\mathbb{E}\left[(\mathcal{R}_n - K)^+\right] = Z_n\mathcal{R}_n(0)\,\Phi(d_+) - K\Phi(d_-), \tag{5.2.84}$$

$$d_\pm = \frac{\ln(Z_n\mathcal{R}_n(0)/K) \pm \sigma^2(K)\mathcal{T}_n/2}{\sigma_n(K)\sqrt{\mathcal{T}_n}}, \tag{5.2.85}$$

with implied volatility of the nth caplet/floorlet given by

$$\sigma_n(K) = Z_n\alpha_n\frac{z_n}{x(z_n)}\left[1 + \left(\frac{\rho\nu_n\alpha_n}{4} + \nu_n^2\frac{2-3\rho^2}{24}\right)\mathcal{T}_n\right], \tag{5.2.86}$$

$$z_n = \frac{\nu_n}{\alpha_n}\ln\left(\frac{Z_n\mathcal{R}_n(0)}{K}\right), \tag{5.2.87}$$

$$x(z) = \ln\left(\frac{\sqrt{1-2\rho z + z^2} + z - \rho}{1-\rho}\right). \tag{5.2.88}$$

Research revealed that, while SABR formulas provide a good approximation as long as stochastic volatilities/maturities are not too large, the approximation of frozen drift is not always applicable. Significant deviations from Monte Carlo results are observed, for example, at maturities more than 10y, with other parameters are close to moderate market values: $\gamma \sim 1\%$, $\nu^2 T_m \sim 0.5$.

The Numerix SABR IMM implementation is based on the extension of Hagan et al. approximations beyond the freezing of the second term in Eq. (5.2.77). This approach is presented in Section 5.2.8.

5.2.7 IMM Calibration

Standalone IMM is calibrated to the following instruments:

- ZCII Swaps
- YOYII Swaps
- ZCII Options
- YOYII Caps/Floors

One should keep in mind that the values of a cap, floor and a swap are coupled by a parity relation (for the same term structure and strike). Therefore, one should avoid inclusion of all three instruments into the calibration target.

Calibration of the hybrid IMM+LMM model is performed in two steps. First, one calibrates the standalone LMM to the IR market sector. The calibrated LMM component is used to create the hybrid model. Then one calibrates the hybrid model to the inflation data sector in the same way as one calibrates the standalone IMM. Below we describe the calibration procedure of the standalone IMM.

Calibration to a term structure of ZCII swaps is performed separately and prior to calibrating the model to other inflation instruments. Indeed, the break-even spreads (K in Eq. (5.2.3)) allow one to obtain the values of forward indices for each swap maturity $\mathcal{T}_i$. These values, in turn, determine initial values of the CPI forwards $F_i(0)$ since each index forward $F_i(t)$ is a martingale in the forward measure associated with its maturity. Thus, we assume that the initial values of CPI forwards are obtained from ZCII swap market quotes and provided to the model.

The rest of the calibration procedure consists of tuning the unknown model parameters in order to achieve the best agreement with market prices of the inflation instruments that are included into the calibration target. In particular, in the case of index-based models, one has to find the optimal values of the volatility $\gamma_n^I(t)$ and of correlation vectors $\mathcal{N}_{n,f}^I$ in Eq. (5.2.12). (Correlation angles are assumed to be constant in time). In the case of rate-based models, the unknowns are $\gamma_n^R(t)$ and $\mathcal{N}_{n,f}^R$ in Eq. (5.2.19). In the presence of stochastic volatility, one should also optimize over parameters that define the stochastic volatility dynamics. Those are a and $\gamma(t)$ in Eq. (5.2.51), in the case of the Heston model, and ν_m in Eq. (5.2.74) in the case of the SABR model. Apparently, the configuration space of model paraments is much larger than the number of market quotes. Thus, to proceed, one should impose some restrictions on the form of the adjusted parameters/functions.

At present, two schemes are available for calibration of IMM. In the first scheme, one assumes that volatility and correlation vectors have some simple functional forms that are set by a few parameters. The second scheme is similar to the first one, but it tries to adopt the number of needed calibration parameters to maturities of the instruments included into the target. Below we describe the two schemes in more detail. To be concrete, we consider the calibration of the index-based model, but all considerations are applicable to the rate-based model as well. In the latter case, all consideration should refer to the rate volatility (5.2.19), rather than to the index volatility (5.2.12).

Calibration Scheme I

In the first scheme, we present the volatility $\gamma_n^I(t)$ in Eq. (5.2.12) as a product of two functions that depend on the index maturity $\mathcal{T}_n$ and time t:

$$\gamma_n^I(t) = f_1(\mathcal{T}_n) f_2(\mathcal{T}_n - t). \tag{5.2.89}$$

Each of the functions f_1, f_2 is a curve going through a set of points. To avoid non-uniqueness of the parametrization, we set $f_2(0) = 0$. In practice, it is good enough to use piecewise-linear interpolation. To set the entire volatility matrix (5.2.12), we have to specify the correlation vectors $\mathcal{N}_n^I$. We assume that there are two factors and present $\mathcal{N}_n^I$ in the form

$$\mathcal{N}_n^I = (\cos(\theta_n), \sin(\theta_n)), \tag{5.2.90}$$

where θ_n is the correlation angle. In this case, the correlation matrix (5.2.13) takes the simple form

$$C^I_{nm} = \cos(\theta_n - \theta_m). \tag{5.2.91}$$

Further, we assume that θ_n depends on the index n through the maturity time $\mathcal{T}_n$:

$$\theta_n(t) = f_3(\mathcal{T}_n). \tag{5.2.92}$$

The function f_3 is also presented as a curve going through a set of points. A constant shift of the correlation angles $\theta_n(t)$ does not change the correlation matrix, and thus we assume $f_3(T_0)$ to be frozen during the calibration.

To summarize, in calibration scheme I, we find the functions f_1, f_2 and f_3 by fitting the market values of target instruments. Each function is set via interpolation through a fixed number of points.

Calibration Scheme II

In the second calibration scheme, we assume a particular analytic dependence of the volatility $\gamma_n(t)$ on time that is set by two parameters per single forward index:

$$\gamma_n(t) = a_n(1 - e^{-k_n(\mathcal{T}_n - t)}). \tag{5.2.93}$$

As in the case of scheme I, we assume the two-factor case and use time independent correlation angles (5.2.90) for setting the correlation vectors $\mathcal{N}^I_n$. Thus, there are three parameters (a_n, k_n and θ_n) per each forward index, which allows for fitting three different instruments (ZC option, YOY swap and cap/floor) per index. Usually, inflation instrument maturities do not cover forward times from the set (5.2.8), so the problem is still over parametrized. To further reduce the number of parameters, we assume that functions a_n, k_n and θ_n are piecewise-constant functions of forward maturities. For example, if the first available maturity is $\mathcal{T}_n$, then all coefficients with $\mathcal{T}_n \leq \mathcal{T}_k$ are constant. It may also happen that for a certain forward index n, there is more than zero but less than three instruments that mature at $\mathcal{T}_n$. In that case, we do not tune all three parameters. In particular, in the case of a single instrument maturing at $\mathcal{T}_n$, we only adjust the coefficient a_n, and in the case of two instruments, we only adjust a_n and k_n.

Stochastic Volatility

Quotes on ZCII options and YOYII caps/floors are often available at several strikes per instrument (with the other parameters kept the same). Calibration to such sets of instruments cannot be achieved within IMM with deterministic volatility. IMM with stochastic volatility allows capturing the volatility smile/skew by tuning the stochastic volatility parameters. When calibrating the Heston model, we assume time independent functions $a(t) = a$ and $\gamma(t) = \gamma$ in Eq. (5.2.51). For the Heston model we therefore have two more optimization parameters, in addition to those considered in previous sections. In the case of SABR model, we assume that ν_n in Eq. (5.2.74) depends on the index n through the index maturity time, $\nu_n = \nu(\mathcal{T}_n)$. Further, we assume that the function $\nu(\mathcal{T}_n)$ is either constant or linear. So, in the case of the SABR model, the stochastic volatility can be parametrized by either one or two parameters.

5.2.8 SABR IMM Model Beyond Frozen Drift Approximation

In this section, we derive approximate option pricing formulas for the model defined by Eqs. (5.2.74), (5.2.75 and (5.2.77). We will assume that transformation (5.2.6) has been performed. The dynamics of $\tilde{R}$ then become

$$d\tilde{\mathcal{R}}_n(t) = \tilde{\mathcal{R}}_n(t)\,(\tilde{r}_n(t)\,dt + \gamma_n^R(t)\,dW_n^R), \tag{5.2.94}$$

where

$$\tilde{r}_n(t) = r_n(t) - \bar{r}_n(t). \tag{5.2.95}$$

The first term in expression (5.2.77), which appears due to correlation with interest rates, will still be treated in the frozen drift approximation, so that $\tilde{r}_n$ becomes

$$\tilde{r}_n(t) = -\sum_{m=\tilde{\eta}(t)}^{n-1} C_{nm}^R\,(\gamma_n^R(t)\,\gamma_m^R(t) - \gamma_n^R(0)\,\gamma_m^R(0)). \tag{5.2.96}$$

We stress that the improvement of the frozen drift approximation considered in this section refers to the second term in expression (5.2.77) only.

Introducing the logarithmic variables $x_n(t) = \ln \gamma_n(t)$, $f_n(t) = \ln \tilde{R}_n(t)$, we write Eqs. (5.2.74), (5.2.94) and (5.2.96) as

$$dx_n(t) = -\frac{\nu_n^2}{2}\,dt + \nu_n\,dV, \tag{5.2.97}$$

$$f_n(t) = (\tilde{r}_n(t) - e^{2x_n(t)}/2)\,dt + e^{x_n(t)}\,dW, \tag{5.2.98}$$

$$\tilde{r}_n(t) = -\sum_{m=\tilde{\eta}(t)}^{n-1} C_{nm}^R\,(e^{x_n(t)+x_m(t)} - e^{x_n(0)+x_m(0)}). \tag{5.2.99}$$

The corresponding forward Kolmogorov equation is

$$\left(\frac{\partial}{\partial t} - \frac{e^{2x_n}}{2}\left(\frac{\partial^2}{\partial f_n^2} + \frac{\partial}{\partial f_n}\right) + r_n\frac{\partial}{\partial f_n} - \frac{1}{2}\sum_{m_1,m_2}\nu_{m_1}\nu_{m_2}\left(\frac{\partial^2}{\partial x_{m_1}\partial x_{m_2}} + \delta_{m_1,m_2}\frac{\partial}{\partial x_{m_1}}\right)\right.$$
$$\left. - \rho_n^{\text{iv}}\sum_m \nu_m \frac{\partial^2}{\partial f_n \partial x_m} e^{x_n}\right) P(\ldots; x_1 \ldots x_N, f_n, t) = 0. \tag{5.2.100}$$

As in [32], we will derive approximate formulae for the implied volatility. Under the approximation used by Hagan et al., the implied volatility can be written in the form

$$\sigma(K) = \tilde{\sigma}(K)\,(1 + g\,T). \tag{5.2.101}$$

This expression is composed of non-perturbative and perturbative factors. The non-perturbative factor, $\tilde{\sigma}(K)$, represents the exact answer for implied volatility in the limit $T \to 0$. The second perturbative factor contains a leading order correction in T. We note that the two factors of the answer can be considered separately, and even using different techniques. We begin with derivation of implied volatility at $T \to 0$ (that is, $\tilde{\sigma}(K)$) within the Wentzel, Kramers, and Brillouin (WKB) approach. In the following section, we demonstrate this technique by rederiving the result by Hagan et al. for implied volatility in the limit $T \to 0$.

Derivation of the SABR Result in the Limit $T \to 0$ via WKB Approach

The SABR model with $\beta = 1$ is defined by stochastic equations

$$dF(t) = \alpha\, F(t)\, dW_1(t), \tag{5.2.102}$$

$$d\alpha(t) = \nu\, \alpha(t)\, dW_2(t), \tag{5.2.103}$$

where the Brownian noises $W_1(t)$ and $W_2(t)$ are correlated by $dW_1(t)\, dW_2(t) = \rho dt$. The stochastic variable F usually represents the underlying financial instrument, and α models the stochastic volatility of the underlying. According to [32], for the SABR model above $\tilde{\sigma}(K)$ in Eq. (5.2.101) is

$$\tilde{\sigma}(K) = \alpha \frac{z}{x(z)}, \tag{5.2.104}$$

where

$$z = \frac{\nu}{\alpha} \ln\left(\frac{F(0)}{K}\right), \tag{5.2.105}$$

$$x(z) = \ln\left(\frac{\sqrt{1 - 2\rho z + z^2} + z - \rho}{1 - \rho}\right). \tag{5.2.106}$$

Below we rederive the result (5.2.104) using the WKB approach. Let us fix the strike at a certain level above or below the initial stock value and assume a small maturity. Under such conditions, the probability for a stock to reach the strike becomes negligibly small. Presenting the probability density as

$$P = e^{-A}, \tag{5.2.107}$$

we conclude that the exponent A grows to infinity with lowering maturity. This allows solving forward/backward Kolmogorov equations in WKB approximations in a manner similar to, for example, solving Schrodinger equations in the quasi-classical limit $\hbar \to 0$. The forward Kolmogorov equation in the logarithmic variables $f = \ln F$ and $x = \ln \alpha$ is

$$\left(\frac{\partial}{\partial t} - \frac{e^{2x}}{2}\left(\frac{\partial^2}{\partial f^2} + \frac{\partial}{\partial f}\right) - \frac{\nu^2}{2}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial x}\right) - \nu\rho \frac{\partial^2}{\partial f \partial x} e^x\right) P(...; f, x, t) = 0. \tag{5.2.108}$$

Substituting $P = e^{-A}$, one can neglect the terms $\partial^2 A/\partial f^2$ and $\partial^2 A/\partial x^2$ since in the limit $T \to 0$ they negligible compared to $(\partial A/\partial f)^2$ and $(\partial A/\partial x)^2$, respectively. We obtain

$$-\frac{\partial A}{\partial t} - \frac{e^{2x}}{2}\left(\frac{\partial A}{\partial f}\right)^2 - \frac{\nu^2}{2}\left(\frac{\partial A}{\partial x}\right)^2 - \nu\rho\, e^x \frac{\partial A}{\partial f}\frac{\partial A}{\partial x} = 0. \tag{5.2.109}$$

Eq. (5.2.109) has the form of a Hamilton-Jacobi equation. It can be solved as follows. We introduce the new variables

$$p = \partial A/\partial x, \qquad P = \partial A/\partial f, \tag{5.2.110}$$

which are analogous to the momenta. We write the corresponding Hamiltonian

$$H(x, P, p) = \frac{e^{2x}}{2} P^2 + \frac{\nu^2}{2} p^2 + \nu\, \rho\, e^x\, P\, p, \tag{5.2.111}$$

so t Eq. (5.2.109) becomes $\frac{\partial A}{\partial t} = -H$. The action A can be found as a value of the functional

$$A = \int \left(-H + \sum_i p_i q_i' \right) dt, \tag{5.2.112}$$

along the trajectory that minimizes the action. Here $\vec{q} = (x, f)$ and $\vec{p} = (p, P)$ are coordinate and momentum vectors, respectively, and the prime ($\prime$) means a derivative with respect to time. The trajectory that minimizes the action satisfies Hamilton equations

$$q_i' = \frac{\partial H}{\partial p_i}, \qquad p_i' = -\frac{\partial H}{\partial q_i}. \tag{5.2.113}$$

In the case of (5.2.111), these equations are

$$x'(t) = \nu^2 p(t) + \nu\,\rho\, e^{x(t)} P(t), \tag{5.2.114}$$
$$f'(t) = P(t)\, e^{2x(t)} + \nu\,\rho\, e^{x(t)} p(t), \tag{5.2.115}$$
$$p'(t) = -P^2(t) e^{2x(t)} - \nu\,\rho\, e^{x(t)}\, P(t)\, p(t), \tag{5.2.116}$$
$$P'(t) = 0. \tag{5.2.117}$$

Eqs. (5.2.114)–(5.2.117) must be supplemented by four boundary conditions. We know the initial and final values of the stock and the initial value of the stochastic volatility:

$$f(0) = \ln F(0), \tag{5.2.118}$$
$$f(T) = \ln K, \tag{5.2.119}$$
$$x(0) = \ln \alpha_0. \tag{5.2.120}$$

The missing fourth condition could be taken as the value of stochastic volatility at option maturity. Calculating the option value would then require an additional integration of the probability density that was obtained using the WKB method over the boundary value $x(T)$. It is more convenient, however, to incorporate this integration into the scheme. In the limit $T \to 0$, such integration is reduced to minimizing the action with respect to $x(T)$. This, in turn, means that

$$\left.\frac{\partial A}{\partial x}\right|_{t=T} = p(T) = 0. \tag{5.2.121}$$

One can see from Eq. (5.2.117) that P is time independent. Keeping in mind the mechanical analogy, one can say that the f-component of the momentum is conserved because the Hamiltonian (5.2.111) is invariant with respect to translations along the f-coordinate. One also concludes that the energy

$$E = \frac{e^{2x(t)}}{2} P^2(t) + \frac{\nu^2}{2} p^2(t) + \nu\rho\, e^{x(t)}\, P(t)\, p(t) \tag{5.2.122}$$

is conserved along the trajectory, since the Hamiltonian (5.2.111) does not depend explicitly on time. Taking a derivative of the energy along the trajectory defined by Eqs. (5.2.114)–(5.2.117), one can see that dE/dt is indeed zero. Due to the boundary condition (5.2.121), we have at

maturity $p(T) = 0$. From Eq. (5.2.122), we find that

$$E = \frac{P^2(T)}{2}\, e^{2x(T)}. \tag{5.2.123}$$

Energy conservation allows simplifying the answer for the action along the optimal trajectory: The Hamiltonian H in Eq. (5.2.112) is time independent and equals E along the trajectory. One can show that$\sum_i p_i q_i' = 2E$ along the trajectory as well. Indeed,

$$\sum_i p_i q_i' = \sum_i p_i \frac{\partial H}{\partial p_i} = P(e^{2x}P + \nu\rho e^x p) + p(\nu^2 p + \nu\,\rho\, e^x P) = 2E. \tag{5.2.124}$$

The action takes a simple form,

$$A = E\,T. \tag{5.2.125}$$

Thus, the probability density is determined by the energy E of the optimal trajectory, which satisfies Eqs. (5.2.114)–(5.2.117) and boundary conditions (5.2.118)–(5.2.121). We note that Eq. (5.2.125) should be interpreted as the leading-order term of the logarithm of the stock probability density in the limit $T \to 0$. Therefore, the implied volatility of the stock option can be found from

$$-ET = -\frac{(f(T) - f(0))^2}{2\sigma^2(K)T}, \tag{5.2.126}$$

and is given by

$$\sigma(K) = \pm\frac{\left|f(T) - f(0)\right|}{\sqrt{2E}\,T}, \tag{5.2.127}$$

where the upper (lower) sign is for a call (put) option. Note that we expect $\sigma(K)$ to converge to a finite, nonzero value in the limit $T \to 0$. Thus, the energy E is expected to scale with maturity as $E \sim 1/T^2$.

Now we turn to finding the optimal trajectory. We solve Eq. (5.2.122) with respect to p:

$$p = \frac{1}{\nu}\left(-\rho\, e^x\, P \pm \sqrt{(\rho^2 - 1)\, P^2\, e^{2x} + 2E}\right) \tag{5.2.128}$$

and use this equation in Eq. (5.2.114). The resulting equation,

$$x' = \pm\nu\sqrt{(\rho^2 - 1)\, P^2\, e^{2x} + 2E}, \tag{5.2.129}$$

can be integrated from the origin to t, resulting in

$$\text{arctanh}\,\frac{\sqrt{2E - (1-\rho^2)P^2e^{2x}}}{\sqrt{2E}}\Bigg|_{x_0}^{x(t)} = \mp\sqrt{2E}\nu t. \tag{5.2.130}$$

Setting the upper time to maturity ($t = T$) and using Eq. (5.2.123), we get

$$\ln\frac{\sqrt{2E} - \sqrt{2E - (1-\rho^2)\, P^2\, e^{2x_0}}}{(1-\rho)\, P\, e^{x_0}} = \mp\sqrt{2E}\nu T. \tag{5.2.131}$$

This equation is very useful since it relates the energy E with the momentum P. Yet the energy is not determined by Eq. (5.2.131) alone. To complete the task, we will integrate Eq. (5.2.115).

Inserting p, given by Eq. (5.2.128), into s Eq. (5.2.115), we obtain

$$f' = e^{2x}\,(1-\rho^2)\,P + \rho\,e^{x}\,\sqrt{(\rho^2-1)\,P^2\,e^{2x}+2E}. \tag{5.2.132}$$

The derivative over time can be turned into a derivative over x with the help of Eq. (5.2.129):

$$\frac{df}{dx} = \frac{1}{\nu}\left(\rho\,e^{x} \pm \frac{(1-\rho^2)\,P\,e^{2x}}{\sqrt{(\rho^2-1)\,P^2\,e^{2x}+2E}}\right). \tag{5.2.133}$$

Integrating this equation from $x(0)$ to $x(T)$, and using Eq. (5.2.123), we obtain

$$\nu\,(f(T)-f(0)) = \pm\sqrt{2E/P^2-(1-\rho^2)\,e^{2x(0)}} - \rho\,e^{x(0)}. \tag{5.2.134}$$

This equation can be solved for $2E/P^2$:

$$\frac{2E}{P^2} = \nu^2\,(f(T)-f(0))^2 + 2\nu\rho(f(T)-f(0))\,e^{x(0)} + \alpha_0^2 = \alpha_0^2\,(z^2-2\rho z+1), \tag{5.2.135}$$

where z is given by

$$z = \frac{\nu}{\alpha}\,\ln\left(\frac{F(0)}{K}\right). \tag{5.2.136}$$

Now inserting the expression for $2E/P^2$ from Eq. (5.2.135) into Eq. (5.2.131), we obtain the energy:

$$\sqrt{2E}\,\nu\,T = \mp x(z), \tag{5.2.137}$$

$$x(z) = \ln\left(\frac{\sqrt{1-2\rho z+z^2}+z-\rho}{1-\rho}\right). \tag{5.2.138}$$

According to Eq. (5.2.127), the implied volatility becomes

$$\sigma(K) = \alpha_0\frac{z}{x(z)}, \tag{5.2.139}$$

which agrees with Eq. (5.2.104). Concluding this subsection, we present an expression for the optimal trajectory $x(t)$ at any time below maturity:

$$e^{x-x_0} = \frac{1}{\cosh(x(z)\,t/T) - \frac{z-\rho}{\sqrt{z^2-2\rho z+1}}\sinh(x(z)t/T)}. \tag{5.2.140}$$

This equation can be derived from Eqs. (5.2.130), (5.2.135) and (5.2.137).

IMM SABR Model in the Limit $T \to 0$

Having demonstrated the application of the WKB technique on an example of the standard SABR model, here we will derive the implied volatility $\tilde{\sigma}$ in the case of IMM SABR model. We substitute the probability density in the form (5.2.107) into Eq. (5.2.100). Neglecting the terms that are

small in the limit $T \to 0$, we obtain the Hamilton-Jacobi equation

$$-\frac{\partial A}{\partial t} - \frac{e^{2x_n}}{2}\left(\frac{\partial A}{\partial f_n}\right)^2 - \frac{1}{2}\sum_{m_1,m_2}\nu_{m_1}\nu_{m_2}\frac{\partial A}{\partial x_{m_1}}\frac{\partial A}{\partial x_{m_2}} - \rho_n^{\text{iv}}\, e^{x_n}\frac{\partial A}{\partial f_n}\sum_m \nu_m \frac{\partial A}{\partial x_m} + \tilde{r}_n(t)\frac{\partial A}{\partial f_n} = 0. \quad (5.2.141)$$

Denoting partial derivatives as

$$P_n = \frac{\partial A}{\partial f_n}, \quad p_m = \frac{\partial A}{\partial x_m}, \quad (5.2.142)$$

we define the Hamiltonian as

$$H(\mathbf{x}, P_n, \mathbf{p}) = H_0(x_n, P_n, \mathbf{p}) - P_n\,\tilde{r}_n(\mathbf{x}, t), \quad (5.2.143)$$

$$H_0(x_n, P_n, \mathbf{p}) = \frac{e^{2x_n}}{2}P_n^2 + \frac{1}{2}\sum_{m_1,m_2}\nu_{m_1}\nu_{m_2}p_{m_1}p_{m_2} + \rho_n^{\text{iv}}e^{x_n}\,P_n\sum_m \nu_m\, p_m, \quad (5.2.144)$$

where $\mathbf{x} = (x_1, ..., x_N)$, $\mathbf{p} = (p_1, ..., p_N)$. To obtain an analytic solution, we will consider the term containing the drift $\tilde{r}_n(\mathbf{x}, t)$ as a perturbation. This term is small indeed, since $P_n \sim 1/T$. (This follows from Eq. (5.2.123) and that $E \sim 1/T^2$.) Up to leading order in $\tilde{r}_n(\mathbf{x}, t)$, the action can be found by adding the corresponding term to the zero order action:

$$A(\mathbf{x}; f_n; T) = A_0(\mathbf{x}; f_n; T) + P_n\int \tilde{r}_n(\mathbf{x}, t)dt, \quad (5.2.145)$$

Here we used the fact that momentum P_n, defined with respect to the zero-order Hamiltonian H_0, is time independent. The perturbative term (the second term in Eq. (5.2.145)) can be absorbed into the shift of the value $f_n(T)$. Indeed,

$$A_0(f_n - \delta; \mathbf{x}; T) - A_0(f_n; \mathbf{x}; T) \approx -\delta\frac{\partial}{\partial f_n}A_0(f_n; \mathbf{x}; T) = -P_n\delta. \quad (5.2.146)$$

According to Eqs. (5.2.99) and (5.2.145), the shift δ is

$$\delta = -\int_0^T \sum_{m=\tilde{\eta}(t)}^{n-1} C_{nm}^R\,(e^{x_n(t)+x_m(t)} - e^{x_n(0)+x_m(0)})dt. \quad (5.2.147)$$

Note that to zeroth order, the option value is given by Eqs. (5.2.86)–(5.2.88). The correction (5.2.147) leads to the following contribution to the coefficient Z_n in Eqs. (5.2.86) and (5.2.87):

$$Z_n \to Z_n\, e^{\delta}. \quad (5.2.148)$$

This result has a straightforward meaning of the drift term integrated over time along the optimal trajectory. The optimal trajectory can be calculated with respect to the non-perturbed

Hamiltonian H_0. Thus, it is given by the Hamilton equations

$$x'_m(t) = \nu_m \sum_{m_1} \nu_{m_1}\, p_{m_1}(t) + \nu_m\, \rho_n^{\mathsf{iv}}\, e^{x_n(t)}\, P_n(t), \tag{5.2.149}$$

$$f'_n(t) = e^{2x_n(t)} P_n(t) + \rho_n^{\mathsf{iv}} e^{x_n} \sum_m \nu_m\, p_m, \tag{5.2.150}$$

$$p'_n(t) = -e^{2x_n(t)} P_n^2(t) - \rho_n^{\mathsf{iv}}\, e^{x_n} P_n \sum_m \nu_m\, p_m, \tag{5.2.151}$$

$$p'_m(t) = 0, \qquad m \neq n, \tag{5.2.152}$$

$$P'_n(t) = 0. \tag{5.2.153}$$

We see that all momenta p_m with $m \neq n$ are conserved. The boundary condition (5.2.121) is applicable for all x_n. Thus, we conclude that

$$p_m = 0, \quad \text{for} \quad m \neq n. \tag{5.2.154}$$

This allows simplifying Eqs. (5.2.149)–(5.2.153):

$$x'_m(t) = \nu_m \nu_n\, p_n(t) + \nu_m\, \rho_n^{\mathsf{iv}}\, e^{x_n(t)}\, P_n, \tag{5.2.155}$$

$$f'_n(t) = e^{2x_n(t)} P_n + \nu_n\, \rho_n^{\mathsf{iv}}\, e^{x_n(t)}\, p_n(t), \tag{5.2.156}$$

$$p'_n(t) = -e^{2x_n(t)} P_n^2 - \nu_n\, \rho_n^{\mathsf{iv}}\, e^{x_n(t)}\, P_n\, p_n(t). \tag{5.2.157}$$

Note that at $m = n$ the above equations coincide with Eqs. (5.2.114)–(5.2.116). So, as one expects, the dynamics of the nth component is described by the standard SABR expression. Also, from Eq. (5.2.155), one can see that

$$x_m(t) - x_m(0) = \frac{\nu_m}{\nu_n}(x_n(t) - x_n(0)). \tag{5.2.158}$$

This equation allows getting all trajectories with $m \neq n$ from the trajectory $x_n(t)$ that is given by Eq. (5.2.140). Plugging these results into the expression for shift δ, we obtain

$$\delta = -\int_0^T \sum_{m=\tilde{\eta}(t)}^{n-1} \frac{C_{nm}\, \gamma_n(0)\, \gamma_m(0)}{\left(\cosh(x(z_n)\, t/T_n) - g_n \sinh(x(z_n)\, t/T_n)\right)^{1+\nu_m/\nu_n}} dt, \tag{5.2.159}$$

$$g_n = \frac{z_n - \rho_n^{\mathsf{iv}}}{\sqrt{z_n^2 - 2\rho_n^{\mathsf{iv}} z_n + 1}}. \tag{5.2.160}$$

IMM SABR Model in the Limit $T \to 0$.

Now we turn to calculating the perturbation factor in Eq. (5.2.101). This term can be obtained by applying standard perturbation techniques to the implied volatility, considering the dynamics of stochastic volatility as a perturbation. In particular, in the case of the $\beta = 1$ SABR model, the coefficient g is

$$g = \frac{\rho\nu\alpha}{4} + \nu^2 \frac{2 - 3\rho^2}{24}. \tag{5.2.161}$$

In addition to the frozen drift approximation, one should include the contribution of the factor (5.2.99). It turns out that, to leading order, one can simply average this term over dynamics

of x_m given by Eq. (5.2.97). The resulting value of the shift is given by

$$\delta_n = -\frac{1}{2}\gamma_n(0)\,\nu_n \sum_{m=\tilde{\eta}(t)}^{n-1} C_{nm}\,\gamma_m(0)\,\nu_m T_m^2. \tag{5.2.162}$$

5.2.9 Appendices for Inflation Market Models

Appendix A

Calculation of Expected Value of a Ratio of Index Forwards

Here we consider the stochastic differential equation (5.2.9):

$$dF_n(t) = F_n(t)\,\lambda_n^I(t)\cdot(\rho\,\tilde{\mu}_n(t)\,dt + dW^I(t)). \tag{5.2.163}$$

With the help of Ito's lemma, for the dynamics of logarithm of the index forward $y_n(t) = \ln F_n(t)$, we get

$$dy_n(t) = \lambda_n^I(t)\cdot\left(\rho\,\tilde{\mu}_n(t)\,dt + dW^I(t)\right) - \frac{1}{2}\,D_{nn}^I(t)\,dt, \tag{5.2.164}$$

where $D_{nm}^I(t) = \lambda_n^I(t)\cdot\lambda_m^I(t)$. The solution of this equation is

$$y_n(t) = y_n^{(0)}(t) + y_n'(t), \tag{5.2.165}$$

$$y_n^{(0)}(t) = y_n(0) + \int_0^t \left(\lambda_n^I(t')\,\rho\cdot\tilde{\mu}_n(t') - \frac{1}{2}D_{nn}(t')\right)dt', \tag{5.2.166}$$

$$y_n'(t) = \int_0^t \lambda_n(t')\cdot\frac{dW(t')}{dt'}dt'. \tag{5.2.167}$$

Considering $\tilde{\mu}_n(t)$ as deterministic, the covariance of logarithmic variables is

$$V_{nm}(t) = \mathbb{E}\left[y_n'(t)\cdot y_m'(t)\,|\,\mathcal{F}_t\right] = \int_0^t D_{nm}(t)dt. \tag{5.2.168}$$

The expectation of nth index forward is volatility independent:

$$\mathbb{E}\left[e^{y_n(t)}\right] = y_n(0)e^{\int_0^t \lambda_n^I(t')\,\rho\cdot\tilde{\mu}_n(t')dt'}. \tag{5.2.169}$$

The expectation of the ratio of two index forwards is

$$\mathbb{E}\left[e^{y_n(t_1)-y_m(t_2)}\right] = e^{y_n^{(0)}(t_1)-y_m^{(0)}(t_2)+\frac{1}{2}\tilde{V}_{nm}(t_1,t_2)}, \tag{5.2.170}$$

where the $\tilde{V}_{nm}(t_1,t_2)$ is variance of the difference of the two log-forwards:

$$\tilde{V}_{nm}(t_1,t_2) = V\left[(y_n(t_1)-y_m(t_2))^2\right] = V_{nn}(t_1) + V_{mm}(t_2) - 2V_{nm}(t), \qquad t = \min\{t_1,t_2\}. \tag{5.2.171}$$

Substituting Eq. (5.2.166) into Eq. (5.2.170) gives

$$\mathbb{E}\left[e^{y_n(t_1)-y_m(t_2)}\right] = e^{\rho\int_0^{t_1}\lambda_n^I(t')\cdot\tilde{\mu}_n(t')dt' - \rho\int_0^{t_2}\lambda_m^I(t')\cdot\tilde{\mu}_m(t')dt' + V_{mm}(t_2) - V_{nm}(t)}. \tag{5.2.172}$$

Appendix B

Option Value in Terms of the Characteristic Function

Consider a call/put option on the underlying $z = X_0\, e^y$ with option value

$$(X_0 e^y - K)^{\pm}. \tag{5.2.173}$$

The variable y is normalized such that it begins its stochastic evolution from zero value. Our goal here is to derive the expression for the call/put option in terms of the function

$$\phi(\xi) = \mathbb{E}\left[e^{\xi y}\right]. \tag{5.2.174}$$

We begin with the identity

$$(X_0 e^y - K)^{\pm} = \frac{K}{2\pi i}\int_{C^{\pm}} \frac{1}{\xi(\xi-1)} e^{\xi(y-\ln(K/X_0))} d\xi, \tag{5.2.175}$$

where the contour $C^{\pm}$ is parallel to the imaginary axis and goes to the right/left of the poles. Taking the expectation of both sides, we obtain

$$\mathbb{E}\left[(X_0 e^y - K)^{\pm}\right] = \frac{K}{2\pi i}\int_{C^{\pm}} \frac{1}{\xi(\xi-1)} e^{-\xi \ln(K/X_0)} \phi(\xi) d\xi. \tag{5.2.176}$$

Now we move the integration contour such that it remains parallel to the imaginary axis and crosses the real value at $0 < \delta < 1$. I.e., the contour is $\xi = \delta + i\omega$, with $\omega \in \mathbb{R}$. Taking care of pole contributions, we have

$$\mathbb{E}\left[(X_0 e^y - K)^{\pm}\right] = \left\{ \begin{array}{c} X_0\phi(1) \\ K \end{array} \right\} + \frac{K}{2\pi i}\int_{C_{\delta\pm}} \frac{1}{\xi(\xi-1)} e^{-\xi \ln(K/X_0)} \phi(\xi) d\xi. \tag{5.2.177}$$

One can recognize the parity relation of put/call options. In the present approach, we will take $\delta \to 0$. To take care of the pole at $\xi = 0$, let us present the second term in the above expression as

$$\frac{K}{2\pi i}\int_{C_{\delta\pm}} \cdots = -\frac{K}{2} + \frac{K}{4\pi i}\int_{C_{\delta}} \cdots + \frac{K}{4\pi i}\int_{C_{-\delta}} \ldots, \tag{5.2.178}$$

where $C_{-\delta}$ is the contour $\xi = -\delta + i\omega$. Making the transformation $\xi \to -\xi$ in the last term, we obtain

$$\mathbb{E}\left[(X_0 e^y - K)^{\pm}\right] = \left\{ \begin{array}{c} X_0\phi(1) \\ K \end{array} \right\} - \frac{K}{2} + \frac{K}{4\pi i}\int_{C_{\delta}} \left(\frac{1}{\xi(\xi-1)} e^{-\xi \ln(K/X_0)} \phi(\xi) + (\xi \to -\xi) \right) d\xi. \tag{5.2.179}$$

The advantage of this representation is that the pole at $\xi = 0$ does not contribute to the integral, allowing us to set $\delta = 0$. Taking $\xi = i\omega$ we obtain

$$\mathbb{E}\left[(X_0 e^y - K)^{\pm}\right] = \left\{ \begin{array}{c} X_0\phi(1) \\ K \end{array} \right\} - \frac{K}{2} - \frac{K}{2\pi}\Re \int \frac{1}{\omega(\omega + i)} e^{-i\omega \ln(K/X_0)} \phi(i\omega) d\omega. \tag{5.2.180}$$

Appendix C

Solution of Equation (5.2.59)

Here we solve Eq. (5.2.59) in the case where parameters are piecewise constant. We begin with solving the equation for B writing it in the form

$$B' = -\frac{\gamma^2}{2}(B-\alpha)^2 + \frac{1}{2}(D_1\xi - D_2\xi^2) + \frac{\gamma^2}{2}\alpha^2, \tag{5.2.181}$$

where

$$\alpha = \frac{a}{\gamma^2} - \frac{\xi}{\gamma}M. \tag{5.2.182}$$

We rewrite this as

$$\frac{dB}{(B-\alpha)^2 - \frac{\Omega^2}{\gamma^4}} = -\frac{\gamma^2}{2}dt, \tag{5.2.183}$$

where

$$\Omega^2 = \gamma^2\left(D_1\xi - D_2\xi^2 + \gamma^2\alpha^2\right). \tag{5.2.184}$$

Integrating both parts of Eq. (5.2.183) over (t_1, t_2), we obtain

$$\operatorname{arctanh}\frac{(B_2-\alpha)\gamma^2}{\Omega} - \operatorname{arctanh}\frac{(B_1-\alpha)\gamma^2}{\Omega} = \frac{\Omega}{2}(t_2-t_1), \tag{5.2.185}$$

where $B_2 = B(t_2), B_1 = B(t_1)$. Making use of the trigonometric identity

$$\operatorname{arctanh} a - \operatorname{arctanh} b = \operatorname{arctanh}\left(\frac{a-b}{1-ab}\right) \tag{5.2.186}$$

we obtain

$$B(t_1) = B_1 = \alpha + \frac{\Omega}{\gamma^2}\,\frac{(B_2-\alpha)\cosh\frac{\Omega}{2}(t_2-t_1) - \frac{\Omega}{\gamma^2}\sinh\frac{\Omega}{2}(t_2-t_1)}{\frac{\Omega}{\gamma^2}\cosh\frac{\Omega}{2}(t_2-t_1) - (B_2-\alpha)\sinh\frac{\Omega}{2}(t_2-t_1)} \tag{5.2.187}$$

Knowing the solution for B, now we can find the function A. From Eq. (5.2.64) we obtain

$$\begin{aligned} A(t_2) - A(t_1) &= -a\int_{t_1}^{t_2} B(t)dt \\ &= -a\alpha(t_2-t_1) - \frac{\Omega a}{\gamma^2}\int_0^{t_2-t_1}\frac{(B_2-\alpha)\cosh\frac{\Omega}{2}t - \frac{\Omega}{\gamma^2}\sinh\frac{\Omega}{2}t}{\frac{\Omega}{\gamma^2}\cosh\frac{\Omega}{2}t - (B_2-\alpha)\sinh\frac{\Omega}{2}t}\,dt \\ &= -a\alpha(t_2-t_1) + \frac{2a}{\gamma^2}\ln\left[\cosh\frac{\Omega}{2}(t_2-t_1) - \frac{(B_2-\alpha)\gamma^2}{\Omega}\sinh\frac{\Omega}{2}(t_2-t_1)\right]. \end{aligned} \tag{5.2.188}$$

Finally,

$$A(t_1) = A(t_2) + a\alpha(t_2-t_1) - \frac{2a}{\gamma^2}\ln\left[\cosh\frac{\Omega}{2}(t_2-t_1) - \frac{(B_2-\alpha)\gamma^2}{\Omega}\sinh\frac{\Omega}{2}(t_2-t_1)\right]. \tag{5.2.189}$$

Analytic properties of Ω. The above calculation relies on the fact that the square root of Ω^2, given by Eq. (5.2.184), is analytic along the integration contour. Here we show the branch cut

of the square root can always be chosen along real negative axis, which assures the required analytic properties for integration over $\xi = i\omega$. Inserting $\xi = i\omega$ into Eq. (5.2.184), we obtain

$$\frac{\Omega^2}{\gamma^2} = \frac{a^2}{\gamma^2} + \left(D_2 - M^2\right)\omega^2 + \left(D_1 - \frac{2aM}{\gamma}\right) i\omega. \tag{5.2.190}$$

It is sufficient to show that

$$D_2 - M^2 > 0. \tag{5.2.191}$$

Consider first the case where $D_2 = \lambda_n^I \cdot \lambda_n^I$, $M = \lambda_n^I \cdot \rho$. The full correlation matrix of the Brownian noises $dW_1^I, dW_2^I, \ldots, dW_N^I, dV$ can be written as

$$\tilde{\rho} = \begin{pmatrix} \hat{1} & \rho \\ \rho & 1 \end{pmatrix}, \tag{5.2.192}$$

where $\hat{1}$ is the $N \times N$ unity matrix. The full correlation matrix should be positive definite, therefore

$$(\lambda, -\lambda \cdot \rho) \begin{pmatrix} \hat{1} & \rho \\ \rho & 1 \end{pmatrix} \begin{pmatrix} \lambda \\ -\lambda \cdot \rho \end{pmatrix} = \lambda \cdot \lambda - (\lambda \cdot \rho)^2 = D_2 - M^2 > 0. \tag{5.2.193}$$

The case where D_2 and M are given by Eqs. (5.2.72) and (5.2.73) is considered analogously.

Solution as $\omega \to 0$. At low frequency, the solution for A and B have the form

$$A = i\omega\tilde{A}, \tag{5.2.194}$$

$$B = i\omega\tilde{B}. \tag{5.2.195}$$

The equation for $\tilde{B}$ is

$$\tilde{B}' = a\tilde{B} + \frac{1}{2}D_1. \tag{5.2.196}$$

Integrating this equation from t_1 to t_2 results in

$$\tilde{B}_1 = -\frac{D_1}{2a} + \left(\tilde{B}_2 + \frac{D_1}{2a}\right) e^{a(t_1 - t_2)}, \tag{5.2.197}$$

and A is given by

$$\tilde{A}(t_1) = \tilde{A}(t_2) - \frac{D_1}{2}(t_2 - t_1) + \left(\tilde{B}_2 + \frac{D_1}{2a}\right)\left(1 - e^{a(t_1 - t_2)}\right). \tag{5.2.198}$$

Appendix D

Calculation of $\tilde{\mu}_n$ in the case of Hull-White

Here we calculate the function $\tilde{\mu}_n$ in Eq. (5.2.14) for the case of the Hull-White model.

For the Hull-White model, the short rate is given by

$$r(t) = \theta(t) + \sum_{i=0}^{F_{\text{IR}}} s_i(t),$$

where θ is chosen to match the yield curve and

$$ds_i(t) = a_i(t)s_i(t)dt + \sigma_i(t)dW_i^{\text{IR}}(t).$$

The Brownian motions dW_i^{IR} and dW_j^{IR} are correlated by

$$dW_I^{\text{IR}}(t)dW_j^{\text{IR}}(t) = \rho_{i,j}^{\text{IR}}dt,$$

and and the correlation between the CPI and IR sectors is

$$dW_f(t)dW_g^{\text{IR}}(t) = \rho_{f,g}dt.$$

The SDE the Hull-White process s_i is written in the risk-neutral measure. To convert between T-forward and risk-neutral measures, we need the dynamics of zero-coupon bonds. Let $B(t,T)$ denote the value of a zero-coupon bond at time t with maturity T. Then

$$d\log B(t,T) = \cdots + \sigma_B(t,T)dW^{\text{IR}}(t), \tag{5.2.199}$$

where $\ldots$ denotes a drift term and σ_B is the volatility of the bond. The value the bond with maturity T that starts at the value s_i at time t is

$$B(t,s;T) = \mathbb{E}\left(e^{-\int_t^T r(t')dt'} \,\middle|\, x_i(t) = x_i\right), \tag{5.2.200}$$

This definition together with the SDE for s_i then implies

$$\sigma_B^i(t,T) = \frac{\partial \log B(t,s;T)}{\partial s_i}\sigma_i(t). \tag{5.2.201}$$

To calculate the derivative with respect to s_i, we first consider the trajectories that begin their dynamics at $s_i(t)$ and $s_i(t) + \delta s_i(t)$, for some small offset $\delta s_i(t)$. Denote the difference of these two trajectories by $\delta s_i(t)$. Then along a fixed stochastic path, $\delta s_i(t)$ obeys

$$d\left(\delta s_i(t)\right) = -a_i(t)\delta s_i(t)dt, \tag{5.2.202}$$

which gives

$$\delta s_i(\tau) = \delta s_i(t)e^{-\int_t^\tau a_i(t')dt'} \tag{5.2.203}$$

for $\tau \geq t$. Note that Eq. (5.2.203) is deterministic (path independent). Now r is linear in the

state variables s_i, so Eq. (5.2.200) gives

$$\frac{\partial \log B(t,s;T)}{\partial s_i} = -\int_t^T e^{-\int_t^{T'} a_i(t')dt'} dT'. \tag{5.2.204}$$

To change form the risk-neutral measure to the T-forward measure, we make the substitution $dW(t) \to dW(t) + \rho\sigma_B(t,\mathcal{T})dt$, where ρ is the correlation between the CPI and interest rates. Then in the risk-neutral measure, we have

$$\tilde{\mu}_n(t) = -\sigma_B(t,\mathcal{T}_n).$$

Then using Eq. (5.2.204) in Eq. (5.2.201) gives

$$\tilde{\mu}_n^i(t) = \sum_{j=0}^{F_{\text{IR}}} \sigma_j(t) \int_t^{\mathcal{T}_n} e^{-\int_t^{T'} a_j(t')dt'} dT' \tag{5.2.205}$$

in the risk-neutral measure. In the $\mathcal{T}_n$-forward measure, we have

$$\tilde{\mu}_m^i(t) = -\sum_{j=0}^{F_{\text{IR}}} \sigma_j(t) \int_{\mathcal{T}_m}^{\mathcal{T}_n} e^{-\int_t^{T'} a_j(t')dt'} dT'.$$

Chapter 6

Basket Models

6.1 Basket Options of Equities and FX Rates

In this section, we consider the most popular FX or equity basket models with deterministic IRs. Along with their definition, we briefly describe a dimensional reduction technique leading to an effective numerical treatment.

6.1.1 Black-Scholes Equity and FX Basket Models with Forward Monte Carlo Method

Here, we consider Black-Scholes equity basket models with a forward Monte Carlo method. (The FX basket models follow a similar logic.) These models implement a Monte Carlo simulation of products which may depend on a large number of equities. Each equity follows the Black-Scholes process and a correlation matrix of the different equities is defined:

$$dS_i(t) = (r(t) - u_i(t))S_i(t)dt + \sigma_i(t)S_i(t)dW_i(t), \tag{6.1.1}$$

$$\langle dW_i dW_j \rangle = \rho_{ij}. \tag{6.1.2}$$

Internally, the model simulates q Brownian factors dV_j (where q is the quality factor defined by users) and converts these into equity prices using

$$dF_i(t) = \sigma_i(t)F_i(t)\sum_{j=1}^{q} X_{ij}dV_j(t), \tag{6.1.3}$$

where $F_i(t)$ is the equity forward price given by $F_i(t) = S_i(t)\, e^{-\int_0^t (r(s)-u_i(s))\, ds}$.

The transformation matrix X_{ij} is constructed as follows. Calculate the covariance matrix M_{ij} using

$$M_{ij} = \rho_{ij}\bar{\sigma_i}\bar{\sigma_j}, \tag{6.1.4}$$

$$\bar{\sigma_i}^2 = \frac{1}{T_{max}}\int_0^{T_{max}} \sigma_i^2(t)dt, \tag{6.1.5}$$

where T_{max} is the largest event date in the schedule. Take the matrix square root by digitalizing

the covariance matrix and keeping only the q largest eigenvectors:

$$M_{ij} = \sum_{k=1}^{N} \lambda_k e_{ik} e_{jk},$$
$$T_{ik} = \lambda_k^{1/2} e_{ik} \quad (k = 1 \dots q). \tag{6.1.6}$$

We then have the option to rescale the matrix square root to ensure the correct variance for each of the N equities even when $q < N$:

$$T_{ik} \to T_{ik} \sqrt{\frac{M_{ii}}{\sum_{l=1}^{q} T_{il}^2}}. \tag{6.1.7}$$

If this option is selected, the convergence of options on single equities will be improved, but, in general, the convergence of options which depend on weighted averages of multiple equities will get worse. Finally, the transformation matrix is given by

$$X_{ij} = \frac{1}{\bar{\sigma}_i} T_{ij}. \tag{6.1.8}$$

Heston Equity and FX Basket Model and its Two-Factor Dimensional Reduction

Here, we describe a Heston Equity and FX basket model and its two-factor effective approximation obtained from the full multifactor Heston equity basket model by dimensional reduction. Consider $S(t) = \sum_i w_i S_i(t)$, a basket of equities $S_i(t)$ with notionals w_i, where each equity process is driven by a Heston model with time-dependent coefficients:

$$dS_i(t) = (r(t) - u_i(t)) S_i(t) dt + S_i(t) \sqrt{z_i(t)} \lambda_i(t) \cdot dW(t), \tag{6.1.9}$$
$$dz_i(t) = a_i(t) (1 - z_i(t)) dt + \sqrt{z_i(t)} \gamma_i(t) \cdot dW(t), \qquad z_i(0) = 1. \tag{6.1.10}$$

An arbitrary correlation matrix can be used to link together different Heston models. For the moment, the correlations are considered as input. In the future, we plan to add the possibility of their calibration to European options on a basket index.

Using Markovian projection, we approximate a normalized forward basket process

$$F(t) = \frac{S(t)}{\sum_i w_i \; S_i(0) \, e^{\int_0^t (r(s) - u_i(s)) \, ds}},$$

with a single effective Heston process $F(t) \simeq F^*(t)$ that follows the dynamics

$$dF^*(t) = (1 + \beta(t) (F^*(t) - F^*(0))) \sqrt{z(t)} \, \sigma_H(t) \cdot dW(t), \qquad F^*(0) = 1,$$
$$dz(t) = \alpha(t) (1 - z(t)) dt + \sqrt{z(t)} \, \sigma_z(t) \cdot dW(t), \qquad z(0) = 1. \tag{6.1.11}$$

See [12] for details. This approximation is used for basket European pricing as well as numerical evaluation of more complicated basket index instruments.

The Heston basket model is available with full Monte Carlo simulation and also in its reduced two-factor form.

Dupire Equity and FX Basket Model and its One-Factor Dimensional Reduction

Here, we describe a Dupire equity and FX basket model and its one-factor effective approximation obtained from the full multifactor Dupire equity basket model by dimensional reduction. Consider $S(t) = \sum_i w_i\, S_i(t)$, where $S_i(t)$ are equities with notionals w_i and each equity process is driven by a Dupire model (3.2.1):

$$dS_i(t) = (r(t) - u_i(t))S_i(t)dt + S_i(t)\,\Sigma_i(t, S_i(t))dW_i(t). \tag{6.1.12}$$

An arbitrary correlation matrix can be used to link together the Brownian motionss W_i of different Dupire models.

Using the technique of Avellaneda et al. [16], which is based on a short-time PDF expansion, we approximate the basket process with a *single* effective Dupire process

$$dS(t) = (r(t) - u(t))S(t)dt + S(t)\,\Sigma(t, S(t))dW(t), \tag{6.1.13}$$

where $u(t)$ is the effective dividend yield and $\Sigma(t, S(t))$ is the effective vol surface.

The Dupire basket model is available with full Monte Carlo simulation, and also in its reduced one-factor form.

Chapter 7

Hybrid Models

In this chapter, we describe the universal hybrid model construction in Numerix products, which includes cross-currency models, equities in the presence of deterministic or stochastic interest rates, foreign equities and many others. We will give details of correlation setups between different multi-factor models and explain calibration techniques.

In general, modulo several exceptions, a hybrid model can contain multiple IR models with exchange processes and several equity models in domestic or foreign currencies. Below, we present a list of model types for each category:

- **IR models**
  Hull-White (HW), Black-Karasinski (BK), Shifted Black-Karasinski (SBK) one- or multi-factor, LMM[1]
- **FX models**
  Black-Scholes (BS)
- **Equity models**
  Black-Scholes (BS), Dupire, Heston
- **Commodity models**
  Black, Schwartz (S1F), Gibson Schwartz (GS)
- **Credit models**
  Black-Karasinski (BK), Black-Derman-Toy (BDT), Hull-White (HW), Cox-Ingersoll-Ross (CIR)
  Multiple credit processes (up to 5) per currency.

An important class of hybrid models is equity basket models (that is, the models that include multiple equity processes), possibly in several currencies.

7.1 Model Setup

For the moment, we consider the multi-factor uncorrelated vector Brownian factors

$$W(t) = \{W_1(t), W_2(t), \dots\}$$

[1]The CC LMM model (Section A.1) supports IR LMM in both domestic and foreign currencies. In the hybrid models, IR LMM model can only be a domestic component.

that drive all the processes and comment on correlation structure in a separate section. Where no confusion is possible, we will write scalar products in the space of factors as $f \cdot g$.

7.1.1 Short-Rate IR Component

Consider a general F-factor short-rate model that has the evolution (see Chapter 1)

$$r(t) = G(y(t)),$$

where

$$y(t) = \sum_{i=1}^{d} x_i(t) + \alpha(t). \tag{7.1.1}$$

Each auxiliary process obeys the Ornstein-Uhlenbeck equation

$$dx_i(t) = -\theta_i x_i(t)dt + \sigma_i(t) \cdot dW(t), \tag{7.1.2}$$

and the function $\alpha(t)$ is fixed to satisfy a given yield curve. Namely, for the risk-neutral measure, the savings account numeraire process

$$N(t) = e^{\int_0^t r(s)ds}$$

matches the yield curve discount factor: $\mathbb{E}[1/N(t)] = Y(t)$.

Let us introduce a currency label c for IR models; c can take the value d (domestic) or f (foreign). The hybrid model can contain multiple foreign currencies which we denote by f_1,f_2, etc. In this notation, the general short rate model $r_c(t) = G_c\left(\sum_{i=1}^{d} x_{c,i}(t) + \alpha_c(t)\right)$ satisfies the SDEs

$$dx_{c,i}(t) = -\theta_{c,i} x_{c,i}(t)dt + \sigma_{c,i}(t) \cdot dW_c(t). \tag{7.1.3}$$

Let $X_f(t)$ denote an exchange rate between a domestic model with rate $r_d(t)$ and a foreign model with rate $r_f(t)$. For the moment, the FX model can be Black-Scholes:

$$dX_f(t) = X_f(t)\left(r_d(t) - r_f(t)\right)dt + X_f(t)\,\sigma_{X_f}(t) \cdot dW_d(t). \tag{7.1.4}$$

Each interest rate's model can have several equity models $S_{e,c}$ based on it. Here, the subscript e is an equity label while the subscript c denotes a given equity currency. In the currency c risk-neutral measure, the equity SDE can be:

$$\text{BS} \quad \rightarrow \quad \frac{dS_{e,c}(t)}{S_{e,c}(t)} = \left(r_c(t) - d_e(t)\right)dt + \sigma_{e,c}(t) \cdot dW_c(t),$$

$$\text{Dupire} \quad \rightarrow \quad \frac{dS_{e,c}(t)}{S_{e,c}(t)} = \left(r_c(t) - d_e(t)\right)dt + \Sigma_{e,c}(t, S_{e,c}) \cdot dW_c(t),$$

$$\text{Heston} \quad \rightarrow \quad \frac{dS_{e,c}(t)}{S_{e,c}(t)} = \left(r_c(t) - d_e(t)\right)dt + \sqrt{z_{e,c}(t)}\sigma_{e,c}(t) \cdot dW_c(t),$$

$$dz_{e,c}(t) = a_{e,c}(t)\left(1 - z_{e,c}(t)\right)dt + \sqrt{z_{e,c}(t)}\,\gamma_{e,c}(t) \cdot dW_c(t),$$

where $d_e(t)$ is the equity-e continuous proportional dividend. Discrete dividends are covered in the beginning of Chapter 3.

To link the domestic IR, the foreign IR, their FX rate and the corresponding equity models, we fix the risk-neutral domestic measure, $W \equiv W_d$, and perform a foreign measure shift. Namely, if the FX-rate is defined by (7.1.4), the foreign Brownian motion dW_f will have a drift in the domestic measure:

$$dW_f(t) = dW(t) - \sigma_{X_f}(t)\, dt. \tag{7.1.5}$$

Thus, the domestic IR, the foreign IR and the FX rate will satisfy the following SDEs in the domestic measure:

$$\text{domestic short rate} \quad \rightarrow \quad r_d(t) = G_d\left(\sum_{i=1}^{d} x_{d,i}(t) + \alpha_d(t)\right),$$
$$dx_{d,i}(t) = -\theta_{d,i} x_{d,i}(t)dt + \sigma_{d,i}(t)\cdot dW(t),$$

$$\text{foreign short rate} \quad \rightarrow \quad r_f(t) = G_f\left(\sum_{i=1}^{d} x_{f,i}(t) + \alpha_f(t)\right),$$
$$dx_{f,i}(t) = -(\theta_{f,i} x_{f,i}(t) + \sigma_{X_f}(t)\cdot\sigma_{f,i}(t))dt + \sigma_{f,i}(t)\cdot dW(t),$$

$$\text{FX rate} \quad \rightarrow \quad dX_f(t) = X_f(t)\,(r_d(t) - r_f(t))\,dt + X_f(t)\,\sigma_{X_f}(t)\cdot dW(t).$$

The equity evolution remains unchanged for domestic equities and gets the measure shift (7.1.5) for foreign equities in the domestic measure, which is driven by the Brownian motion W. Below, we give a list of SDEs for the three *foreign* equity models $S_{e,f}(t)$ with their drift modification by the quanto correction (7.1.5):

$$\text{BS} \quad \rightarrow \quad \frac{dS_{e,f}(t)}{S_{e,f}(t)} = (r_f(t) - d_e(t) - \sigma_{e,f}(t)\cdot\sigma_{X_f}(t))\,dt + \sigma_{e,f}(t)\cdot dW(t),$$

$$\text{Dupire} \quad \rightarrow \quad \frac{dS_{e,f}(t)}{S_{e,f}(t)} = \left(r_f(t) - d_e(t) - \Sigma_{e,f}(t, S_{e,f})\cdot\sigma_{X_f}(t)\right)\,dt + \Sigma_{e,f}(t, S_{e,f})\cdot dW(t),$$

$$\text{Heston} \quad \rightarrow \quad \frac{dS_{e,f}(t)}{S_{e,f}(t)} = \left(r_f(t) - d_e(t) - \sqrt{z_{e,f}(t)}\sigma_{e,f}(t)\cdot\sigma_{X_f}(t)\right)\,dt + \sqrt{z_{e,f}(t)}\sigma_{e,f}(t)\cdot dW(t),$$
$$dz_{e,f}(t) = \left(a_{e,f}(t)\,(1 - z_{e,f}(t)) - \sqrt{z_{e,f}(t)}\,\gamma_{e,f}(t)\cdot\sigma_{X_f}(t)\right)dt$$
$$+ \sqrt{z_{e,f}(t)}\,\gamma_{e,f}(t)\cdot dW(t).$$

7.2 Correlation Conventions

In this section, we provide details on the correlation setup in the Numerix hybrid module. In the section above, we considered models that share multi-factor uncorrelated Brownian motion $B(t) = \{B_1(t), B_2(t), \dots\}$ where the *vector* volatilities give the correlations. To avoid misunder-

standing, we have adopted the letter "B" instead of the largely overloaded "W".

To separate inner model correlations from cross (outer) correlation between models, we have adapted the following structure. Let $s_i^{(m)}(t)$ denote the i-state process of the mth model. For example, the state process can be the Ornstein-Uhlenbeck process x_i for a multi-factor short-rate process. A state process $s_i^{(m)}(t)$ is driven by a dot product of the vector Brownian motion with its vector volatility $\sigma_i^{(m)}(t)$:

$$s_i^{(m)}(t) \quad \rightarrow \quad \sigma_i^{(m)}(t) \cdot dB(t).$$

This theoretically attractive form is equivalent to the more practical one dealing with the *correlated* Brownian motions (BM) $W_i^{(m)}(t)$:

$$\left\langle dW_i^{(m)}(t)\, dW_{i'}^{(m')}(t) \right\rangle = \frac{\sigma_i^{(m)}(t) \cdot \sigma_{i'}^{(m')}(t)}{|\sigma_i^{(m)}(t)|\, |\sigma_{i'}^{(m')}(t)|} dt.$$

Here $\sigma_i^{(m)}$ are the increments for states:

$$\sigma_i^{(m)}(t) \cdot dB(t) = \left|\sigma_i^{(m)}(t)\right| dW_i^{(m)}. \tag{7.2.1}$$

The volatility module becomes standard scalar volatility. We will refer to W as the *process* Brownian motion; i.e., $dW_i^{(m)}$ underlies the standard model process $s_i^{(m)}(t)$.

The *inner* model correlations

$$C_{i,i'}^{(m)}(t) = \left\langle dW_i^{(m)}(t)\, dW_{i'}^{(m)}(t) \right\rangle / dt,$$

which are correlations between different processes of the *same* model, are an *intrinsic* property of the model. The corresponding correlation matrix $C^{(m)}(t)$ has a model number superscript and its size is the number of different processes in the model. The inner correlations are chosen or calibrated without respect to other models in the hybrid.

Below, for better readability, we will adapt matrix/vector notations for correlation matrices. We will write

$$C^{(m)}(t) = \left\langle dW^{(m)}(t)\, dW^{(m)^T}(t) \right\rangle / dt$$

instead of the element-by-element form. (Here dW is considered as a column-vector while dW^T is a row-vector.)

Consider now a *process* correlation matrix, $C_{pr}(t)$, i.e., a correlation matrix between all pro-

cess BMs that correspond to the same or different models. We can represent it in block form:

$$C_{pr}(t) = \left(\begin{array}{c|c|c|c} C^{(1)}(t) & C^{(1,2)}(t) & \cdots & C^{(1,M)}(t) \\ \hline C^{(2,1)}(t) & C^{(2)}(t) & \cdots & C^{(2,M)}(t) \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline C^{(M,1)}(t) & C^{(M,2)}(t) & \cdots & C^{(M)}(t) \end{array}\right).$$

The diagonal blocks $C^{(m)}(t)$ contain the inner-correlation matrices inside the model m. The off-diagonal blocks represent correlation matrices between different models and are given by

$$C^{(m,m')}(t) = \left\langle dW^{(m)}(t)\, dW^{(m')^T}(t) \right\rangle / dt.$$

To separate the correlations between different models and inner model correlations, we supply each model with uncorrelated Brownian motions $Z^{(m)} = \{Z_1^{(m)}, \ldots, Z_{F_m}^{(m)}\}$, where F_m is the number of factors of the mth model. As the Brownian motions are uncorrelated, we have

$$\left\langle dZ_i^{(m)}(t)\, dZ_{i'}^{(m)}(t) \right\rangle = 0 \quad \text{for} \quad i \neq i'. \tag{7.2.2}$$

Taking a *generalized square root*[2] $U^{(m)}(t)$ of the inner-correlation matrix,

$$C^{(m)}(t) = U^{(m)}(t)\, U^{(m)}(t)^T, \tag{7.2.3}$$

one can link the Brownian motions

$$dW^{(m)} = U^{(m)}(t)\, dZ^{(m)}. \tag{7.2.4}$$

[2]For a symmetric positive semi-definite (SPSD) matrix C there exists a unique (up to trivial sign flip) SPSD *square root* matrix $U_0 = \sqrt{C}$, which solves both $C = U_0 U_0$ and $C = U_0 U_0^T$. Any matrix $U = U_0 Q$, where Q is an arbitrary orthogonal matrix, is a *generalized square root* of C, which solves $C = U U^T$.

We can represent dW in a block form:

$$\begin{pmatrix} dW^{(1)} \\ \hline dW^{(2)} \\ \hline \cdots \\ \hline dW^{(M)} \end{pmatrix} = \left(\begin{array}{c|c|c|c} U^{(1)}(t) & 0 & \cdots & 0 \\ \hline 0 & U^{(2)}(t) & \cdots & 0 \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline 0 & 0 & \cdots & U^{(M)}(t) \end{array}\right) \begin{pmatrix} dZ^{(1)} \\ \hline dZ^{(2)} \\ \hline \cdots \\ \hline dZ^{(M)} \end{pmatrix}.$$

The transition matrix is block diagonal:

$$A(t) = \left(\begin{array}{c|c|c|c} U^{(1)}(t) & 0 & \cdots & 0 \\ \hline 0 & U^{(2)}(t) & \cdots & 0 \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline 0 & 0 & \cdots & U^{(M)}(t) \end{array}\right).$$

Using block vector notation for both BMs:

$$dW \equiv \left(\begin{array}{c} dW^{(1)} \\ \hline dW^{(2)} \\ \hline \cdots \\ \hline dW^{(M)} \end{array} \right)$$

and

$$dZ \equiv \left(\begin{array}{c} dZ^{(1)} \\ \hline dZ^{(2)} \\ \hline \cdots \\ \hline dZ^{(M)} \end{array} \right),$$

we obtain the compact transition relation

$$dW(t) = A(t)\, dZ(t).$$

The square-root of the correlation matrix $C^{(m)}$ is taken differently for different models:

- Short-rate IR models
  The square root is a symmetric matrix. For example, for two-factor case with a correlation $\rho(t)$ between two process Brownian motion, dW_1 and dW_2, the variable transformation is

given by

$$dW_1(t) = dZ_1(t)\, a(t) + dZ_2(t)\, b(t), \tag{7.2.5}$$
$$dW_2(t) = dZ_1(t)\, b(t) + dZ_2(t)\, a(t), \tag{7.2.6}$$

for $a(t) = \left(\sqrt{1+\rho(t)} + \sqrt{1-\rho(t)}\right)/2$ and $b(t) = \left(\sqrt{1+\rho(t)} - \sqrt{1-\rho(t)}\right)/2$.

- Heston model
  For Heston correlation $\rho(t)$, the square root is an asymmetric transformation:

$$dW_1(t) = dZ_1(t), \tag{7.2.7}$$
$$dW_2(t) = dZ_1(t)\, \rho(t) + dZ_2(t)\, \sqrt{1-\rho^2(t)}. \tag{7.2.8}$$

  In correlated Brownian motion terms, the Heston evolution looks is

$$\frac{dS(t)}{S(t)} = (r_c(t) - d_e(t))\, dt + \sqrt{z(t)}\sigma(t)\, dW_1(t), \tag{7.2.9}$$
$$dz(t) = a(t)\,(1 - z(t))\, dt + \sqrt{z(t)}\, \gamma(t)\, dW_2(t), \tag{7.2.10}$$

  where $\sigma(t)$ and $\gamma(t)$ are scalar volatilities.

Thus, we define *cross* (or *outer*) correlations as the correlations between the vector Brownian factors for different models, $Z^{(1)}, Z^{(2)}, \ldots$,:

$$C_{cr}(t) = \left\langle dZ(t)\, dZ^T(t) \right\rangle / dt.$$

In other words, the cross correlation matrix can be presented in a block form:

$$C_{cr}(t) = \left(\begin{array}{c|c|c|c} I & \rho^{(1,2)}(t) & \cdots & \rho^{(1,M)}(t) \\ \hline \rho^{(2,1)}(t) & I & \cdots & \rho^{(2,M)}(t) \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline \rho^{(M,1)}(t) & \rho^{(M,2)}(t) & \cdots & I \end{array}\right),$$

where the off-diagonal blocks are

$$\rho^{(m,m')}(t) = \left\langle dZ^{(m)}(t)\, dZ^{(m')^T}(t) \right\rangle / dt.$$

Note that unit diagonal blocks ($\rho^{(m,m)}(t) = I$) reflect the mutual independence of inner-model Brownian motions (7.2.2). Obviously, the corresponding matrices are related by a transposition

$$\rho^{(m,m')}(t) = \rho^{(m',m)}(t)^T.$$

For example, a hybrid model containing

- Model 1 (2-factor) with Brownian factors $\{Z_1^{(1)}, Z_2^{(1)}\}$,
- Model 2 (3-factor) with Brownian factors $\{Z_1^{(2)}, Z_2^{(2)}, Z_3^{(2)}\}$, and
- Model 3 (1-factor) with Brownian factor $\{Z_1^{(3)}\}$

should be equipped with the correlation matrix

$$C_{cr} = \left(\begin{array}{cc|ccc|c} 1 & 0 & \rho_{11}^{(1,2)} & \rho_{12}^{(1,2)} & \rho_{13}^{(1,2)} & \rho_{11}^{(1,3)} \\ 0 & 1 & \rho_{21}^{(1,2)} & \rho_{22}^{(1,2)} & \rho_{23}^{(1,2)} & \rho_{12}^{(1,3)} \\ \hline \rho_{11}^{(1,2)} & \rho_{21}^{(1,2)} & 1 & 0 & 0 & \rho_{11}^{(2,3)} \\ \rho_{12}^{(1,2)} & \rho_{22}^{(1,2)} & 0 & 1 & 0 & \rho_{12}^{(2,3)} \\ \rho_{13}^{(1,2)} & \rho_{23}^{(1,2)} & 0 & 0 & 1 & \rho_{13}^{(2,3)} \\ \hline \rho_{11}^{(1,3)} & \rho_{12}^{(1,3)} & \rho_{11}^{(2,3)} & \rho_{12}^{(2,3)} & \rho_{13}^{(2,3)} & 1 \end{array}\right),$$

where the correlation between $Z_i^{(m)}$ (the ith component of the Brownian motion of the mth model) and $Z_{i'}^{(m')}$ (the i'th component of the Brownian motion of the m'th model) is denoted by $\rho_{ii'}^{(m,m')}$.

We can easily represent the *process* correlation matrix in terms of the transform A and the cross correlation matrix as

$$C_{pr}(t) = A(t)\, C_{cr}(t)\, A^T(t),$$

or in terms of its block elements,

$$C_{pr}^{(m,m')}(t) = U^{(m)}(t)\, \rho^{(m,m')}(t)\, U^{(m')}(t)^T. \tag{7.2.11}$$

Inversely, one can obtain the cross correlation matrix from the process correlation matrix,

$$\rho^{(m,m')}(t) = \left\{U^{(m)}(t)\right\}^{-1} C_{pr}^{(m,m')}(t) \left\{U^{(m')}(t)^T\right\}^{-1}. \tag{7.2.12}$$

when the inner correlation matrices $C^{(m)}(t)$ (and consequently, their square-roots $U^{(m)}(t)$) are *invertible*. Obviously, his property does not hold for degenerate inner-correlation matrices (with unit correlations between certain factors).

For the moment,[3] the correlation input to a hybrid model coincides with the cross correlation matrix C_{cr}.

[3] We plan to enlarge the acceptable input matrix.

The formal procedure of correlation construction consists of the following steps:

- Calibrate or set the inner-correlation matrix for individual models ($C^{(m)}$ for the mth model).
- Set the cross-correlation matrix C_{cr} as an input that determines correlations between model factors $Z^{(1)}, Z^{(2)}, \ldots$.
- Transform[4] the mutually independent factors $Z^{(m)}$ into the correlated process Brownian motions $W^{(m)}$ for each individual model by (7.2.4).
- Set up the hybrid model according to its SDEs.

Below, we give correlation structure examples for various setups.

7.2.1 LMM/BS/LMM

Here, we give an example of a setup for a the correlation matrix for a two-factor domestic ($F = 2$) and three-factor foreign ($\tilde{F} = 3$) model). Using the correlated Brownian motions (7.2.1), we can rewrite our initial vector volatility structure (A.1.9)–(A.1.11) in terms of standard volatilities,

$$dL_n(t) = (L_n(t) + s_n)\left(\sum_{j=\eta(t)}^{n} \frac{\delta_j\,(L_j(t) + s_j)}{1 + \delta_j\,L_j(t)}\,\gamma_n(t)\cdot\gamma_j(t)\right)dt + (L_n(t) + s_n)\,\gamma_n(t)\cdot dZ(t), \tag{7.2.13}$$

$$d\tilde{L}_n(t) = (\tilde{L}_n(t) + \tilde{s}_n)\left(\sum_{j=\tilde{\eta}(t)}^{n} \frac{\tilde{\delta}_j\,(\tilde{L}_j(t) + \tilde{s}_j)}{1 + \tilde{\delta}_j\,\tilde{L}_j(t)}\,\tilde{\gamma}_n(t)\cdot\tilde{\gamma}_j(t) - \gamma_n(t)\cdot\tilde{\rho}_X\,\sigma_X(t)\right)dt + (\tilde{L}_n(t) + \tilde{s}_n)\,\tilde{\gamma}_n(t)\cdot d\tilde{Z}(t), \tag{7.2.14}$$

$$dY(t) = Y(t)\,\sigma_X(t)dZ_X(t). \tag{7.2.15}$$

We use $\tilde{}$ (tilde) for foreign market quantities instead of a currency-label subscript for better legibility. Now, the domestic market is driven with the mutually independent collection of vector Brownian motions $Z \equiv Z^{(d)}$ such that $\lambda_n(t)\cdot dW(t) = \gamma_n(t)\cdot dZ(t)$. Analogously, the foreign market Brownian motion $\tilde{Z} \equiv Z^{(f)}$ satisfies $\tilde{\lambda}_n(t)\cdot dW(t) = \tilde{\gamma}_n(t)\cdot d\tilde{Z}(t)$. In the same way, the FX-rate Brownian motion Z_X corresponds to $\sigma_Y(t)\cdot dW(t) = \sigma_X(t)dZ_X(t)$ with scalar volatility $\sigma_X(t)$.

In spite of formal similarities, one should distinguish the six-dimensional initial quantities—the vector Brownian motion $W(t)$, vector volatilities $\lambda_n(t)$, $\tilde{\lambda}_n(t)$, and $\sigma_Y(t)$—from the two-dimensional domestic vector Brownian motion $Z(t)$ with volatilities $\gamma_n(t)$ and the three-dimensional foreign vector Brownian motion $\tilde{Z}(t)$ with volatilities $\tilde{\gamma}_n(t)$.

The general correlation structure between:

- the domestic and foreign Brownian motions:

$$\mathbb{E}\left[dZ_f\,d\tilde{Z}_{f'}\right] = \rho_{ff'}\,dt \text{ for any } 0 \le f \le F-1 \text{ and } 0 \le f' \le \tilde{F}-1,$$

- the FX-rate Brownian motion and domestic Brownian motion:

$$\mathbb{E}\left[dZ_f\,dZ_X\right] = \rho_{X\,f}\,dt \text{ for any } 0 \le f \le F-1,$$

[4] For LMM, this step is omitted, since the model requires mutually independent Brownian motions $Z^{(m)}$.

- the FX-rate Brownian motion and foreign Brownian motion:

$$E[d\tilde{Z}_f \, dZ_X] = \tilde{\rho}_{X\,f} \, dt \text{ for any } 0 \leq f \leq \tilde{F} - 1,$$

corresponds to the input correlation matrix

$$C = \left(\begin{array}{cc|ccc|c} 1 & 0 & \rho_{00} & \rho_{01} & \rho_{02} & \rho_{X\,0} \\ 0 & 1 & \rho_{10} & \rho_{11} & \rho_{12} & \rho_{X\,1} \\ \hline \rho_{00} & \rho_{10} & 1 & 0 & 0 & \tilde{\rho}_{X\,0} \\ \rho_{01} & \rho_{11} & 0 & 1 & 0 & \tilde{\rho}_{X\,1} \\ \rho_{02} & \rho_{12} & 0 & 0 & 1 & \tilde{\rho}_{X\,2} \\ \hline \rho_{X\,0} & \rho_{X\,1} & \tilde{\rho}_{X\,0} & \tilde{\rho}_{X\,1} & \tilde{\rho}_{X\,2} & 1 \end{array}\right).$$

For completeness, note that the *correlated* Brownian motions U can be expressed as

$$|\gamma_n| \, dU_n(t) = \gamma_n(t) \cdot dZ(t)$$

for the nth domestic Libor, and similarly for the other states.

7.2.2 Heston Basket with Hull-White

Here, we consider a correlation structure example of two Heston models for equities coupled with a Hull-White domestic interest rate defined as follows:

- Heston models[5] with processes $S^{(i)}$ for $i = 1, 2$ and parameters:
  - volatility of equity $\sigma^{(i)}$ for $i = 1, 2$
  - correlation between equity and corresponding stochastic volatility $\rho^{(i)}$ for $i = 1, 2$
  - stochastic volatility mean reversion $a^{(i)}$ for $i = 1, 2$
  - volatility of volatility $\gamma^{(i)}$ for $i = 1, 2$
  - dividend yield $d^{(i)}(t)$ for $i = 1, 2$
- HW1F model with short rate $r(t)$ and parameters:
  - volatility $\lambda(t)$
  - mean reversion $\theta(t)$

Each Heston model has a mutually independent two-dimensional Brownian motion $Z^{(i)}$ for $i = 1, 2$ (so $\langle dZ_1^{(i)} dZ_2^{(i)} \rangle = 0$. We also define the HW one-dimensional Brownian motion $Z^{(3)}$. The input

[5] The superscript corresponds to Heston model index $i = 1, 2$.

correlation matrix

$$C_{cr} = \left(\begin{array}{cc|cc|c} 1 & 0 & \rho_{11}^{(1,2)} & \rho_{12}^{(1,2)} & \rho_{11}^{(1,3)} \\ 0 & 1 & \rho_{21}^{(1,2)} & \rho_{22}^{(1,2)} & \rho_{12}^{(1,3)} \\ \hline \rho_{11}^{(1,2)} & \rho_{21}^{(1,2)} & 1 & 0 & \rho_{11}^{(2,3)} \\ \rho_{12}^{(1,2)} & \rho_{22}^{(1,2)} & 0 & 1 & \rho_{12}^{(2,3)} \\ \hline \rho_{11}^{(1,3)} & \rho_{12}^{(1,3)} & \rho_{11}^{(2,3)} & \rho_{12}^{(2,3)} & 1 \end{array}\right)$$

defines the correlations between $Z_1^{(1)}$, $Z_2^{(1)}$, $Z_1^{(2)}$, $Z_2^{(2)}$, $Z^{(3)}$. We have the following SDEs:

$$\frac{dS^{(i)}(t)}{S^{(i)}(t)} = (r(t) - d^{(i)}(t))\,dt + \sqrt{z^{(i)}(t)}\sigma^{(i)}(t)\,dZ_1^{(i)}(t), \tag{7.2.16}$$

$$dz^{(i)}(t) = a^{(i)}(t)\,(1 - z^{(i)}(t))\,dt + \sqrt{z^{(i)}(t)}\,\gamma^{(i)}(t)\,\left(dZ_1^{(i)}\rho^{(i)}(t) + dZ_2^{(i)}\sqrt{1 - (\rho^{(i)}(t))^2}\right), \tag{7.2.17}$$

$$dr(t) = (\zeta(t) - \theta(t)r(t))dt + \lambda(t)dZ^{(3)}(t), \tag{7.2.18}$$

where $\zeta(t)$ is the yield curve compensator.

7.3 Calibration

In this section, we consider calibration issues. Any hybrid model can be calibrated to appropriate European options by brute-force Monte Carlo. At the same time, the following large class of models is covered with analytical calibration:

- Calibration of any combination of Gaussian models, multi-factor HW for IR, BS for FX or equity, to European FX- or equity-options.
- Cross-currency model calibration to FX-options:
  - SBK/HW for interest rates, BS for FX-rate,
  - LMM for interest rates, BS for FX-rate.
- Hybrid equity-IR model calibration to equity-options:
  - Heston for equity, HW for interest rates.
- Calibration[6] to equity basket index options (with deterministic interest rates):
  - BS for basket equities,
  - Heston for basket equities.

The calibration procedure is organized in a universal and flexible way using efficient iterative solving given the analytical (or numerical for brute force calibration) way of European option calculation. All the models listed above have European analytics implemented. The analytical methods are described below.

[6]For the moment, only analytical approximation for index options is available. The calibration procedure will be implemented in the next release.

7.3.1 European FX-Option Analytics for IR-LMM_FX-BS_IR-LMM Model

In [9] and [10], the authors uses Markovian projection to calculate the price of a European FX option with a hybrid model that uses an LMM IR model for both currencies and a Black Scholes FX model for the FX rate between the two currencies (i.e., IR-LMM_FX-BS_IR-LMM). Here, we describe briefly the method, omitting technical details which can be found in the cited papers. To compute the option price with maturity T_M and strike K, we change the measure to the T_M-forward measure with the domestic bond $P(t,T_M)$ as a numeraire. The forward FX rate

$$F(t,T_M) = \frac{X(t)\,\tilde{P}(t,T_M)}{P(t,T_M)} \tag{7.3.1}$$

is a martingale under the T_M-forward measure. The option price reduces to an expectation under this measure and is given by

$$\mathbb{E}\left[\frac{\left(X(T_M)-K\right)^+}{N(T_M)}\right] = P(0,T_M)E_{T_M}\Big[\left(F(T_M,T_M)-K\right)^+\Big]. \tag{7.3.2}$$

The forward FX-rate process $F(\cdot,\cdot)$ can be expressed in terms of the bond ratios, $R_n(t)$ and $\tilde{R}_n(t)$, which are given by

$$\begin{aligned} R_n(t) &= \frac{P(t,T_n)}{P(t,T_{n+1})} = 1+\delta_n\,L_n(t), \\ \tilde{R}_n(t) &= \frac{\tilde{P}(t,T_n)}{\tilde{P}(t,T_{n+1})} = 1+\delta_n\,\tilde{L}_n(t), \end{aligned} \tag{7.3.3}$$

and the process $Y(t)$ from (A.1.11):

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

Below, we assume that the maturity T_M of the forward FX rate is fixed, and omit the corresponding argument, writing $F(t)$ instead of $F(t,T_M)$. Note that we know without calculation that the drift terms will cancel because we now work in a martingale measure for $F(t)$. An SDE for the forward process can be obtained by Ito's lemma in the T-forward measure:

$$dF(t,T_M) = F(t,T_M)\Lambda(t)dW(t),$$

where $\Lambda(t)$ is a non-Markovian process described in [9].

The forward-rate process is obviously non-Markovian. The Markovian projection logic is to replace it with an *effective* Markovian process $F^*(t)$ that follows a time-dependent displaced diffusion,

$$dF^*(t) = \big(F^*(t)\beta(t)+(1-\beta(t))F^*(0)\big)\,\sigma(t)dW(t), \qquad F^*(0)=F(0,T_M), \tag{7.3.4}$$

such that its marginal distributions well approximate those of the initial true process $F(t)$. The optimal skew $\beta(t)$ and volatility $\sigma(t)$ were explicitly calculated in [9]. Once this is done, we invoke

the formula for shift averaging derived by Piterbarg (2005),

$$\bar{\beta}_{T_M} = \frac{\int_0^{T_M} \beta(t)\sigma^2(t) \int_0^t \sigma^2(\tau) d\tau dt}{\int_0^{T_M} \sigma^2(t) \int_0^t \sigma^2(\tau) d\tau dt}, \tag{7.3.5}$$

which reduces the pricing of an FX option with maturity T_M to a simple modification of the Black-Scholes formula:

$$\mathbb{E}\left[\frac{(X(T_M)-K)^+}{N(T_M)}\right] = P(0,T_M)\left(\frac{F}{\bar{\beta}_{T_M}}\mathcal{N}(d_+) - \left(K + \frac{F(1-\bar{\beta}_{T_M})}{\bar{\beta}_{T_M}}\right)\mathcal{N}(d_-)\right), \tag{7.3.6}$$

$$d_\pm = \frac{\ln F/(K\bar{\beta}_{T_M} + F(1-\bar{\beta}_{T_M})) \pm V/2}{\sqrt{V}}, \quad V = \bar{\beta}_{T_M}^2 \int_0^{T_M} |\sigma(t)|^2 dt, \tag{7.3.7}$$

where $F = F(0,T_M)$ is the forward rate, and $\bar{\beta}_{T_M}$ is the effective shift obtained by applying the averaging formula (7.3.5) to the time-dependent shift.

7.3.2 FX or Equity Option for Gaussian Interest Rates

In this subsection, we consider pricing FX or equity options using BS/Heston models coupled with HW multi-factor interest rates. We will treat the general case of a one-factor HW with a Heston model for the exchange rate. The other models listed below can be related to the general setup as follows:

- The equity case corresponds to zero foreign volatility in the FX setup.
- The BS exchange rate is the Heston model degeneration with zero vol-of-vol.

Consider a Hull-White model for domestic and foreign interest rates. Take, for simplicity, the one-factor case[7] for domestic and foreign short rates (denoted $r_d(t)$ and $r_f(t)$, respectively). In the respective risk-neutral measures, the Hull-White SDEs are of the form

$$dr_c(t) = (\zeta_c(t) - r_c(t)\, a_c(t))\, dt + \sigma_c(t) \cdot dW_c(t), \tag{7.3.8}$$

where the currency label $c = d, f$ corresponds to a market number (domestic/foreign) and the parameters a_c and σ_c are the time-dependent mean-reversion and volatility. The parameter ζ_c is the usual yield curve adaptor, and W_d (respectively W_f) is the risk-neutral domestic (respectively, risk-neutral foreign) vector Brownian motion, mutually uncorrelated as usual. For further applications, we introduce a zero-coupon bond $P_c(t,T)$ that matures at T and satisfies standard lognormal evolution

$$\frac{dP_c(t,T)}{P_c(t,T)} = r_c(t)\, dt - \Sigma_c(t,T)\, dW_c(t),$$

Here

$$\Sigma_c(t,T) = \sigma_c(t) \int_t^T e^{-\int_t^\tau a_c(s)\, ds} d\tau.$$

are the zero-coupon bond volatilities.

[7] The multi-factor case can be treated in the same manner.

Let $X(t)$ be the FX rate between domestic and foreign markets following a Heston process. Then, the hybrid model under the domestic Brownian motion $W^* = W_d$ will have the SDEs

$$dr_d(t) = (\zeta_d(t) - r_d(t)\, a_d(t))\, dt + \sigma_d(t) \cdot dW^*(t), \tag{7.3.9}$$

$$dr_f(t) = (\zeta_f(t) - r_f(t)\, a_f(t) - \sqrt{z(t)}\, \lambda(t) \cdot \sigma_f(t))\, dt + \sigma_f(t) \cdot dW^*(t), \tag{7.3.10}$$

$$dX(t) = X(t)(r_d(t) - r_f(t))\, dt + X(t)\sqrt{z(t)}\, \lambda(t)\ \cdot dW^*(t), \tag{7.3.11}$$

$$dz(t) = \alpha(t)\,(1 - z(t))\, dt + \sqrt{z(t)}\, \gamma(t)\ \cdot dW^*(t), \qquad z(0) = 1. \tag{7.3.12}$$

Correlations in the model are introduced via vector volatilities.

The price of a European option with strike K and maturity T under the risk neutral-measure is given by

$$\mathbb{E}\left[\frac{(X(T) - K)^+}{N_d(t)}\right],$$

where the domestic savings account is $N_d(t) = \exp\left(\int_0^t r_d(s)ds\right)$. As usual, we can represent the price in a more convenient domestic T-forward measure, which is associated with domestic zero-coupon bond $P_d(t,T)$ that matures at T:

$$\mathbb{E}\left[\frac{(X(T) - K)^+}{N(t)}\right] = P_d(0,T)\mathbb{E}_T\left[(X(T) - K)^+\right].$$

Introducing the forward FX-rate

$$S(t) = \frac{X(t)\, P_f(t,T)}{P_d(t,T)},$$

which is a martingale process under the domestic T-forward measure, we can further simplify the option price:

$$\mathbb{E}\left[\frac{(X(T) - K)^+}{N_d(t)}\right] = P(0,T)\, E_T\left[(S(T) - K)^+\right].$$

Given domestic and foreign zero-coupon bond volatilities $\Sigma_c(t,T)$, one can easily construct the SDE for the forward $S(t)$:

$$dS(t) = S(t)\left(\sqrt{z_T(t)}\, \lambda(t) + \Sigma_d(t,T) - \Sigma_f(t,T)\right) \cdot dW_T. \tag{7.3.13}$$

Here W_T is T-forward Brownian motion, which is related to risk-neutral motion by $dW_T(t) = dW^*(t) + \Sigma_d(t,T)\, dt$, and z_T is the stochastic variance in the T-forward measure and satisfies the SDE

$$dz_T(t) = \left(\alpha(t)\,(1 - z_T(t)) - \Sigma_d(t,T) \cdot \gamma(t)\,\sqrt{z_T(t)}\right)\, dt + \sqrt{z_T(t)}\, \gamma(t)\ \cdot dW_T(t), \qquad z(0) = 1.$$

Below, we consider separately the Black-Scholes case (for zero vol-of-vol) and the much more complicated Heston case.

Black-Scholes Case

The BS FX rate can be obtained from the Heston model by setting its vol-of-vol, $\gamma(t)$, to zero so that the stochastic volatility multiplier becomes a constant ($z(t) = 1$). Thus, the forward rate

SDE (7.3.13) degenerates to lognormal:

$$dS(t) = S(t)\left(\lambda(t) + \Sigma_d(t,T) - \Sigma_f(t,T)\right) \cdot dW_T.$$

This immediately gives the implied BS volatility for the European option with maturity T and any strike as

$$\sigma_I(T) = \int_0^T \left|\lambda(t) + \Sigma_d(t,T) - \Sigma_f(t,T)\right|^2 dt,$$

leading to a flat smile.

Heston Case

To calculate option prices for Heston models, one should approximate the process (7.3.13). This was done using Markovian projection in [7], wherein the authors used a displaced volatility Heston model as the projection target. Technical details can be found in [7].

References

[1] Andersen, L. B. Efficient simulation of the heston stochastic volatility model. http://ssrn.com/abstract=946405, 2007. Available at SSRN. Accessed 2014-10-03.

[2] Andersen, L. B., and Andreasen, J. Volatile volatilities. *Risk Magazine 15*, 12 (December 2002), 163–168.

[3] Andersen, L. B., and Buffum, D. Calibration and implementation of convertible bond models. *Journal of Computational Finance 7* (2004), 1–34.

[4] Andersen, L. B., and Piterbarg, V. Atlantic Financial Press, 2010.

[5] Andersen, L. B., and Sidenius, J. Extensions to the gaussian copula: random recovery and random factor loadings. *Journal of Credit Risk 1*, 1 (2005), 29–70.

[6] Antonov, A., and Arneguy, M. Analytical formulas for pricing cms products in the lmm model. http://ssrn.com/abstract=1352606, 2009. Accessed 2014-10-04.

[7] Antonov, A., Arneguy, M., and Audet, N. Markovian projection to a displaced volatility heston model. http://ssrn.com/abstract=1106223, 2008. Accessed 2014-10-04.

[8] Antonov, A., Mechkov, S., and Misirpashaev, T. Analytical techniques for synthetic cdos and credit default risk measures, 2005. Working paper, available at www.defaultrisk.com.

[9] Antonov, A., and Misirpashaev, T. Efficient calibration to fx options by Markovian projection in cross-currency libor market models. http://ssrn.com/abstract=936087, 2006. Accessed 2014-10-04.

[10] Antonov, A., and Misirpashaev, T. Markovian projection onto a displaced diffusion: Generic formulas with applications. http://ssrn.com/abstract=937860, 2006. Accessed 2014-10-04.

[11] Antonov, A., and Misirpashaev, T. Projection on a quadratic model by asymptotic expansion with an application to lmm swaption. http://ssrn.com/abstract=1421139, 2009. Accessed 2014-10-04.

[12] Antonov, A., Misirpashaev, T., and Piterbarg, V. Markovian projection onto a heston model. http://ssrn.com/abstract=997001, 2007. Accessed 2014-10-04.

[13] Antonov, A., and Spector, M. Analytical approximations for short rate models. http://ssrn.com/abstract=1692624, 2010. Accessed 2014-10-04.

[14] Antonov, A., and Spector, M. General short-rate analytics. *Risk Magazine* (May 2011), 66–71.

[15] Arnsdorf, M., and Halperin, I. Bslp: Markovian bivariate spread-loss model for portfolio credit derivatives, 2007. Available at defaultrisk.com.

[16] Avellaneda, M., Boyer-Olson, D., Busca, J., and Friz, P. Reconstructing volatility. *Risk Magazine* (October 2002), 87–91.

[17] Balduzzi, P., Das, S. R., and Foresi, S. The central tendency: A second factor in bond yields, 1996. NBER working paper no. 6325.

[18] Bates, D. S. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *REVIEW OF FINANCIAL STUDIES 9*, 1 (1996). Available at SSRN: http://ssrn.com/abstract=6913.

[19] Belgrade, N., Benhamou, E., and Koehler, E. A market model for inflation. http://ssrn.com/abstract=576081, 2004. Accessed 2014-10-04.

[20] Brace, A. *Engineering BGM*. CRC Financial Mathematics Series. Chapman & Hall, Boca Raton, 2007.

[21] Brace, A., Gatarek, D., and Musiela, M. The market model of interest rate dynamics. *Mathematical Finance 7* (1996), 127–154.

[22] Brigo, D., and Mercurio, F. *Interest Rate Models - Theory and Practice: With Smile, Inflation and Credit*, 2nd ed. Springer, Berlin, 2006.

[23] Brigo, D., Pallavicini, A., and Torresetti, R. Calibration of cdo tranches with the dynamical generalized-poisson loss model, 2006. Working paper. Available at defaultrisk.com.

[24] Cox, J. C., Ingersoll, J. E., and Ross, S. A theory of the term structure of interest rates. *Econometrica 2* (1985), 385–407.

[25] Davis, M., and Lo, V. Infectious defaults. *Quantitative Finance 1* (2001), 382–387.

[26] Dragulescu, A., and Yakovenko, V. Probability distribution of returns in the heston model with stochastic volatility. *Quantitative Finance 2* (2002), 443–453.

[27] Duffie, D., and Singleton, K. Modeling term structure of defaultable bonds. *Review of Financial Studies 12* (1999), 687–720.

[28] Errais, E., Giesecke, K., and Goldberg, L. Pricing credit from the top down with affine point processes, 2006. Working paper, available at defaultrisk.com.

[29] Feller, W. Two singular diffusion problems. *The Annals of Mathematics, Second Series 54*, 1 (July 1951), 173–182.

[30] Gatheral, J. *The Volatility Surface: A Practitioner's Guide*. Wiley Finance, 2006.

[31] Geman, H. *Commodities and commodity derivatives*. John Wiley & Sons, Ltd., London, 2005.

[32] Hagan, P. S., Kumar, D., Lesniewski, A. S., and Woodward, D. E. Managing smile risk. *Wilmott Magazine 1*.

[33] Halperin, I. Using Markov chain models for pricing credit sensitive derivatives.

[34] Heston, S. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies 6*, 2 (1993).

[35] Hull, J., Predescu, M., and White, A. The valuation of correlation dependent credit derivatives using a structural model, 2005. Working paper, University of Toronto. Available at defaultrisk.com.

[36] Hull, J. C. *Options, Futures and Other Derivatives*, 8th ed. Prentice Hall, January 2011.

[37] Jarrow, R., Lando, D., and Turnbull, S. A Markov model of the term structure of credit risk spread.

[38] Kazziha, S. *Interest Rate Models, Inflation-based derivatives, Trigger Notes And Cross-Currency Swaptions*. PhD thesis, Imperial College of Science, Technology, and Medicine, 1999.

[39] Kenyon, C. Inflation is normal. *Risk Magazine* (July 2008).

[40] Kijima, M., and Komoribayashi, K. A Markov chain model for valuing credit risk derivatives. *The Journal of Derivatives 6*, 1 (1998), 97.

[41] Kou, S. G., and Wang, H. Option pricing under a double exponential jump diffusion model. http://ssrn.com/abstract=284202, 2001. Accessed 2014-10-05.

[42] Lewis, A. *Option Valuation under Stochastic Volatility*. Finance Press, Newport Beach, California, 2005.

[43] Li, D. X. A copula function approach. *The Journal of Fixed Income 9*, 4 (2000), 43–54.

[44] Longstaff, F. A., and Schwartz, E. S. Interest rate volatility and the term structure: A two-factor general equilibrium model. *Journal of Finance 47*, 4, 1259–1282.

[45] Lopatin, A. V. A simple dynamic model for pricing and hedging heterogenous cdos, 2008. Available at defaultrisk.com.

[46] Lopatin, A. V., and Misirpashaev, T. Two-dimensional Markovian model for dynamics of aggregate credit loss. *Advances in Economics 22* (2007).

[47] Maghsoodi, Y. Solution of the extended Cox, Ingersoll, and Ross model of the term structure and analytical bond option valuation. *Mathematical Finance 6* (1996), 89–109.

[48] Mercurio, F. Pricing inflation-indexed derivatives. *Quantitative Finance* (2005), 289–302.

[49] Mercurio, F., and Moren, N. Inflation with a smile. *Risk Magazine 19*, 3 (2006), 70–75.

[50] Mercurio, F., and Moreni, N. Inflation modeling with SABR dynamics. *Risk Magazine* (June 2009), 106–111.

[51] Merton, R. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics 3* (1976), 125–144.

[52] Mikkelsen, P. Cross-currency Libor market models. http://old-hha.asb.dk/bs/wp/fin/glfin/D01_6.pdf, 2001. Accessed 2014-10-05.

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

[54] Piterbarg, V. Time to smile. *Risk Magazine* (May 2005), 71–75.

[55] Piterbarg, V. Markovian projection method for volatility calibration. http://ssrn.com/abstract=906473, 2006. Accessed 2014-10-05.

[56] Piterbarg, V. Markovian projection for volatility calibration, April 2007.

[57] R. Bos, A. Gairat, A. S. Dealing with discrete dividends. risk. *Risk Magazine* (January 2003), 109–112.

[58] R. Jarrow, Y. Y. Pricing treasury inflation protected securities and related derivatives using an hjm model. *Journal of Financial and Quantitative Analysis 38*, 2 (2003), 409–430.

[59] Schlögl, E. A multicurrency extension of the lognormal interest rate market models. *Finance and Stochastics 6*, 2 (2002), 173–196.

Appendix A

Models to be Deprecated

This section covers the models that either have been deprecated or are marked for deprecation at some point in the future.

A.1 Cross-Currency LMM

In this section, we describe the cross-currency LMM model. The CC LMM model supports IR LMM in both domestic and foreign currencies. The supported combination of models is LMM/BS/LMM. Combining LMM with other types of IR models in a cross-currency setup is not allowed.

Extending the LMM framework to cross-currency instruments is conceptually transparent and was addressed in [52] and [59].

As in the previous section, we consider the multi-factor uncorrelated vector Brownian motion $W(t) = \{W_1(t), W_2(t), \dots\}$ that drives all the processes, and comment on correlation structure below.

Recall the basic facts of the LMM theory from Section 1.2. The forward domestic Libor $L_n(t)$ is a rate starting at T_n and ending at T_{n+1} and is given by

$$L_n(t) = \frac{1}{\delta_n}\left(\frac{P(t,T_n)}{P(t,T_{n+1})} - 1\right), \tag{A.1.1}$$

where $P(t,T_n)$ is the time-t price of a zero-coupon bond with maturity T_n. The domestic LMM is defined by

- A set of maturities T_n, $n = 0, \dots, N$ such that $0 < T_0 < T_1 < \cdots < T_N$.
- Initial forward LIBOR rates $L_n(0)$.
- A volatility vector function $\lambda_n(t)$, $n = 0, \dots, N-1$.
- Time-independent shifts s_n for $n = 0, \dots, N-1$.

Similarly, a foreign LIBOR[1] $\tilde{L}_n(t)$ is a rate starting at $\tilde{T}_n$ and ending at $\tilde{T}_{n+1}$ and is given by

$$\tilde{L}_n(t) = \frac{1}{\tilde{\delta}_n}\left(\frac{\tilde{P}(t,\tilde{T}_n)}{\tilde{P}(t,\tilde{T}_{n+1})} - 1\right), \tag{A.1.2}$$

[1] We reserve ˜ (tilde) for foreign market quantities instead of using subscripts for ease of notation.

where $\tilde{P}(t,T_n)$ is the foreign zero-coupon bond maturing at T_n. The foreign model definition includes

- A set of maturities $\tilde{T}_n$, $n=0,\ldots,\tilde{N}$ such that $0<\tilde{T}_0<\tilde{T}_1<\cdots<\tilde{T}_{\tilde{N}}$.
- Initial forward LIBOR rates $\tilde{L}_n(0)$.
- A volatility vector function $\tilde{\lambda}_n(t)$, $n=0,\ldots,\tilde{N}-1$.
- Time-independent shifts $\tilde{s}_n$ for $n=0,\ldots,\tilde{N}-1$.

Note that domestic and foreign maturity sets can be different.

In their respective rolling-spot measures,the domestic and foreign LMM satisfy the SDE (1.2.10).

The numeraire associated with the rolling-spot measure is

$$N(t)=\frac{1}{P(T_0,T_1)}\cdots\frac{1}{P(T_{i-1},T_i)}\,\frac{P(t,T_{i+1})}{P(T_i,T_{i+1})},\qquad T_n\le t\le T_{n+1}, \tag{A.1.3}$$

for the domestic market, and

$$\tilde{N}(t)=\frac{1}{\tilde{P}(T_0,T_1)}\cdots\frac{1}{\tilde{P}(T_{i-1},T_i)}\,\frac{\tilde{P}(t,T_{i+1})}{\tilde{P}(T_i,T_{i+1})},\qquad \tilde{T}_n\le t\le \tilde{T}_{n+1}, \tag{A.1.4}$$

for the foreign market. As was shown by Schlögl (2002), the rolling-spot numeraire is equivalent to the continuously compounded savings account in the sense that they lead to the same measure. Another property shared by the rolling-spot numeraire and the savings account is diffusion-free dynamics

$$dN(t)=N(t)\,\mu_N(t)dt \quad \text{and} \quad d\tilde{N}(t)=\tilde{N}(t)\,\tilde{\mu}_N(t)dt. \tag{A.1.5}$$

The FX dynamics linking the single-currency and cross-currency models are defined by postulating lognormal dynamics for a numeraire-discounted asset $Y(t)$, which is related to the foreign exchange rate $X(t)$ by

$$Y(t)=\frac{X(t)\,\tilde{N}(t)}{N(t)}. \tag{A.1.6}$$

We set

$$dY(t)=Y(t)\sigma_Y(t)\cdot dW(t) \tag{A.1.7}$$

in the spot-Libor measure with a deterministic F-component volatility σ_Y. The dynamics of the FX rate $X(t)$ follow from (A.1.6) and (A.1.7) and the absence of diffusion terms in the numeraire dynamics (A.1.5):

$$dX(t)/X(t)=(\mu_N(t)-\tilde{\mu}_N(t))\,dt+\sigma_Y(t)\cdot dW(t). \tag{A.1.8}$$

Note that we have not introduced any skew in the process for $Y(t)$. We will see that the FX options skew will be naturally generated by the interest rate components, even if we started with purely lognormal Libor dynamics.

We complete the operational definition of the model by restating the dynamics of the LIBORs

in the **domestic** spot-LIBOR measure:

$$dL_n(t) = (L_n(t) + s_n)\left(\sum_{j=\eta(t)}^{n} \frac{\delta_j\,(L_j(t)+s_j)}{1+\delta_j\,L_j(t)}\,\lambda_n(t)\cdot\lambda_j(t)\right)dt + (L_n(t)+s_n)\,\lambda_n(t)\cdot dW(t), \tag{A.1.9}$$

$$d\tilde{L}_n(t) = (\tilde{L}_n(t) + \tilde{s}_n)\left(\sum_{j=\tilde{\eta}(t)}^{n} \frac{\tilde{\delta}_j\,(\tilde{L}_j(t)+\tilde{s}_j)}{1+\tilde{\delta}_j\,\tilde{L}_j(t)}\,\tilde{\lambda}_n(t)\cdot\tilde{\lambda}_j(t) - \tilde{\lambda}_n(t)\cdot\sigma_Y(t)\right)dt + (\tilde{L}_n(t)+\tilde{s}_n)\,\tilde{\lambda}_n(t)\cdot dW(t), \tag{A.1.10}$$

$$dY(t) = Y(t)\,\sigma_Y(t)\cdot dW(t). \tag{A.1.11}$$

The model can be simulated by inductively repeating the following steps after initialization at the origin of time.

- Get domestic and foreign states for the next date from a discretization of Eqs. (A.1.9) and (A.1.10).
- Get domestic and foreign numeraires, $N(t)$ and $\tilde{N}(t)$, from Eq. (A.1.5).
- Get the value of the process $Y(t)$ for the next date from a discretization of Eq. (A.1.11).
- Obtain the FX rate at the next date from $X(t) = Y(t)N(t)/\tilde{N}(t)$.

Corporate Headquarters
New York
99 Park Avenue
5th Floor
New York, NY 10016
Tel: +1 212 302 2220

Sao Paulo—Brasil
Avenida Brigadeiro
Faria Lima, 3729
5 Andar
Sao Paulo
04538905, Brasil
Tel: +5511 3443 7331

London
Numerix Software Limited
41 Eastcheap, 2nd Floor
London, UK EC3M 1DT
Tel: +44 20 7648 6100

Singapore
8 Eu Tong Sen Street
#24-85 The Central
Singapore 059818
Tel: +65 6501 0788

Hong Kong
Level 16
Man Yee Building
60-68 Des Voeux Road Central
Central, Hong Kong
Tel: +852 3796 2720

Tokyo
Marunouchi Trust
Tower Main, 20th Floor
1-8-3 Marunouchi, Chiyoda-ku
Tokyo, Japan 100-0005
Tel: +81 3 5208 1130

Numerix is the award winning, leading independent analytics institution providing cross-asset solutions for structuring, pre-trade price discovery, trade capture, valuation and portfolio management of derivatives and structured products. Since its inception in 1996, over 700 clients and 75 partners across more than 25 countries have come to rely on Numerix analytics for speed and accuracy in valuing and managing the most sophisticated financial instruments.

numerix®
Taking Analytics Further