Bank i Kredyt 50(2), 2019, 107-148

BANK I KREDYT
BANK & CREDIT

Markovian and multi-curve friendly parametrisation of a HJM model used in valuation adjustment of interest rate derivatives

Marcin Dec*

Submitted: 19 October 2018. Accepted: 13 February 2019.

Abstract

We consider feasible Heath-Jarrow-Morton framework specifications that are easily implementable in XVA engines when pricing linear and non-linear interest rate derivatives in a multi-curve environment. Our particular focus is on relatively less liquid markets (Polish PLN) and the calibration problems arising from that fact. We first develop the necessary tool-kit for multi--curve construction and XVA integration and then show and discuss various specifications of the HJM model with regard to their practical usage. We demonstrate the importance of the Cheyette subclass and derive the dynamics of instantaneous forward rates in generic forms of different models. We performed calibrations of several one-factor models of that form and found that even with a relatively simple specification, i.e. Hull-White with two summands, we may achieve satisfactory results in terms of the quality of the calibration and calculation time.

Keywords: instantaneous forward rate models, multi-curve valuation, XVA, HJM framework, Cheyette model

JEL: G12, G13, E43

* Warsaw School of Economics; e-mail: decmarcin@gmail.com.

1 Introduction

The Heath-Jarrow-Morton interest rate framework builds on the concept of instantaneous forward rates, allowing for movements of the whole yield curve in a non-arbitrage manner. In this article, we will be looking for a simple yet rich enough specification of the HJM class models that could be used in pricing engines for the calculation of miscellaneous valuation adjustments of interest rate derivatives. These add-ons (XVA) constitute the common approach to expressing additional risk factors involved in the valuation of derivative instruments. As with every project of high complexity, XVA calculation is also a predominantly IT hardware task, but here we will concentrate on particular algorithms, characteristics of models and the conditions necessary in subsequent phases of XVA development for the whole framework to be consistent, calibrable and capable of producing results within reasonable computational time. Because of the problem's high dimensionality and at least weak path dependence, the only suitable toolbox in practice is Monte Carlo simulation. The full deployment and implementation of XVA in a financial institution requires high-performance simulations, therefore we concentrate on a subclass of models with Markovian characteristics. The financial crisis of 2007–2008 (henceforth: FC) changed entirely the way all financial instruments are valued, especially in the interest rate world, by separating discounting and forwarding (also called: projection) curves, adding even more to the complexity of the task described so far.

There is a scarcity of after-crisis research on variants of multi-curve HJM implementations for XVA engines, and the latter topic has become very important in the industry. The article is organised as follows. Firstly, we set up the scene and outline general recipes for multi-curve construction and also show after crisis changes in the valuation of plain vanilla interest rate derivatives. Then we introduce all the definitions necessary to accurately characterise XVA integral and list the desired features and requirements of an interest rate model to be used as a workhorse in the XVA engine. The major contribution of this article is a concise presentation, discussion and practical implementation of one of the HJM's subclasses with a special focus on calibration in a multicurve environment. We demonstrate the importance of the Cheyette subclass and derive general dynamics of instantaneous forward rates in generic form. In the search for a tractable model and at least semi-analytical pricing formulae to exist at every state of the world of our simulations, we start with propositions of Brace and Musiela and modify the results to the multicurve environment via multiplicative spreads.[1]

The FC was the turning point in many aspects of valuation methodology and risk management of financial instruments. Particularly in the interest rate products domain, it has triggered a revolution caused by the breaking of the no-arbitrage assumption which served for years as the foundation of single curve valuation of these products. During the crisis there dramatic changes were observed in the levels, volatilities and liquidity of interbank money market products (uncollateralised deposits, short-term repo, fx swaps, OIS[2]) and interest rate derivatives (FRA, IRS, basis swaps, caps, floors, swaptions) which may be summarised as follows (cf. Bianchetti, Carlicchi 2011: Bianchetti, Carlicchi 2013; Bianchetti 2009):

[1] We would like to express our gratitude to Thomson Reuters (and Tullett Prebon) for their co-operation in making financial market data for an extensive list of instruments available for my research. This database included daily observations of prices and volatilities of all OTC quoted interest rate derivatives in PLN, EUR and USD from the period of 2014–2017. The opportunity to work with real market prices is always a good basis for developing practical and implementable solutions.

[2] OIS stands for Overnight Index Swap in which counterparts exchange fixed rate for a compound rolling overnight xONIA rate in an agreed period of time.

- Explosion of spreads between money market deposit (xIBOR based[3]) and OIS (xONIA based[4]) rates. The spreads abruptly increased from a long-term plateau of several basis points before FC to more than 200 bps at some point after FC, still hovering around several tens of basis points 10 years later.
- Divergence between FRA par rates and the corresponding forward rates implied by relevant unsecured deposits of far and near leg's maturity, as a consequence of ceased arbitrage possibilities mainly due to liquidity and capital constraints.
- Re-emergence of basis swap spreads risks (differences between different tenor xIBOR rates). The sheer existence of non-zero spreads (i.e. 3M–6M) meant that no-arbitrage relationship in classical terms stopped to hold.
- Outburst of credit spreads (measured by CDS spreads) especially of the biggest banks in the world – including xIBOR panellists – from near-zero levels to 80–300 bps range after FC.
- Evaporation of credibility and trust, which lead to dramatic squeeze of liquidity in many market segments in almost all instruments but the ones traded with central banks or highly collateralised.
- Shift towards CSA[5] discounting for collateralized cash flows and a strong market division into funded and unfunded products followed.

As a result of a diffusion of collateral agreements reducing credit risk between counterparts of OTC derivatives, these instruments' quotes in the market may now be regarded as risk-free. Since almost all exposures have to be collateralised now, the market cannot use the same discount curve to calculate net present values and to forecast forward rates. Hence the market moved to the best available proxy of risk-free discounting, namely OIS discounting.[6] Another important consequence of FC is that every tenor of a reference rate (i.e. 3M or 6M) may now be treated as a separate underlying asset which leads to challenges of the multi-curve environment. Some authors (i.e. Moreni, Pallavicini 2014) suggest that a good approach is to see different curves as if they were different currencies and refrain from trying to model why the curves differ, but rather describe how to incorporate multi-curve reality into one model. This proves to be a difficult task, mainly because of proper no-arbitrage conditions formulation and consistency in risk-neutral measures used in pricing. What is making this task even more cumbersome is the fact that the market actively and reliably prices only a few tenors and derivatives based on them.[7]

2 Construction of yield curves

Let's define key concepts and yield curves which we will use throughout this article by merging and simplifying the nomenclature used by Ametrano and Bianchetti (2009), Bianchetti (2010), Bianchetti

[3] xIBOR is a trimmed average reference rate for OTC money market unsecured deposits in currency x, usually calculated at 11:00 AM local time or different maturities (from O/N to 12 months, depending on the currency) on the basis of a questionnaire amongst the highest credit rated market participants (panelists). It is based on real transactions.

[4] xONIA is a weighted average overnight rate in currency x usually calculated by the central bank relevant for that currency from real O/N deposits settled between banks.

[5] CSA denotes Credit Support Annex of standard bilateral ISDA agreement regulating the rules of collateral posting against negative mark-to-market valuation of derivatives portfolios dealt between these counterparts.

[6] In the literature also referred to as CSA discounting.

[7] I.e. interest rate options are usually based on 6M tenors only, with some exception to 3M tenor being used in shorter maturity products.

and Carlicchi (2011), Chibane, Selvaraj, and Sheldon (2009) Kienitz (2013). Denote a discount curve based on instruments with an underlying tenor of j as:

$$\mathrm{C}_P^j(t)=\{T\to P^j(t,T)\} \tag{1}$$

where $P^j(t,T)$ is a discount factor for the period between today t and a certain date T in the future.

Assuming $j=\{ON, 1M, 3M, 6M, 12M\}$ and defining $F^j(t,T)$ as a forward rate implied by j-tenor curve observed at time t and "working" between date $T-\delta_j$ and T, where δ_j is a year fraction equal to the tenor length, we may specify the forward curve of j-tenor rates: $\mathrm{C}_F^j(t)=\{T\to F^j(t,T)\}$.

We may also define continuously compounded zero-coupon rates
$Z^j(t,T)=\frac{-\log P^j(t,T)}{(T-t)}$
and therefore the zero-coupon curve as $\mathrm{C}_Z^j(t)=\{T\to Z^j(t,T)\}$. It is crucial for our HJM framework later in this paper to propose an instantaneous forward rate concept $f^j(t,T)$:

$$P^j(t,T)=exp\left(-\int_t^T \mathrm{f}^j(t,u)\,du\right)\Rightarrow \mathrm{f}^j(t,T)=-\frac{\partial \ln P^j(t,T)}{\partial T} \tag{2}$$

$$\Rightarrow \mathrm{f}^j(t,T)=\mathrm{Z}^j(t,T)+(T-t)\frac{\partial Z^j(t,T)}{\partial T} \tag{3}$$

The instantaneous forward rate curve of tenor-j is then $\mathrm{C}_f^j(t)=\{T\to f^j(t,T)\}$. We have to establish some intra-tenor connection between different curves by setting a forward basis (assuming that the day count conventions on the two curves are the same and hence the year fractions):

$$\beta_F^{j,d}(t,T)=\frac{F^j(t,T)}{F^d(t,T)}=\frac{P^d(t,T)\left(P^j(t,T-\delta_j)-P^j(t,T)\right)}{P^j(t,T)\left(P^d(t,T-\delta_j)-P^d(t,T)\right)} \tag{4}$$

and the forward basis curve would be: $\mathrm{B}_F^{j,d}(t)=\{T\to \beta_F^{j,d}(t,T)\}$.

For instance, if we take overnight OIS curve as a discount curve and would like to price some instruments based on xIBOR3M and xIBOR6M we would need to construct a collection of curves: $\mathcal{C}=\{\mathrm{C}_P^d(t),\mathrm{C}_F^{3M}(t),\mathrm{C}_F^{6M}(t)\}$ and two basis curves will result from this construction as well: $\mathcal{B}=\{\mathrm{B}_F^{3M,d}(t),\mathrm{B}_F^{6M,d}(t)\}$.

Having the new multi-curve set-up defined above, we may outline a general pricing algorithm in this environment and compare it with the single-curve set-up where appropriate. There are many simple, macro-level algorithms presented in the literature after FC by Ametrano and Bianchetti (2013), Bianchetti (2009), Bianchetti and Carlicchi (2013) or Henrard (2014), from which the recipe below originates:

- Construct a single discounting curve $\mathrm{C}_P^d(t)$ using liquid vanilla interest rate instruments traded in the market, with increasing maturities (mainly OIS swaps) and a chosen bootstrapping scheme. Choose an interpolation method. In a single curve world, we do construct one curve which serves both as a discounting and forwarding curve, and we do not use OIS swaps to construct it, but rather a mix of the most liquid $xIBOR$-based derivatives.
- Construct multiple forwarding curves i.e.: $\mathrm{C}_F^{1M}(t), \mathrm{C}_F^{3M}(t), \mathrm{C}_F^{6M}(t), \mathrm{C}_F^{12M}(t)$ depending on needs and market data availability. Each curve is constructed using simple linear interest rate instruments, homogeneous in rate tenor, i.e.: for $\mathrm{C}_F^{3M}(t)$ we would take: xIBOR3M, FRA3x6, FRA6x9, ..., FRA18x24, IRS2Y3M, IRS3Y3M, ..., IRS10Y3M. Decide on the interpolation method. In the single curve case we skipped this step of construction.
- For each variable cash flow c_k of a derivative to be priced, compute an estimate of the relevant forward rate, $F^j(t,T)$, from the relevant forwarding curve $\mathrm{C}_F^j(t)$.
- Compute the expected cash flows as the time-t expectation of the interest rate related payoff forward measure Q^{T_k} associated to a corresponding discount factor $P^d(t,T_k)$. Compute the relevant discount factors $P^d(t,T_k)$ as well.
- The value of the derivative is just the sum of the discounted expected cashflows.

The last three bullets are the same in the classical and modern approach, but the discounting curve is, obviously, different. For the sake of proper calibration and practical use we need to specify also:

– the choice between bootstrap and root-finding Jacobian procedure,

– the interpolation scheme,

– the methods of dealing with gaps or lack of data in certain segments of a curve.

For reasons of space and the fact that we will be pricing relatively simple instruments, we will stick to bootstrapping as a method of extracting curves from market data, rather than a global root-finding procedure (as proposed in Henrard 2014), which may be more relevant if one faces a problem of more intertwined curves and instruments. Detailed recursive algorithms for bootstrapping the discount and forwarding curves will follow in the subsequent sections.

With regard to the choice of the interpolation method, we will follow the recommendations of Hagan and West (2006), who found that based on the following criteria:

– continuity and positivity of forward rates,

– minimisation of little spill-over effect (locality of interpolation),

– stability of forwards (bumping does not change much in the shape of the curve),

– locality of hedges (delta risk of hedge concentrates near the underlying with no filtering to other areas of the curve),

the best results, although for distinct purposes, are achieved using a linear interpolation on the logarithms of discount factors[8] and a monotone convex interpolation on the logarithms of discount factors (which is much more demanding numerically, hence usually implemented directly in a software package[9]). The raw linear method of interpolation between two points T_i and T_{i+1} at T may be summarized as follows:

$$P(t,T) = P(t,T_{i+1})^{\frac{T-T_i}{T_{i+1}-T_i}} P(t,T_i)^{\frac{T_{i+1}-T}{T_{i+1}-T_i}} \tag{5}$$

[8] Which Hagan and West called originally: raw.

[9] I.e. there are classes: *scipy.interpolate.PchipInterpolator and scipy.interpolate.UnivariateSpline* in Python which handles monotonic cubic splines (piecewise cubic hermite interpolating polynomial) and tension splines (via smoothing parameter *k*).

This method, despite being very helpful in a sensitivity analysis, results in piecewise constant instantaneous forward rates, which in turn is not a desirable feature, especially in Heath-Jarrow-Morton framework. On the other hand, the monotone cubic splines class has a very tempting characteristic of smoothness and produces visually "round" forward curves, however, one should be very careful using it for hedging and sensitivity analysis, as the method suffers from such problems as spurious inflection points, excessive convexity and lack of locality (when curve bumping). These drawbacks are mitigated in the tension splines method as proposed by Hagan and West (2006).

In practice the maximum tenors of the instruments used to construct discounting and forwarding curves differ. In the case of Polish PLN interest rate derivatives it is very well pronounced, since OIS swaps are quoted up to 2 years and IRS swaps up to 20. Henrard (2014) suggests to use the spread-over-existing method, which means freezing the longest available basis from market data (2 years in the case of PLN) and applying it to calculate the rest of discounting and forwarding curves. The bootstrapping algorithms of these two curves are therefore entangled and have to be adjusted accordingly.

2.1 Discounting curve

Nowadays it is general market practice that one uses OIS swap market rates and xONIA to build the best proxy of risk-free rates – OIS discounting curve $\mathrm{C}_P^d(t)$. Every curve (discounting and forwarding) has to start somewhere, so the following nearest point on the curve should be selected:

$$P^x(t,T) = \frac{1}{1 + R_{depo}^x(t,T)\delta_x} \tag{6}$$

where $R_{depo}^x(t,T)$ is a *normal* unsecured deposit rate (for $x = d$ we have $R_{depo}^x = R_{depo}^d = \text{xONIA}$) and δ_x is a year fraction for the period of (t, T) with proper day count convention (i.e. 360 for EUR and 365 for PLN).

Let's denote different rate schedules in any interest rate swap as $\{T_0, T_1, \ldots, T_n\}$ for floating leg and $\{S_0, S_1, \ldots, S_m\}$ for fixed leg, with additional conditions $T_0 = S_0$ and $T_n = S_m$. Define an auxiliary variable – a swap annuity as:

$$\mathrm{A}^d(t,S) = \sum_{j=1}^{m} P^d(t,S_j)\delta_{S_{j-1},S_j} \tag{7}$$

Setting the schedules of floating and fixed leg to be the same we may obtain from the par OIS rate formula:

$$P^d(t,T_i) = \frac{1 - R_{OIS}^d(t,T_i)\sum_{j=1}^{i-1} P^d(t,T_j)\delta_{T_{j-1},Tj}}{1 + R_{OIS}^d(t,T_i)\delta_{T_{i-1},Ti}} \tag{8}$$

where $R_{OIS}^d(t,T_i)$ being the fixed rate of OIS swap with T_i maturity.

In such a way, consecutive, longer discount factors of $\mathrm{C}_P^d(t)$ may be obtained recursively in each step i.

2.2 Forwarding curves

To calculate forwarding curves, we start with the shortest discount factors[10] as in the discounting curve case, but this time using for example xIBOR3M:

$$P^{3M}(t,T)=\frac{1}{1+R_{depo}^{3M}(t,T)\delta_{3M}} \tag{9}$$

Then we may use consecutively adjacent FRA contracts i.e. 3×6, 6×9, 9×12,..., 18×24 to get discount factors for $C_P^{3M}(t)$ curve from a recursive formula:

$$P^{3M}(t,T_i)=\frac{P^{3M}(t,T_{i-1})}{1+R_{FRA}^{3M}(t,T_i)\delta_{T_{i-1},Ti}} \tag{10}$$

where $R_{FRA}^{3M}(t,T_i)$ is a market FRA par rate for a xIBOR3M to be fixed at $T_{i\text{-}1}$ and $\delta_{T_{i-1},Ti}\approx 3M$ depending on the precise day count convention.

The liquidity of FRA contracts usually dries up above the 2 year mark, therefore, we should use longer contracts to continue building our curve. In the case of IRS swaps where fixed and floating rates are exchanged periodically, we will have to use the OIS discount curve from the previous section. This is the key difference to the other instruments so far discussed. Again from par rate equivalence in the IRS contract at curve's time pillars where $T_i = S_j$ we have:

$$P^{3M}(t,T_i)=\frac{P^d(t,T_i)P^{3M}(t,T_{i-1})}{R_{IRS}^{3M}(t,T_i)\,\mathrm{A}^d(t,S_j)-\sum_{j=1}^{i-1}P^d(t,T_j)F^{3M}(t,T_j)\delta_{T_{j-1},Tj}+P^d(t,T_i)} \tag{11}$$

which, in fact, is recursive as $F^{3M}(t,T_j)\delta_{T_{j-1},Tj}=\dfrac{P^{3M}(t,T_{j-1})}{P^{3M}(t,T_j)}-1$ is evaluated in the sum up to $i-1$ term.

The problem of not equal frequency of interest periods on fixed and floating legs, which leads to necessity of interpolation during bootstrapping, may be easily overcome by introducing intermediate synthetic instruments that are interpolated first from the market data. For example, if we are constructing $C_F^{3M}(t)$ and have market prices of IRS6Y3M and IRS7Y3M, we simply add 3 synthetic IRS swaps with maturities 6.25Y, 6.5Y and 6.75Y with interpolated prices using the chosen interpolation method. Only then do we follow the recursive formula (11).

In the case of the spread-over-existing method to cope with no OIS data for longer periods (as for example in the case of the Polish zloty), we assume a constant multiplicative spread β for some $i > i^*$ where i^* is the time index of the longest maturity OIS, or in other words:

$$\beta=\frac{P^d(t,T_i)\left(P^{3M}(t,T_{i-1})-P^{3M}(t,T_i)\right)}{P^{3M}(t,T_i)\left(P^d(t,T_{i-1})-P^d(t,T_i)\right)} \tag{12}$$

[10] Henrard (2014) calls it pseudo discount factors as it is not used in real discounting any more.

It is easily obtainable that:

$$P^{3M}\left(t,T_i\right)=\frac{P^d\left(t,T_i\right)P^{3M}\left(t,T_{i-1}\right)}{P^d\left(t,T_i\right)\left(1-\beta\right)+\beta P^d\left(t,T_{i-1}\right)} \tag{13}$$

Plugging this result into (11) and solving for $P^d\left(t,T_i\right)$ while noticing that $\mathrm{A}^d\left(t,S_j\right)=\mathrm{A}^d\left(t,S_{j-1}\right)+P^d\left(t,T_j\right)\delta_{T_{j-1},T_j}$ we get:

$$P^d\left(t,T_i\right)=\frac{\sum_{j=1}^{i-1}P^d\left(t,T_j\right)F^{3M}\left(t,T_j\right)\delta_{T_{j-1},Tj}+\beta P^d\left(t,T_{i-1}\right)-R_{IRS}^{3M}\left(t,T_i\right)\mathrm{A}^d\left(t,T_{i-1}\right)}{R_{IRS}^{3M}\left(t,T_i\right)\delta_{T_{i-1},Ti}+\beta} \tag{14}$$

Knowing $P^d\left(t,T_i\right)$ we may retrieve $P^{3M}\left(t,T_i\right)$ from (13). Having $\mathrm{C}_P^{3M}\left(t\right)$ and $\mathrm{C}_P^d\left(t\right)$ it is trivial to get $\mathrm{C}_F^{3M}\left(t\right)$ or $\mathrm{C}_Z^{3M}\left(t\right)$ in the time points which are a multiple of a tenor (i.e 0.25), but it is not so in the whole domain of *T* or in the cases of instantaneous forwards $\mathrm{C}_f^d\left(t\right)$ or $\mathrm{C}_f^{3M}\left(t\right)$. As indicated in the previous subsection, we have to decide on interpolation first. For comparison, we use two interpolation methods (raw and monotone cubic spline) on the logarithms of discount factors of $\mathrm{C}_P^{3M}\left(t\right)$ and $\mathrm{C}_P^d\left(t\right)$ and then we are able to calculate forwarding curves for all *T*:

$$F^{3M}\left(t,T\right)=\frac{1}{\delta_{3M}}\left(\frac{P^{3M}\left(t,T-\delta_{3M}\right)}{P^{3M}\left(t,T\right)}-1\right) \tag{15}$$

$$f^d\left(t,T\right)=-\frac{\partial \ln P^d\left(t,T\right)}{\partial T}\approx-\left(\frac{\ln P^d\left(t,T+h\right)-\ln P^d\left(t,T\right)}{\mathrm{h}}\right) \tag{16}$$

letting *h* be a small year fraction (i.e. 1/365) used in the first derivative of log discount function's approximation.

3 Valuation adjustments (XVA)

In the broadest sense, XVA is an adjustment of a risk-free financial instrument's valuation for some types of risks that the primary valuation does not account for. The most popular valuation adjustments (XVA) are: Credit (CVA) – to account for counterparty risk, Debit (DVA) – to account for bank's own credit risk,[11] Funding (FVA) – to account for funding risk, including cost of liquidity buffers, Collateral (ColVA) – to measure the impact of collateral effects, Margin (MVA) – to account for funding risk of initial margin, Capital (KVA) – to measure the impact of regulatory capital, Tax (TVA) – to model the influence of tax on the valuation.

The economic value $\hat{V}$ (including XVAs) of a financial instrument is equal to the sum of risk-free valuation (usually based on OIS discounting) and all of the adjustments a bank should or would like to include:

[11] There was a widespread critique of DVA, questioning the possibility of monetisation of one's own default risk. For more details see: Green (2015).

$$\hat{V} = V + \sum_{i=1}^{n} XVA_i \tag{17}$$

where n is the number of adjustments used in a particular case.

This approach and practice may be treated as violations of the law of one price in the market, since a valuation of every derivative instrument is somehow "localised" to the particular counterparty, bank and collateral arrangement and it would be very hard to imagine that everybody agrees on the way a certain deal is evaluated. We propose a detailed set of XVA definitions below which is a result of combining the best parts from Brigo, Morini, and Pallavicini (2013), Green (2015), Gregory (2012), Kienitz and Caspers (2017) and Lu (2015).

Let's call $V(t, T)$ a position of cash-flows at $t_0 \le t \le T$ with final maturity T. Then the mark-to-market valuation at time t in risk-neutral measure Q would be $\mathrm{E}_t^{\mathrm{Q}}\left[V(t,T)\right]$. Exposure at time t is then a positive part of that valuation:

$$\varepsilon(t)^+ = \left(\mathrm{E}_t^{\mathrm{Q}}\left[V(t,T)\right]\right)^+ \tag{18}$$

Obviously, if we are at time t_0 the term $\varepsilon(t_0)^+$ can be called a current exposure. The exposure at default (τ_C) of a counterparty C is:

$$\varepsilon(\tau_C)^+ = \left(\mathrm{E}_{\tau_C}^{\mathrm{Q}}\left[V(\tau_C,T)\right]\right)^+ \tag{19}$$

The expected (at time t_0) exposure EE that we would have at time t is then given by:

$$EE_{t_0}(t) = \mathrm{E}_{t_0}^{\mathrm{P}}\left[\varepsilon(t)^+\right] = \mathrm{E}_{t_0}^{\mathrm{P}}\left[\left(\mathrm{E}_t^{\mathrm{Q}}\left[V(t,T)\right]\right)^+\right] \tag{20}$$

Note that the external expectation is taken with respect to physical measure P. It is common to visualise expected exposures via a mapping $t \longmapsto EE_{t_0}(t)$ for $t_0 \le t \le T$ and call it an expected exposure profile.[12]

In valuation adjustment the world would like to know not only the expected exposures but also to have some idea about their distributions, hence we introduce PFE (potential future exposure) and *MPFE* (maximum PFE in an interval):

$$PFE_q(t) = \alpha_q(\varepsilon(t)^+) \tag{21}$$

$$MPFE_q(t) = \sup_{s\in[t_0,t]} \alpha_q(\varepsilon(s)^+) \tag{22}$$

where $\alpha_q(\cdot)$ is a quantile function (or inverse cumulative distribution function) and q is a confidence level required. A mapping $t \longmapsto PFE_q(t)$ is called potential future exposure profile.

12 Sometimes also called: credit equivalent exposure curve or loan equivalent exposure curve.

The ultimate possible loss that one may suffer from a default of a counterparty C is equal to the multiplication of exposure at default $\varepsilon(\tau_C)^+$ and a loss-given-default usually defined as $LGD = 1-R_C$, where R_C stands for a deterministic assumed recovery rate after C defaults. Hence unilateral credit valuation adjustment (CVA) is defined as:

$$\begin{aligned}\mathrm{CVA}_{t_0} &= \left[(1-R_C)P(t_0,\tau_C)1_{\{\tau_C<T\}}\varepsilon(\tau_C)^+\right] \\ &= \Phi(\tau_C<T)\mathrm{E}^{\mathrm{Q}}_{t_0}\left[(1-R_C)P(t_0,\tau_C)|\tau_C<T\right]\end{aligned} \tag{23}$$

where T is a maximum time of scheduled occurrence of any of the cash-flows in a position[13] under valuation, and $\Phi(\cdot)$ is a cumulative default probability.

Then unilateral CVA is essentially a product of a cumulative probability of default of a counterpart and expected discounted loss given that default.

In practice, there are some assumptions usually made for the sake of a balance between model complexity and availability of data (cf. Green 2015; Brigo, Morini, Pallavicini 2013 or Gregory 2012), which may be relaxed at some point:

- discretisation of the default timeline in the form of a partition $\pi_{t_0,T}$

$$\pi_{t_0,T} = \{(0=t_0,t_1],(t_1,t_2],\ldots,(t_{n-2},t_{n-1}],(t_{n-1},t_n=T]\} \tag{24}$$

- approximation that whenever a default happens within a period $(t_{i-1}, t_i]$ we assume that it occurs at the precise end of that segment: $\tau = t_i$,
- credit risk is independent of any other market risks, which means that for any i a default τ within a period $(t_{i-1}, t_i]$ is independent of a current exposure $\varepsilon(t_i)^+$,
- default times of a bank τ_B that is evaluating its positions and the counterparts τ_C are independent.

Under these assumptions we have the following bucketed approximation of CVA:

$$\begin{aligned}\mathrm{CVA}_{t_0} &= (1-R_C)\sum_{i=1}^{n}\mathrm{E}^{\mathrm{Q}}_{t_0}\left[P(t_0,\tau_i)1_{\{\tau_C\in(t_{i-1},t_i]\}}\varepsilon(\tau_C)^+\right] = \\ &= (1-R_C)\sum_{i=1}^{n}\Phi(t_{i-1}<\tau_C<t_i)\mathrm{E}^{\mathrm{Q}}_{t_0}\left[P(t_0,t_i)\varepsilon(\tau_C)^+|\tau_C\in(t_{i-1},t_i]\right] = \\ &= (1-R_C)\sum_{i=1}^{n}\Phi(t_{i-1}<\tau_C<t_i)\mathrm{E}^{\mathrm{Q}}_{t_0}\left[P(t_0,t_i)\varepsilon(\tau_C)^+\right]\end{aligned} \tag{25}$$

Changing risk-neutral expectations to physical ones and taking out the discount factors from under the expectation is yet another approximation very common in the literature and practice. In such a case, unilateral CVA is a weighted (by the probability of default in different buckets) sum of discounted expected exposures.

Also for the sake of simplicity, we assume that the counterpart and bank's hazard rates and therefore cumulative probabilities are given or bootstrapped from market data.

[13] Or portfolio as rarely CVA is calculated on single trade level.

After FC the role and usage of collateralisation in derivative trading both in the form of daily margining and so-called initial margin under CSA credit annexes to ISDA frame agreement spread extensively, resulting in the current situation in which the vast majority of trades are dealt within collateral frameworks (bilateral or CCP).

Posting collateral against mark-to-market value yields in some additional financial costs to the counterparty for which the deal generates negative exposure and may potentially bring some funding benefits otherwise, mainly due to the global unification of the rates paid on the collateral accounts (xONIA based[14]), which are much lower that the bank's funding costs. It is quintessential to have an effective model that generates future states of the world, whereas consequential calculations of XVAs are model-dependent but do not require any further simulations.

Green (2015) summarises other than CVA and DVA common XVA formulae in a stylised manner assuming that every adjustment is in the form of:

$$XVA(\alpha,\beta)=\int_t^T \alpha(u)e^{\int_t^u r(u)+\lambda_C(u)+\lambda_B(u)ds}\,\mathrm{E}_t\left[\beta(u)\right]du \tag{26}$$

where $r(\cdot)$ is a risk-free rate, $\lambda_c(\cdot)$ is a hazard rate of a counterparty, $\lambda_B(\cdot)$ is a hazard rate of a bank and the $\alpha(\cdot)$ and $\beta(\cdot)$ are different functions depending on the type of adjustment we would like to calculate.

Table 1 displays examples of such functions and their meanings.

Table 1

Variety of other than CVA/DVA value adjustments' forms

XVA	$\alpha(\cdot)$	$\beta(\cdot)$	**Meaning**
FVA	S_F	$\varepsilon(u)$	derivatives funding costs
$ColVA_X$	SX		collateral cost or benefit
KVA	$\gamma_K - r_B\phi$	$K(u)$	capital charge costs
MVA	$S_F - S_{I_B}$	$I_B(u)$	initial margin impact
TVA_E	γE	$E(u)$	tax effects

S_F is a funding spread defined as a difference between bank's bond yield and risk-free rate,

S_X is a collateral spread – a difference between a yield on collateral account and risk-free,

S_{I_B} is an initial margin spread – a difference between a yield on initial margin and risk-free,

γ_K is a cost of capital,

ϕ is a fraction of capital available for derivative funding,

$K(u)$ is a capital requirement at time u,

$X(u)$ is a collateral at time u,

$E(u)$ is a cash-flow liable to tax at time,

u, $I_B(u)$ is margin account value at u.

[14] Also referred to as: OIS collateralization.

4 Requirements for interest rate model in XVA calculations

It is essential that a proposed interest rate model framework is able to generate an object sometimes called CVA cube (cf. Kienitz, Caspers 2017), which is a three-dimensional matrix:

$$\boxplus = \left\{ \left(V_i, S_j, t_k \right) : \mathrm{R} \times \mathrm{N} \times \mathrm{R}^+ ; i = 1, .., I; j = 1, \ldots, J; k = 1, \ldots, K \right\} \tag{27}$$

where V_i is the valuation of the i-th instrument in a portfolio in a certain scenario and time-step, S_j is a j-th simulation and t_k is k-th time-step.

Having calculated $\boxplus$ it is easy to perform all sorts of XVA related calculations involving expected exposures as well as quantile potential exposures.

In designing HJM or any other interest rate model framework to be used in the XVA engine, considerable attention must be paid to the following (based on Green 2015; Henrard 2014; Kienitz, Caspers 2017):

– the requirement for high-performance Monte Carlo (as the majority of time would be spent on valuations in each time step rather than on path building per se),

– the necessity of the model to be Markovian, that means a need of all model quantities to be state--dependent, but not path-dependent,

– the existence of analytical or semi-analytical pricing formulae for the derivatives instruments of our interest,

– the ability of the framework to produce a fast discount factor curve at each node of simulation,

– the ease of multi-curve extension (i.e. via multiplicative spread),

– the possibility of replicating relatively rich shapes of curves,

– the ability to incorporate negative rates reasonably.[15]

In the ideal world we would also like our model to incorporate all of the following characteristics in the calibration phase:

– allowing for different currencies (with scarce liquidity/data availability),

– fast calibration time,[16]

– stable calibration (i.e. bootstrap wherever possible) that produces vega sensitivities.

5 HJM framework in a multi-curve environment

5.1 Classic proposition of Heath-Jarrow-Morton

The backbone of the Heath, Jarrow and Morton (1992) approach is an instantaneous forward rate process which allows modelling of the whole yield curve movements. The instantaneous forward rate $f(t,T)$ is defined as follows:

$$f(\mathrm{t},\mathrm{T}) = -\frac{\partial \ln \mathrm{P}(\mathrm{t},\mathrm{T})}{\partial \mathrm{T}} \Rightarrow P(\mathrm{t},\mathrm{T}) = \mathrm{e}^{-\int_{\mathrm{t}}^{\mathrm{T}} \mathrm{f}(\mathrm{t},\mathrm{u})\mathrm{du}} \tag{28}$$

where $P(t, T)$ is a zero-coupon bond price at time t with T maturity.

[15] Hence Gaussian models would be preferred.

[16] Although it may be set as a background process and cache the results every, say 10–15 minutes.

The dynamics of instantaneous forward rate in P-measure is given by:

$$f(t,T) = f(0,T) + \int_0^t \mu(s,T)\,ds + \int_0^t \sigma(s,T)\,dW(s) \tag{29}$$

where $dW(t)$ is a one dimensional P-Brownian motion and $\mu(t,T)$ and $\sigma(t,T)$ are one dimensional stochastic processes of drift and volatility of the forward rate respectively.

There are several conditions on regularity of these two functions μ and σ, which are not very restrictive as they only allow for an integral equation (29) to be sensible, namely (after Beyna, Wystup 2010):

$$\int_0^T |\mu(s,T)|\,ds < \infty \;\; P-a.s.\;\; \forall 0 \le T \le T^* \tag{30}$$

$$\int_0^T \sigma^2(s,T)\,ds < \infty \;\; P-a.s.\;\; \forall 0 \le T \le T^* \tag{31}$$

$$\int_0^{T^*} |f(0,s)|\,ds < \infty \;\; P-a.s. \tag{32}$$

$$\int_0^{T^*}\left(\int_0^u |\mu(s,u)|\,ds\right)du < \infty \;\; P-a.s. \tag{33}$$

where $T^* \ge T$ is the maximum time horizon.

Short interest rate process for $t \le T$ then is given by:

$$r(t) = f(t,t) = f(0,t) + \int_0^t \mu(s,t)\,ds + \int_0^t \sigma(s,t)\,dW(s) \tag{34}$$

The last term above (stochastic integral) suggests that the short rate process is non-Markovian, since t appears both as an upper limit of integration and inside the integrand. Using the definitions of $P(t,T)$ from (28) and of $r(t)$ above (34), Fubini's theorem for stochastic integrals and finally Itô lemma on $\ln P(t,T)$ yields in:[17]

$$\begin{aligned} P(t,T) = P(0,T) &+ \int_0^t P(s,T)\left(r(s) - \int_s^T \mu(s,u)\,du + 1/2\left\|\nu(s,T)\right\|^2\right)ds + \\ &+ \int_0^t P(s,T)\,\nu(s,T)\,dW(s) \end{aligned} \tag{35}$$

where $\nu(s,T) = -\int_s^T \sigma(s,u)\,du.$ Under the assumption of no arbitrage and complete market (the existence of a unique martingale measure is guaranteed) the discounted bond price process $\frac{P(t,T)}{B(t)}$ is a Q-local martingale. Q is an equivalent martingale measure if and only if:

[17] For full derivation see (Filipovic 2009).

$$\mu(t,T)=\sigma(t,T)\left(\int_t^T \sigma(t,s)ds-\gamma(t)\right) \tag{36}$$

where $\gamma(t)$ is a process of market price of risk. After the measure change from P to Q by means of Girsanov's theorem we obtain that $\gamma(t) = 0$.

This result is commonly referred to as the HJM drift condition. Incorporating it into forward rate dynamics and extending the results to m-dimensional Brownian motion we get under risk-neutral measure Q:[18]

$$\begin{aligned} f(t,T) &= f(0,T)+\sum_{i=1}^{m}\int_0^t\left(\sigma_i(u,T)\int_u^T\sigma_i(u,s)ds\right)du+\sum_{i=1}^{m}\int_0^t\sigma_i(s,T)dW_i^{\mathrm{Q}}(s) \\ r(t) &= f(0,t)+\sum_{i=1}^{m}\int_0^t\left(\sigma_i(u,t)\int_u^t\sigma_i(u,s)ds\right)du+\sum_{i=1}^{m}\int_0^t\sigma_i(s,t)dW_i^{\mathrm{Q}}(s) \end{aligned} \tag{37–38}$$

which form the key results of the HJM framework.

The key characteristics of the general form HJM framework may be summarised as follows:
– infinite dimensionality of the state space of volatility functions $\sigma(t, T)$,
– instantaneous forwards are not observed on the market,
– no need for initial curve calibration as current term structure is by construction an input to the model since instantaneous forward is calculated from the initial curve of zero coupon bonds,
– the term structure of volatility determines the forward rate at all times since the dynamics of the forward rate is determined by the short rate and the cumulative quadratic variation,
– non-Markovianity of the framework (as $r(t)$ is path dependent),
– $f(t,T)$ and $r(t)$ are Gaussian.[19]

These properties make the general HJM framework very challenging, if not useless to some applications. The PDE, Markov functional or trees methods are based on conditional expectations, hence it is crucial to work with Markov processes to use them. In the Monte Carlo domain there is no need for Markovianity, but since we are looking for a framework and algorithm enabling us to produce valuations at each time-step of the MC routine simulating yield curves for expected exposure calculations, the problem would grow in dimensionality multiplicatively in each period, making the whole idea infeasible.

Nonetheless, there is a vast catalogue of research proposing solutions to these drawbacks. The main aim is to restrict the generic form of volatility function to some class (parametrisation) that breaks down the general high-dimensional HJM framework to a structure of low-dimensional Markov processes with a relatively small number of state variables to keep track of. Having the special volatility function which guarantees the processes involved to be Markovian, one can try to develop pricing formulae (analytic or semi-analytic) for plain vanilla non-linear interest rate derivatives (swaptions,

[18] As we will deal with a risk neutral environment from this point on, we drop the Q upper-script.
[19] If only volatility function is deterministic.

caplets and floorlets as the building blocks of caps/floors) and calibrate volatility surface parameters to market quotes.

In the multi-curve world there have been two major groups of strategies as to how to cope with the IBOR-OIS spreads in the HJM framework. The first one may be summarised as the usage of the single curve set-up (old) to simulate risk-free rates followed by modification of pricing formulae of derivative instruments by the basis spread (usually deterministic multiplicative spread). Whereas the other strategy is to model and simulate all relevant curves for market valuation within the original model and develop new pricing formulae for them.

5.2 Cheyette class of HJM models

In search of Markovianity of the short interest rate process in the HJM framework, Ritchen and Sankarasubramanian (1995) proved the following lemma.

Lemma 5.1 (Ritchken-Sankarasubramanian)

In a one-factor HJM model with volatility function $\sigma(t, T)$ (being differentiable w.r.t T), a necessary and sufficient condition for the price of interest rate derivative to be completely determined by a two-state Markov process $\chi(\cdot)=\left(r(\cdot),\phi(\cdot)\right)$ is that the volatility function is of the following form:

$$\sigma(t,T)=\beta(t)e^{-\int_t^T \kappa(u)du}$$

where β is an adapted process and κ is a deterministic function. Then the second component of the Markov process is $\phi(t)=\int_t^T \sigma^2(s,t)ds$ and, in fact, β is the instantaneous short-rate volatility process.

One can argue that this formulation is a special case of the work and proposal of Cheyette (1995), which may be summarised as follows.

Lemma 5.2 (Cheyette)

In M-factor HJM model with volatility function of a form:

$$\sigma^k(t,T)=\sum_{i=1}^{N_k}\frac{\alpha_i^k(T)}{\alpha_i^k(t)}\beta_i^k(t),\ k=1,\dots,M \tag{39}$$

where N_k denotes the number of volatility summands of each factor k, the dynamics of a forward rate is determined by $n=\sum_{k=1}^{M}N_k$ state variables which are Markov processes and consequently the short rate is given as a sum of the initial forward rate and all state variables, hence the SDE for short interest rate is Markovian.

Due to great interconnection of these two lemmas, the class of the HJM models that stem from them is sometimes referred to as the Cheyette-Ritchen-Sankarasubramanian class (in short: Cheyette). The other names: quasi-Gaussian (cf. Andersen, Piterbarg 2010; Jamshidian 1991) or pseudo-Gaussian[20]

[20] We will use either; Cheyette or quasi-Gaussian.

are used for the models in with we allow $\beta(\cdot)$ – a function of time-t (current time) – to be stochastic. It may be treated as a partial relaxation of a Gaussian requirement on volatility function of the general HJM model to be Gaussian.

Admittedly, there is a trade-off between accuracy driven by the number of summands in each factor and the speed of calibration to market volatility surface or derivatives pricing. The accuracy and speed is also influenced by the number of parameters in each of the elementary functions used α_i and β_i.

General Cheyette model dynamics

By simply plugging the Cheyette separable volatility structure (39) into the forward rate dynamics of each factor from general HJM derivation (38) we obtain (cf. Beyna 2013):

$$f(t,T)=f(0,T)+\sum_{k=1}^{M}\left(\sum_{j=1}^{N_k}\frac{\alpha_j^k(T)}{\alpha_j^k(t)}\left(X_j^k(t)+\sum_{i=1}^{N_k}\frac{A_i^k(T)-A_i^k(t)}{\alpha_i^k(t)}V_{ij}^k(t)\right)\right) \tag{40}$$

where $X_i^k(t), A_i^k(t), V_{ij}^k(t)$ are defined as follows:

$$X_i^k(t)=\int_0^t\frac{\alpha_i^k(t)}{\alpha_i^k(s)}\beta_i^k(s)dW_k(s)+\int_0^t\frac{\alpha_i^k(t)}{\alpha_i^k(s)}\beta_i^k(s)\left(\sum_{j=1}^{N_k}\frac{A_j^k(t)-A_j^k(s)}{\alpha_j^k(s)}\beta_j^k(s)\right)ds \tag{41}$$

$$A_i^k(t)=\int_0^t\alpha_i^k(s)ds \tag{42}$$

$$V_{ij}^k(t)=\int_0^t\frac{\alpha_i^k(t)\alpha_j^k(t)}{\alpha_i^k(s)\alpha_j^k(s)}\beta_i^k(s)\beta_j^k(s)ds=V_{ji}^k(t) \tag{43}$$

where: $k = 1,\dots, M$ and $i, j = 1,\dots, N_k$.

Note that since we used a multi-factor model, the volatility function is a column vector[21] of Cheyette form volatility function for each factor:

$$\sigma(t,T)=\left(\sigma^1(t,T),\sigma^2(t,T),\dots,\sigma^M(t,T)\right)^{\mathrm{T}} \tag{44}$$

The first state variable X_i^k is stochastic as there is a vector of Brownian motions in the first term and in fact it describes short interest rate movement because $r(t)$ is the sum of the initial instantaneous rate $f(0, t)$ and the sum of all $X_i^k(t)$ (for all summands and all factors).[22]

The second state variable $V_{ij}^k(t)$ corresponds to the cumulative quadratic variation of a forward process, but it is deterministic for any $t \leq T$ as soon as we know the deterministic volatility surface of the form proposed by Cheyette at time t. $A_i^k(t)$ is just a technical integral used by the authors for lucid presentation.

---

[21] Where $(\cdot)^{\mathrm{T}}$ is an operator of matrix transposition.

[22] It suffices to plug $r(t) = f(t, t)$ into the forward rate dynamics equation to see that the majority of terms other than X_i^k cancel out.

It is easy to differentiate (41) and (43) to get the differential equations describing the general dynamics of X and V:

$$dX_i^k(t) = \left(X_i^k(t)\frac{\partial}{\partial t}\left(\ln \alpha_i^k(t)\right) + \sum_{j=1}^{N_k} V_{ji}^k(t) \right) dt + \beta_i^k(t)\, dW_k(t) \tag{45}$$

$$dV_{ij}^k(t) = \left(\beta_i^k(t)\beta_j^k(t) + V_{ji}^k(t)\frac{\partial}{\partial t}\left(\ln \alpha_i^k(t)\alpha_j^k(t)\right) \right) dt \tag{46}$$

In particular cases of assumed functional forms of $\sigma(t, T)$ it is easier to represent the state variables first and then to differentiate these special cases' equations.

Gaussian HJM pricing formulae

The foundations of Gaussian HJM model pricing of European contingent claims were laid by Brace, Musiela and Rutkowski (Brace, Musiela 1994, 1997; Musiela, Rutkowski 2005). For the purposes of this article we would need an explicit or semi-explicit formula for swaptions (to calibrate our volatility surface to) and additionally a formula for caps/floors – for the sake of speed and efficiency tests in MC simulations of xVA in different HJM models.

Recall dynamics of $f(t, T)$ and $P(t, T)$ under P from the previous section. Defining the forward price evaluated at time t of a zero coupon bond starting at T and maturing at T_1 to be $F_T(t,T1) = \frac{P(t,T_1)}{P(t,T)}$ and introducing a T. -maturity forward measure of $\mathrm{P}_T = \mathrm{E}(\mathrm{M}_T(\cdot))(T)\mathrm{P}$ where $\mathrm{E}(\cdot)$ is a stochastic exponential operator and the martingale inside is $\mathrm{M}_T(t) = \int_0^t\int_s^T \sigma(s,u)\,du\,dW(s)$. Under this forward measure we will have:

$$\begin{aligned} F_T(T,T_1) = F_T(t,T_1)\exp\Bigg(& \int_t^T\int_T^{T_1} \sigma(s,u)\,du\,dW_T(s) + \\ & -1/2\int_t^T \left| \int_T^{T_1} \sigma(s,u)\,du \right|^2 ds \Bigg) \end{aligned} \tag{47}$$

Let $\xi = g\left(Z_T^1, Z_T^2, \ldots, Z_T^n\right)$ be a European contingent claim at time T, where Z^i are asset prices which under P-measure follow a log-normal processes $dZ_t^i = Z_t^i\left(r_t dt + \eta_t^i \cdot dW_t\right)$. Define the forward price as $F_{Z_i}(t) = \frac{Z_i(t)}{P(t,T)}$ and by the same arguments as above we get:

$$\begin{aligned} Z_i(T) = F_{Z^i}(T,T) = F_{Z^i}(t,T)\exp\Bigg(& \int_t^T \left(\int_s^T \sigma(s,u)\,du + \eta_i(s) \right) dW_T(s) + \\ & -1/2\int_t^T \left| \int_s^T \sigma(s,u)\,du + \eta_i(s) \right|^2 ds \Bigg) \end{aligned} \tag{48}$$

As Brace and Musiela pointed out, the vector of random variables $\int_t^T\left(\int_s^T \sigma(s,u)\,du + \eta_i(s)\right) dW_T(s)$ is normally distributed with zero mean and variance matrix $\Delta_{ij} = Cov\left(\log Z_i(T), \log Z_j(T)\right)$ at t filtration, and therefore we have as follows.

Proposition 5.3 (Brace-Musiela). In M-factor HJM model with deterministic volatility function σ_i for i = 1,..., n the arbitrage price at time i of a European contingent claim $\xi = g\left(Z_T^1, Z_T^2, ..., Z_T^n\right)$ at maturity date T (where $g(\cdot)$ is a pay-off function and Z_t^i are the price processes of zero coupon bonds) equals:

$$\pi(t) = P(t,T)\int_{\mathrm{R}^M} g\left(\frac{Z_t^1 \mathrm{N}^M\left(x+\theta_1\right)}{P(t,T)\mathrm{N}^M\left(x\right)}, \ldots, \frac{Z_t^n \mathrm{N}^M\left(x+\theta_n\right)}{P(t,T)\mathrm{N}^M\left(x\right)}\right)\mathrm{N}^M(x)dx \tag{49}$$

where N^M denotes the standard M-dimensional normal density

$$\mathrm{N}^M(x) = \left(\frac{1}{2\pi}\right)^{M/2} e^{-|x|^2/2}$$

and we define vectors $\theta_1, \theta_2, \ldots, \theta_n \in \mathrm{R}^M$ such that the scalar product of every pair θ_i and θ_j for $i \neq j$, is given by:

$$\theta_i \cdot \theta_j = \int_t^T \sigma_i\left(u,T\right)\cdot \sigma_j\left(u,T\right)du \tag{50}$$

Suppose now that we have to price a swaption with maturity T and cash flow equivalents (cf. Henrard 2014). $c_i = K\delta_i$ for i = 1,...,$n-1$ and $c_n = 1 + K\delta_n$, where K is a strike of the swaption and $\delta_i = T_i - T_{i-1}$ is a year fraction of a particular interest period i. Since the cash flows have different times of scheduled appearance after maturity, the option has to be adjusted to swap annuity measure. The payout function from receiver swaption is therefore $g\left(Z_1, Z_2, \ldots, Z_n\right) = (\sum_{i=1}^{n} c_i Z_i - 1)^+$. It may be treated as a basket option of all the cash-flows executed as one set. The fixed leg is represented by $K\delta_i$ for i = 1,...,n, whereas the floating leg is replaced by the equivalent in a single curve environment – notional paid at the beginning and returned at the end of the underlying swap schedule.

The direct application of proposition 5.3 on this pay-off of receiver swaption (RS) yields in the following arbitrage price (payment dates of the underlying swap are T_0, T_1,...,T_n and the option expires at T_0):

$$RS_{HJM_g}(t) = \int_{\mathrm{R}^M}\left(\sum_{i=1}^{n} c_i P(t,T_i)\mathrm{N}^M\left(x+\theta_i\right) - P(t,T_0)\mathrm{N}^M\left(x\right)\right)^+ dx \tag{51}$$

As Henrard (2003) showed for a one factor model, in the consequence of the above mentioned relation and the evaluation of signs of the components of the integral, we have:

$$RS_{HJM_g}(t) = \sum_{i=1}^{n} c_i P(t,T_i)\mathrm{N}\left(x+\theta_i\right) - P(t,T_0)\mathrm{N}\left(x+\theta_0\right) \tag{52}$$

where vector of θ_i is taken from the rank one Δ matrix of covariance:

$$\theta_i^2 = \int_t^{T_0}\left|\int_{T_0}^{T_i}\sigma(s,u)du\right|^2 ds \tag{53}$$

and x is the unique solution of:

$$\sum_{i=1}^{n} c_i P(t,T_i)e^{-\frac{1}{2}\theta_i^2 - \theta_i x} = P(t,T_0)e^{-\frac{1}{2}\theta_0^2 - \theta_0 x} \tag{54}$$

Analogously, for a payer swaption (PS) we would have the same conditions for x and θ_i^2 but the signs are changed in the pay-off function and inside the cumulative standard normal distributions, so we have:

$$PS_{HJM_g}(t) = P(t,T_0)\,\mathrm{N}\,(-x-\theta_0) - \sum_{i=1}^{n} c_i P(t,T_i)\,\mathrm{N}\,(-x-\theta_i) \tag{55}$$

The idea of a cash-flow equivalent of the swap is very useful and it will allow to adopt these single curve results to the desired multi-curve framework. Henrard (2014) suggests to use a deterministic multiplicative coupon spread β^j of the forward rate F^j (where j is tenor of j-IBOR) over the risk-free rate and its representation in the form of discounting factors, defined as follows:

$$\beta_t^j(u,v) = \left(1+\delta_{t_v-t_u} F^j(u,v)\right)\frac{P(t,v)}{P(t,u)} \tag{56}$$

Hence the present value of j-IBOR payment is given by:

$$P(t,v)\delta_{t_v-t_u} F^j(u,v) = P(t,v)\left(\beta_t^j(u,v)\frac{P(t,u)}{P(t,v)} - 1\right) = \beta_t^j(u,v)P(t,u) - P(t,v) \tag{57}$$

This leads to cash-flow representation of j-IBOR in a multiplicative spread world because the fair value is a sum of discounted positive flow of $\beta_t^j(u,v)$ at u-time and negative cash-flow of -1 at time v.

Table 2

Cash-flow equivalents of an underlying IRS in a 3-year maturity 3-year tenor swaption in single and multi-curve variants

T	**Single curve**			**Floating leg**			**Fixed leg**			**Total multi-curve**		
	$\boldsymbol{i}$	$\boldsymbol{\delta_i}$	$\boldsymbol{c_i}$	**j**	$\boldsymbol{\delta_j}$	$\boldsymbol{c_j}$	**k**	$\boldsymbol{\delta_k}$	$\boldsymbol{c_k}$	$\boldsymbol{\bar{i}}$	$\boldsymbol{\delta_x}$	$\boldsymbol{d_{\bar{i}}}$
3	0		-1	0		$-\beta_0$	0			0		$-\beta_0$
3.5				1	0.5	$1-\beta_1$				1		$1-\beta_1$
4	1	1	$K\delta_1$	2	0.5	$1-\beta_2$	1	1	$K\delta_1$	2	1	$1-\beta_2+K\delta_2$
4.5				3	0.5	$1-\beta_3$				3		$1-\beta_3$
5	2	1	$K\delta_2$	4	0.5	$1-\beta_4$	2	1	$K\delta_2$	4	1	$1-\beta_4+K\delta_4$
5.5				5	0.5	$1-\beta_5$				5		$1-\beta_5$
6	3	1	$1+K\delta_3$	6	0.5	1	3	1	$K\delta_3$	6	1	$1+K\delta_6$

The logic is easily expendable to a stream of floating rates in the IRS as an underlying of a swaption in a multi-curve set-up. One should be very careful when aggregating the cash-flow equivalents though. The example of cash flow schedule d_i for a receiver swaption with maturity of 3 years with underlying swap being 3 year IRS annual fixed (strike = K) for 6 month floating rate in single curve and multi-curve multiplicative spread environments is shown in Table 2.

Therefore we may implement a multi-curve approach to the cap/floor formulae via modification on cash-flow schedule d_i:

$$RS_{HJM_g,m}(t)=\sum_{\bar{i}=0}^{\bar{n}} d_{\bar{i}} P\left(t,\overline{T}_i\right) N\left(x+\theta_{\bar{i}}\right) \tag{58}$$

$$PS_{HJM_g,m}(t)=\sum_{\bar{i}=0}^{\bar{n}} d_{\bar{i}} P\left(t,\overline{T}_i\right) \mathrm{N}\left(-x-\theta_{\bar{i}}\right) \tag{59}$$

where x is the unique solution of:

$$\sum_{\bar{i}=0}^{\bar{n}} d_{\bar{i}} P\left(t,\overline{T}_i\right) e^{-\frac{1}{2}\theta_{\bar{i}}^2-\theta_{\bar{i}}x}=0 \tag{60}$$

Brace and Musiela (1997) developed semi analytic formulae for caps and floors (in a single curve world). The fair value of a cap at time-t, with start at $T=T_0$, strike K for IBORs $L(T_{i-1})$ paid in at times $T_0, T_1, \ldots T_{i-1}, T_i, \ldots, T_n$ for $i = 1\ldots., n$ is (cf. Beyna 2013):

$$CAP_{HJM_g}(t)=\sum_{i=0}^{n-1}\left(P\left(t,T_i\right)\mathrm{N}\left(x_i\right)-\left(1+K\delta\right)P\left(t,T_{i+1}\right)\mathrm{N}\left(x_i-\vartheta_i\right)\right) \tag{61}$$

where:

$$\vartheta_i^2=\int_t^{T_i}\left|\int_{T_i}^{T_{i+1}}\sigma(s,u)\,du\right|^2 ds \tag{62}$$

$$x_i=-\frac{1}{\vartheta_i}log\frac{P\left(t,T_i\right)}{\left(1+K\delta\right)P\left(t,T_{i+1}\right)}+1/2\vartheta_i \tag{63}$$

The following floor price stems from a put-call parity (with the same definitions of the auxiliary variables ϑ_i and x):

$$FLR_{HJM_g}(t)=\sum_{i=0}^{n-1}\left(\left(1+K\delta\right)P\left(t,T_{i+1}\right)\mathrm{N}\left(-x_i+\vartheta_i\right)-P\left(t,T_i\right)\mathrm{N}\left(-x_i\right)\right) \tag{64}$$

Note that a cap price is a sum of caplets or calls for a particular IBORs with a common strike for all the options involved. Also it is important to notice that an IBOR's cash flow equivalence here is 1 paid at T_i and $\left(1+K\delta\right)$received at T_{i+1}. Following the same logic as with cash-flow equivalents of swaptions, in the multi-curve case of cap/floor it would suffice to change only the initial payment 1 to $\beta_{t,i}^j\left(T_{i-1},T_i\right)$ to get the proper pricing formulae in this environment:

$$CAP_{HJM_g,m}(t)=\sum_{i=0}^{n-1}\left(\beta_{t,i}^j\left(T_{i-1},T_i\right)P\left(t,T_i\right)N\left(x_i\right)-\left(1+K\delta\right)P\left(t,T_{i+1}\right)N\left(x_i-\vartheta_i\right)\right) \tag{65}$$

$$FLR_{HJM_g,m}(t)=\sum_{i=0}^{n-1}\left(\left(1+K\delta\right)P\left(t,T_{i+1}\right)N\left(-x_i+\vartheta_i\right)-\beta_{t,i}^j\left(T_{i-1},T_i\right)P\left(t,T_i\right)N\left(-x_i\right)\right) \tag{66}$$

We have cited and developed semi-analytic (for swaptions) and analytic (for caps and floors) pricing formulae in single and multi-curve versions which are relatively easy to evaluate in every step of Monte Carlo simulation, should one know the Cheyette form of volatility surface and find a fast method of solving for x in swaptions' pricing.[23]

Volatility structure parametrisations

Recall from (37) that the volatility function in the M-factor HJM Cheyette model with N_k summands in each factor k for $k = 1,\dots, M$ and $i = 1,\dots, N$ is of a form:

$$\sigma(t,T)=\begin{pmatrix}\sigma^1(t,T)\\ \sigma^2(t,T)\\ \vdots\\ \sigma^k(t,T)\\ \vdots\\ \sigma^M(t,T)\end{pmatrix}=\begin{pmatrix}\sum_{i=1}^{N_1}\frac{\alpha_i^1(T)}{\alpha_i^1(t)}\beta_i^1(t)\\ \sum_{i=1}^{N_2}\frac{\alpha_i^2(T)}{\alpha_i^2(t)}\beta_i^2(t)\\ \vdots\\ \sum_{i=1}^{N_k}\frac{\alpha_i^k(T)}{\alpha_i^k(t)}\beta_i^k(t)\\ \vdots\\ \sum_{i=1}^{N_M}\frac{\alpha_i^M(T)}{\alpha_i^M(t)}\beta_i^M(t)\end{pmatrix} \tag{67}$$

The volatility function at each summand's level is separable into time and lifetime of the option dependent factors. Generally speaking function $\alpha(\cdot)$ is responsible for handling the remaining lifetime (current maturity $T-t$ of the option at time t) impact on volatility, whereas $\beta(\cdot)$ is strictly a time--dependent function.

Table 3 summarises the most frequent choices of $\sigma(T, t)$ components in this class. Its quite striking that $\beta(\cdot)$ is usually in the form of polynomial of different degrees[24] whereas $\alpha(\cdot)$ is an exponential function with constant or piecewise constant parameters in the exponents. If the latter is used, as for example in Beyna and Wystup (2010), the number of nodes taken into account varies from 1–5, contributing to 2–6 different κ parameters per summand. It has been tested there that the more nodes, the better the calibration fit to the market reality but, of course, the bigger the computational time of this part of valuation and simulation.

[23] For the particular volatility function implementation we would have to evaluate double integrals of θ_i^2 and ϑ_i^2 analytically.

[24] Usually $p = \{1, 2\}$, but in the Ho-Lee model we have $p = 0$.

Table 3

Common forms of volatility function within the Cheyette class

Volatility function form	M	N	Name/author(s)
$\sigma^1(t,T)=c$	1	$N_1=1$	Ho-Lee
$\sigma_i^1(t,T)=\beta_i^1(t)\exp\left(-\int_t^T \kappa_i(u)du\right)$	1	$N_1\in \mathrm{N}$	General exponential Cheyette
$\sigma_i^1(t,T)=\beta_i^1\exp(-(T-t)\kappa_i)$	1	$N_1\in \mathrm{N}$	General Hull-White
$\sigma_1^1(t,T)=c\exp(-(T-t)a)$	1	$N_1=1$	Hull-White
$\sigma_1^1(t,T)=c$ $\sigma_2^1(t,T)=Poly^1(t;1)\exp\left(-\int_t^T \underline{\kappa}_2^{1,node}(u)du\right)$	1	$N_1=2$	HW Beyna
$\sigma^1(t,T)=c+Poly^1(t;p)\exp(-(T-t)\kappa_2^1)$ $\sigma^2(t,T)=Poly^2(t;p)\exp(-(T-t)\kappa_1^2)$ $\sigma^3(t,T)=Poly^3(t;p)\exp(-(T-t)\kappa_1^3)$	3	$N_1=2$ $N_2=1$ $N_3=1$	3-factor exponential

where:
$\sigma_i^j(t,T)$ is a volatility function i-th summand for the j-th factor,
$Poly^j(t;p)$ stands for a polynomial function of degree p for the j-th factor,
$\underline{\kappa}_i^{j,node}$ is a parameter of piecewise constant function $\underline{\kappa}(\cdot)$ for a certain *node*.

In the following subsections, we discuss different versions of this class in the context of their particular dynamics (specialising the general formulae from section 5.2.1 and developing special analytical formulae for the quantities needed in pricing swaptions and caps/floors, namely θ_i^2 and ϑ_i^2.

Ho-Lee. The simplest way of coping with volatility is assuming a flat surface at constant level. Using $\sigma_1^1(t,T)=c$ (hence $\alpha_1^1(t)=1$, $\alpha_1^1(T)=1$ and $\beta_1^1(t)=c$) in the general formulae in the Cheyette model we get:

$$A_1^1(x)=\int_0^x 1ds = x \text{ and } V_{11}^1=\int_0^t c^2 ds = c^2 t \tag{68}$$

$$X_1^1(t)=\int_0^t c^2(t-s)ds+\int_0^t c\,dW(s)=\int_0^t c\,dW(s)+1/2c^2t^2 \tag{69}$$

$$f(t,T)=f(0,T)+\int_0^t c dW(s)+1/2\,c^2t^2+(T-t)c^2t \tag{70}$$

and the dynamics of the one state variable is:

$$dX_1^1(t)=c^2tdt+cdW(t) \tag{71}$$

and our auxiliary variables for pricing:

$$\theta_i^2 = \int_0^{T_0} \left| \int_{T_0}^{T_i} \sigma(s,u)\,du \right|^2 ds = \int_0^{T_0} c^2 (T_i - T_0)^2\, ds = c^2 (T_i - T_0)^2 T_0 \tag{72}$$

$$\vartheta_i^2 = \int_0^{T_i} \left| \int_{T_i}^{T_{i+1}} \sigma(s,u)\,du \right|^2 ds = c^2 (T_{i+1} - T_i)^2 T_i \tag{73}$$

These results may be directly used in the derivatives formulae from the previous section.

Hull-White. It is easy to notice that a choice of $\alpha_1^1(t) = e^{-at}$, $\alpha_1^1(T) = e^{-aT}$ and $\beta_1^1(t) = c$ in the Cheyette class would lead to $\sigma_1^1(t,T) = ce^{-(T-t)a}$, which is well known in the literature from the Hull-White model. Analogously to the Ho-Lee model, we have the following:

$$A_1^1(x) = \int_0^x e^{-as}\,ds = \frac{1}{a}\left(1 - e^{-ax}\right) \tag{74}$$

$$V_{11}^1(t) = c^2 \int_0^t e^{-2a(t-s)}\,ds = \frac{c^2}{2a}\left(1 - e^{-2at}\right) \tag{75}$$

$$\begin{aligned} X_1^1(t) &= \frac{c^2}{2a} \int_0^t e^{-a(t-s)} \frac{1 - e^{-at} - 1 + e^{-as}}{e^{-as}}\,ds + c\int_0^t e^{-a(t-s)}\,dW(s) = \\ &= \frac{c^2}{2a^2}\left(e^{-2at} - 2e^{-at} + 1\right) + c\int_0^t e^{-a(t-s)}\,dW(s) \end{aligned} \tag{76}$$

$$f(t,T) = f(0,T) + e^{-a(T-t)}\left(X_1^1(t) + c^2/2\,a^2\left(e^{-a(T-t)} - 1\right)\left(e^{-2at} - 1\right)\right) \tag{77}$$

The variables for pricing are given by:

$$\begin{aligned} \theta_i^2 &= \int_0^{T_0} \left| \int_{T_0}^{T_i} \sigma(s,u)\,du \right|^2 ds = \frac{c^2}{a^2} \int_0^{T_0} \left(e^{a(s-T_i)} - e^{a(s-T_0)}\right)^2 ds = \\ &= c^2/2\,a^3\left(e^{2a(T_0-T_i)} - e^{-2aT_i} - e^{-2aT_0} - 2e^{a(T_0-T_i)} + 2e^{-a(T_i+T_0)} + 1\right) \end{aligned} \tag{78}$$

To calculate ϑ_i^2 we can use formula for θ_i^2 and change T_i for T_{i+1} and T_0 for T_i. Having two free parameters to calibrate, though, is not rich enough to replicate market shapes of volatility surface in this model. Therefore some authors posit an extension within the broadly defined general Hull-White model.

Extended Hull-White (two summands). For a better fit in a general Hull-White environment, we can use two summands of similar form, i.e.: $\sigma(t,T) = \sigma_1^1(t,T) + \sigma_2^1(t,T) = \sigma_1 e^{-\kappa_1(T-t)} + \sigma_2 e^{-\kappa_2(T-t)}$, with $\sigma_1, \sigma_2 > 0$. Then we would have $\alpha_i^1(x) = e^{-\kappa_i x}$, $\beta_i^1(x) = \sigma_i$, for $i = 1, 2$ and a one-factor model as $M = 1$. This model yields a bit more complicated closed formulae but is still tractable:[25]

[25] We skip upper-script as we have only one factor model here.

$$A_i(x)=\int_0^x e^{-\kappa_i s}ds=1/\kappa_i\left(1-e^{-\kappa_i x}\right) \tag{79}$$

$$V_{ij}(t)=\frac{\sigma_i\sigma_j}{\kappa_i+\kappa_j}\left(1-e^{(\kappa_i+\kappa_j)t}\right) \tag{80}$$

$$dX_i(t)=\left(-\kappa_i X_i(t)+\sum_{j=1}^{2}V_{ij}(t)\right)dt+\sigma_i dW(t) \tag{81}$$

$$f(t,T)=f(0,T)+\sum_{i=1}^{2}e^{\kappa_i(T-t)}\left(X_i(t)+\sum_{j=1}^{2}\frac{1-e^{-\kappa_j(T-t)}}{\kappa_j}V_{ij}(t)\right) \tag{82}$$

The term used in pricing of swaptions is then (for ϑ_i^2 we follow the same logic as in the previous section):

$$\begin{aligned}\theta_i^2 &= \sum_{x=1}^{2}\left(\frac{\sigma_x^2}{2\kappa_x^3}\left(e^{2\kappa_x(T_0-T_i)}-e^{-2\kappa_x T_i}-2e^{\kappa_x(T_0-T_i)}+2e^{-\kappa_x(T_0+T_i)}+1-e^{-2\kappa_x T_0}\right)\right)+\\ &+\frac{2\sigma_1\sigma_2}{\kappa_1\kappa_2(\kappa_1+\kappa_2)}\left(e^{(\kappa_1+\kappa_2)(T_0-T_i)}-e^{-(\kappa_1+\kappa_2)T_i}-e^{\kappa_1(T_0-T_i)}+e^{-\kappa_1 T_i-\kappa_2 T_0}+\right.\\ &\left.-e^{\kappa_2(T_0-T_i)}+e^{-\kappa_1 T_0-\kappa_2 T_i}+1-e^{-(\kappa_1+\kappa_2)T_0}\right)\end{aligned} \tag{83}$$

Beyna and Wystup. One of the other possibilities is to use two summands of volatility (constant part from Ho-Lee and general exponential Cheyette) allowing a function in integral in the exponent of one of them to be piecewise constant, instead of a constant as in the pure Hull-White version. Additionally, we can experiment with the polynomial forms of β_2^1 and take $\beta_2^1 = at + b$. The key challenge for implementation and optimisation is to choose the optimal number of nodes on the maturity timescale (hence, to determine the number of constant segments of $\underline{\kappa}(\cdot)$.

Beyna and Wystup proposed

$$\sigma(t,T)=c+(at+b)\exp\left(-\int_t^T\underline{\kappa}^{node}(u)du\right)$$ where $\underline{\kappa}(x)$ is defined as follows:[26]

$$\underline{\kappa}^{node}(x)=\sum_{i=1}^{node+1}\kappa_i\left(\mathrm{H}(x-n_{i-1})-\mathrm{H}(x-n_i)\right) \tag{84}$$

where *node* represents the number of internal nodes or maturities splitting the (0, T) segment into *node* + 1 segments, n_i is an i-th point on this segment s.t. n_0 = 1 and n_{node+1} = T. $H(\cdot)$ stands for a Heaviside function.

Partition $\pi=\{n_0=0,n_1,\dots,n_{node},n_{node+1}=T\}$ should depend on market data and the liquidity of different segments of tenors. The authors proposed *node* = 1 and *node* = 5 for different fit qualities.

[26] Each summand in $\underline{\kappa}$'s definition is a boxcar function defined by the difference in Heaviside functions multiplied by κ_i corresponding to a particular sub-segment $n_{i-1} - n_i$.

In the general form of this piecewise function, the integral in the exponent of β is:

$$\int_t^T \underline{\kappa}^{node}(u)\,du = \sum_{i=1}^{node+1} \int_t^T \kappa_i \left(\mathrm{H}\left(u-n_{i-1}\right) - \mathrm{H}\left(u-n_i\right)\right) du =$$

$$= \sum_{i=1}^{node+1} \kappa_i \left(\left(T-n_{i-1}\right) H\left(T-n_{i-1}\right) - \left(t-n_{i-1}\right) H\left(t-n_{i-1}\right) + \right. \tag{85}$$

$$\left. -\left(T-n_i\right) H\left(T-n_i\right) + \left(t-n_i\right) H\left(t-n_i\right) \right)$$

From that representation it is easier to extract the form of $\alpha(\cdot)$:

$$\alpha(x) = \exp\left(-\sum_{i=1}^{node+1} \kappa_i \left(\left(x-n_{i-1}\right) \mathrm{H}\left(x-n_{i-1}\right) - \left(x-n_i\right) \mathrm{H}\left(x-n_i\right) \right) \right) \tag{86}$$

It is clear that analytical integration with such exponents as the above may be very cumbersome and tedious. Therefore, numerical integration is the only sensible solution, which also means that the analytical forms of θ_i^2 and ϑ_i^2 are not available as well, hence the calibration and pricing would be more demanding.

Three-factor exponential. Beyna, Chiarella and Kang (2012) proposed a three-factor model within the Cheyette class, which they used as a benchmark for the comparative performance of other models. They suggest using polynomials of degree 1 (after some empirical tests), so that as indicated in Table 3, their three-factor exponential model has the volatility function vector of the form:[27]

$$\sigma(t,T) = \begin{pmatrix} \sigma^1(t,T) \\ \sigma^2(t,T) \\ \sigma^3(t,T) \end{pmatrix} = \begin{pmatrix} c + \left(a^1 t + b^1\right)\exp\left(-\left(T-t\right)\kappa^1\right) \\ \left(a^2 t + b^2\right)\exp\left(-\left(T-t\right)\kappa^2\right) \\ \left(a^3 t + b^3\right)\exp\left(-\left(T-t\right)\kappa^3\right) \end{pmatrix} \tag{87}$$

Since the first factor is a sum of constant c and another, common in form with all three factors, extended Hull-White term with linear function β, we have here four state variables, two for the first factor (X_1^1 and X_2^1), one for the second factor (X_1^2) and one for the third (X_1^3). Hence for the first factor (as in the case of extended Hull-White with two summands) we will have four cumulative quadratic terms: $\left\{V_{11}^1, V_{12}^1, V_{21}^1, V_{22}^1\right\}$.

$$V_{11}^1(t) = c^2 t \tag{88}$$

$$V_{12}^1(t) = V_{21}^1(t) = \frac{c}{(\kappa^1)^2}\left(-a^1 + b^1\kappa^1 + e^{-\kappa^1 t}\left(a^1 - b^1\kappa^1 t\right) + a^1\kappa^1 t\right) \tag{89}$$

$$V_{22}^1(t) = \frac{1}{4(\kappa^1)^3}\left(2(b^1\kappa^1)^2 + 2b^1a^1\kappa^1\left(2\kappa^1 t - 1\right) + \right.$$
$$\left. -e^{-2\kappa^1 t}\left((a^1)^2 - 2b^1a^1\kappa^1 + 2(b^1\kappa^1)^2\right) + (a^1)^2(1 + 2\kappa^1 t\left(\kappa^1 t - 1\right))\right) \tag{90}$$

[27] All of the upper-scripts represent factor numbering, not: powers.

$$V_{11}^{x=\{2,3\}}(t)=\frac{1}{4(\kappa^x)^3}\Big(2(b^x\kappa^x)^2+2b^xa^x\kappa^x\left(2\kappa^xt-1\right)+ \\ -e^{-2\kappa^xt}\left((a^x)^2-2b^xa^x\kappa^x+2(b^x\kappa^x)^2\right)+(a^x)^2\left(1+2\kappa^xt\left(\kappa^xt-1\right)\right)\Big) \tag{91}$$

Therefore the dynamics of X_i^j are:

$$dX_1^1(t)=\left(\sum_{k=1}^{2}V_{1k}^1(t)\right)dt+cdW^1(t) \tag{92}$$

$$dX_2^1(t)=\left(-\kappa^1X_2^1(t)+\sum_{k=1}^{2}V_{2k}^1(t)\right)dt+\left(a^1t+b^1\right)dW^1(t) \tag{93}$$

$$dX_1^2(t)=\left(-\kappa^2X_1^2(t)+V_{11}^2(t)\right)dt+\left(a^2t+b^2\right)dW^2(t) \tag{94}$$

$$dX_1^3(t)=\left(-\kappa^3X_1^3(t)+V_{11}^3(t)\right)dt+\left(a^3t+b^3\right)dW^3(t) \tag{95}$$

Beyna, Chiarella, and Kang (2012) developed the price formula of a European caplet in the M-factor HJM model based on the previous works of Brace and Musiela (1994), which is similar to the one-factor version, but the volatility function has to incorporate its multifactorial structure:

$$\vartheta_i^2=\sum_{k=1}^{M}\int_t^{T_i}\left|\int_s^{T_{i+1}}\sigma^k(s,u)\,du-\int_s^{T_i}\sigma^k(s,u)\,du\right|^2ds \tag{96}$$

Extending this result to caps as in (62), we have the following arbitrage-free price of a cap:

$$CAP_{HJM_g,M}(t)=\sum_{i=0}^{n-1}\left(P(t,T_i)\,\mathrm{N}\left(x_i\right)-(1+K\delta)P(t,T_{i+1})\,\mathrm{N}\left(x_i-\vartheta_i\right)\right) \tag{97}$$

where:

$$\vartheta_i{}^2=\sum_{k=1}^{M}\int_t^{T_i}\left|\int_{T_i}^{T_{i+1}}\sigma^k(s,u)\,du\right|^2ds \tag{98}$$

$$x_i=-\frac{1}{\vartheta_i}\log\frac{P(t,T_i)}{(1+K\delta)P(t,T_{i+1})}+1/2\vartheta_i \tag{99}$$

Similarly, we may develop the formula for a floor and extend these into a multi-curve environment via the same arguments as previously, although the analytic calculation of ϑ_i^2 for even a three-factor model is arduous and it may not give any sensible calculation time advantage over a numerical integration of ϑ_i^2. Unfortunately for our purpose – XVA calculation – and to the best of our knowledge, there is no closed formula for swaptions pricing in more than one factor HJM model (Green 2015).

6 Monte Carlo implementation

6.1 Calibration of volatility surface

The importance of the calibration of the volatility function in Markov quasi-Gaussian models cannot be underestimated, because it is a key ingredient of the behaviour of instantaneous forward rates – a cornerstone of HJM. Thanks to the courtesy of Thomson Reuters (and Tullet Prebon), we have been given access to the price data of a wide list of interest rate instruments (daily closing prices) in PLN and EUR in the period 2014–2017, which makes the multiple-curves construction and volatility surface calibration possible. From the non-linear derivatives domain, the list includes at-the-money-forward straddle swaptions for different lifetimes of options (up to 10Y) and different underlying IRS tenors (up to 10Y as well).[28] It is worth noting that swaptions are still quoted in Black76 log-normal implied volatilities, not Bachelier normal volatilities, which are in the same order as the volatility surface resultant from HJM calibration. Many authors[29] suggest to perform calibration in relation to market volatilities (log-normal) rather than derivatives prices, which is the route we follow, but the quality checks of our optimization will be performed both in Black's volatilities and in nominal prices per 1 unit of notional. It should also be underlined that some market data for swaptions volatilities maybe somehow distorted by the dichotomy of negative rates being possible in reality and not possible in the log-normal world. The closer to negative strikes the higher Black's volatilities are, to the point where the market ceases to quote any volatility on a certain swaption with ATMF's strike being negative. This distortion, producing odd volatility surface shapes (hard to calibrate to) may present a challenge in developed markets where interest rates were, are or can be negative in the nearest future.

Suppose O is the space of all possible parameter sets Θ for a certain form (model) of volatility surface $\sigma(t,T)$. Let V_{mkt}^{BS} be the market quoted log-normal volatility of a swaption and $\mathrm{V}_{imp,\Theta}^{BS}$ be implied log-normal volatility calculated from a price of swaption indicated by the model from the set of given particular parameters Θ. We assume equal weight of all instruments in the calibration, although we choose to take the most liquid time pairs of: life time of option $t = \{1, 2, 5, 8, 10\}$ vs. tenor of an underlying IRS $T - t = \{1, 2, 5, 8, 10\}$, namely 25 swaptions per one optimization date.[30] As a measure of distance we propose to take the square of differences in volatilities, hence our optimisation problem may be defined as:

$$\begin{aligned} &\inf_{\Theta \in O} \sum_{i=1}^{25} \left(\mathrm{V}_{mkt,i}^{BS} - \mathrm{V}_{imp,\Theta,i}^{BS}\right)^2 \\ &s.t.\ \forall_i \mathrm{V}_{imp,\Theta,i}^{BS} \geq 0 \end{aligned} \tag{100}$$

where i stands for a particular instrument with expiry and tenor of the underlying from 5 × 5 grid.

The literature on numerical tractability and the practical challenges involved in the optimisation of Cheyette class models, especially with the necessary condition for at least semi-analytical pricing

[28] Unfortunately caps and floors market data especially for PLN are erratic and not-reliable, therefore we haven't used them in any calibration.

[29] Cf. Beyna (2013), Beyna and Wystup (2010), Hirsa (2012), Kienitz and Caspers (2017).

[30] One may also consider taking more instruments, as formally there are 160 swaptions quoted on the market. This approach multiplies calibration time, while not giving better qualitative results.

formulae for swaptions to be known, is rather scarce.[31] The objective function as defined in (101) is non-linear, highly complex, and far from convex or concave, which implies that there are many possible local minima. The sensible solution is therefore (cf. Hirsa 2012) to try several different Θ sets for a particular minimization procedure as starting points. Beyna and Wystup in their works (Beyna, Wystup 2010; Beyna 2013) showed that in the case of the Cheyette class, a simple Newton algorithm is not converging despite step adjustments,[32] the Brent (or Powell) algorithm, which requires no derivatives calculation, is also not suitable because of the increasing computational requirements for high dimensions and the length of the process. The Nelder-Mead algorithm seems to give good results if the initial set Θ happens to be chosen close to the global minimum. The generalisation of this algorithm (via a synthetic temperature optimisation parameter), which is called simulated annealing, is also effective but calibrating the procedure as such may be lengthy as well.

We have implemented here an extended version of the Nelder-Mead algorithm in the sense that we will start the procedure by brute force and calculate objective function values on a hyper-cube of predefined parameter sets (i.e. for 5 parameters and 4 ranges in each of the parameters we will calculate 625 objective function values). The set which would give the lowest objective function value is then used for further polishing by the Nelder-Mead algorithm. For a given date's market data of swaption volatilities we also have to have proper multi-curves bootstrapped previously $C_P^d(t), C_F^{6M}(t)$ and corresponding multiplicative spreads β_t^{6M} (as defined in (59)). Moreover we have to implement the calculation routine to get θ_i^2 s given by $\theta_i^2 = \int_0^{T_0} \left| \int_{T_0}^{T_i} \sigma(s,u)\,du \right|^2 ds$.

For some simple models it can be one analytically, butgenerally we have to integrate it numerically.[33] Then in order to calculate the objective function's value for a certain set of parameters Θ follow the routine:

1. For all instruments calculate (a)–(e):
   a) ATMF forward rate K,
   b) create the schedule of payments $d_{\bar{i}}$ including spread β as in (Table 2),
   c) solve numerically[34] for x the equation (60):

$$\sum_{\bar{i}=0}^{\bar{n}} d_{\bar{i}} P\left(t, \overline{T}_i\right) e^{-\frac{1}{2}\theta_{\bar{i}}^2 - \theta_{\bar{i}} x} = 0$$

   d) calculate the payer swaption value using (59):

$$PS_{HJM_{g},m}(t) = \sum_{\bar{i}=0}^{\bar{n}} d_{\bar{i}} P\left(t, \overline{T}_i\right) \mathrm{N}\left(-x - \theta_{\bar{i}}\right)$$

   e) using the price obtained in (d), solve for implied log-normal volatility $\mathrm{V}_{imp,\Theta}^{BS}$ and compare with corresponding market volatility to calculate the squared difference between the two.
2. Sum all of the squared differences calculated in 1.

[31] Cf. Beyna (2013), Henrard (2014), Andersen and Piterbarg (2010).

[32] A particular difficulty with the Newton algorithm is the necessity of calculation first and second order derivatives in the form of Jacobi and Hessian matrices, which is not available in analytical form for swaptions.

[33] Note that this is not a simple double integral as we have to acknowledge the squared inner integral. For simple models the advantage of analytical integration over numerical is of a factor 1.5, as empirical author's tests shown.

[34] It is a nice, monotone function (cf. Henrard 2014) so we may use *fsolve* cheaply.

This is quite time-consuming and depending on the particular form of the Cheyette model, the number of instruments used and hardware, one objective function evaluation may take some 0.2–0.8 sec.

In the evaluation of the calibration results, we used quite common measures given in pricing software and cited below after Beyna (2013). In order to calculate the value of the measures and evaluate, we need residuals in prices and volatilities distributions and their basic statistics (means and standard deviations: σ_{Px}, σ_{Vol}). The list of measures includes:

– price bias – the percentage of the absolute relative price differences exceeding 30%,
– volatility bias – the percentage of the absolute relative volatility differences exceeding 30%,
– mean of all price residuals,
– mean of all volatility residuals,
– all volatility residuals' bound,
– all price residuals' bound,
– percentile of the volatility residuals within $2\sigma_{Vol}$,
– percentile of the price residuals within $2\sigma_{Px}$,

where M1–M4 constitute measures used in core conditions and M5–M8 – in secondary conditions. Core conditions are: $M1 \le 30\%$ $M2 \le 30\%$, $M3 \le \sigma_{Px}$ and $M4 \le \sigma_{Vol}$. Secondary conditions are: $M5 \le 3\sigma_{Vol}$, $M6 \le 3\sigma_{Px}$, $M7 \ge 0.9$ and $M8 \ge 0.9$. A particular calibration will be considered *Good* if all core and secondary conditions are met, *Passed* if all core conditions are met and at most one of the secondary is breached, and *Failed* – in all other cases. Examples of quality checks for 20 calibrations for Beyna and Wystup and Hull-White two summands models are given in Tables 2 and 3 in the Appendix. We have chosen several Cheyette class models for testing:

Mod1: $(c+bt)\exp(-a(T-t))$ – version of general HW with Poly(1) of $\beta(\cdot)$,
Mod2: $c\exp(-a(T-t))$ – classic Hull-White (HW),
Mod3: $c_1\exp(-a_1(T-t))+c_2\exp(-a_2(T-t))$ – two summands of classic HW,
Mod4: $(c+bt)\exp\left(-\int_t^T \underline{\kappa}(u)\,du\right)$ – version of general HW with Poly(1) of $\beta(\cdot)$ and piecewise constant function $\underline{\kappa}(\cdot)$ with one inner node, $n_1 = 4$,
Mod5: $d+(c+bt)\exp\left(-\int_t^T \underline{\kappa}(u)\,du\right)$ – same as Mod5 but with additional constant summand *d*.

Table 4 in the Appendix exhibits some results of the calibration with satisfactory quality for three selected dates in the period 2014–2017. The general quality of the calibrations for 20 selected dates measured by the average objective function value at the termination of the calibration procedure and its standard deviation may be found in Table 5 in the Appendix.

6.2 Calculating exposures and CVA

Calculations of expected exposures may be divided into five technical parts:

1. Generation of instantaneous forward rates in *J* paths and *K* time steps in each path using a discretised version of a particular model.

2. Reconstruction of discounting curves, forward curves and spread curves[35] i.e.: $B_F^{3M,d}(t)$, $B_F^{6M,d}(t)$ for each path and each time step in the path.

3. Valuation of a certain list of I instruments under each state resulting in 'CVA cube' $\boxplus = \{(V_i, S_j, t_k)\}$.

4. Application of netting, collateralisation or other rules that may govern exposure calculation for a certain counterparty.

5. Calculation of expected exposure and quantile potential future exposure profiles.

In subsequent sections, we present details from the implementation of Hull-White with the two summands (henceforth: HW2) model. The general mechanisms and results prove to be the same for all other models considered here.

Discretisation and necessary adjustments of HW2

The detailed description of the HW2 model allows for almost straightforward MC implementation. As inputs, we need the initial curve $f(0,T)$, which we obtain using the techniques described earlier, and calibrated model parameters that we attained in Table 4 in the Appendix. We should explicitly add that all state variables: $X_1, X_2, V_{11}, V_{12}, V_{21}, V_{22}$ start at $t = 0$ with zero values. Since our simulated at time t instantaneous forward rate $f(t,T)$ is given by:

$$f(t,T) = f(0,T) + \sum_{i=1}^{2} e^{\kappa_i(T-t)} \left(X_i(t) + \sum_{j=1}^{2} \frac{1-e^{-\kappa_j(T-t)}}{\kappa_j} V_{ij}(t) \right) \tag{101}$$

we have to decide on how many pillars (T) in the curve we would like to have, which is a trade-off between the accuracy of subsequent integration (when retrieving discounting factors from instantaneous forward rates) and time spent on a single forward curve simulation.[36]

Yet another consideration is the time step of the whole MC scheme, which is naturally more dependent on the desired granularity of expected exposure calculations.[37] For the X_i state variables we use the standard Euler scheme (where Δ-time step):

$$dX_i(t) = \left(-\kappa_i X_i(t-\Delta) + \sum_{j=1}^{2} V_{ij}(t-\Delta) \right) dt + \sigma_i \sqrt{\Delta} Z_i(t) \tag{102}$$

Since we consider deterministic volatility only, we have to modify the procedure for calculating θ_i^2 because the lower outside integration limit now has to move with horizon steps rather than being constant at zero:

$$\theta_i^2 = \int_{k\Delta}^{T_0} \left| \int_{T_0}^{T_i} \sigma(s,u)\, du \right|^2 ds$$

where k is the number of horizon steps from the simulation date.

---

[35] As defined in Section 2.

[36] We have taken 40 pillars spanning from 0.5 to 20 years.

[37] We have used monthly time steps in the MC schemes.

The second important technical adjustment is for IRS valuation at each horizon step in every path, which is a clear example of how XVA is mildly path-dependent.[38] Every path generated should be used to imply *current* floating coupons of the IRS that we are pricing, therefore certain valuation depends not only on the current state of the world, but a bit of history along the path as well.

Results for HW2

Figure 4 in the Appendix depicts the expected exposures and potential 95% quantile future exposures of different IRS contracts: starting deeply in-the-money, forward starting with fixed rate close to related ATMF and out-of-the-money generated using the methods described and developed in this dissertation. It will suffice to generate 2000 paths to produce results that are in line with those encountered in the literature: Green (2015), Gregory (2012), Lichters, Stamm, and Gallagher (2015) or Lu (2015). The time to generate these profiles is a function of the number of horizon time steps Δ, number of paths J, number of instruments in the portfolio, but also the complexity of instruments priced (i.e. schedule of cash flows) and binary flag if we need quantile exposures or not.[39] The reference times achieved in practical simulations are exhibited in Table 6. The exposure profiles may be generated on a counterparty portfolio level with different sets of netting rules as well, which is not adding much to the execution time.

Within the framework it is easily possible to generate exposures of such instruments as swaptions (both cash equivalent and physically settled), of which an example is depicted in Figure 3. We may use the generated profiles and some survival cumulative probability function (implied from market data or otherwise assumed) to calculate proper valuation adjustments (CVA). Discretisation of the *XVA integral* should rather be done with no finer time grid than the horizon's simulation steps.

7 Conclusions

It has been shown that the choice of a particular HJM's specification for XVA calculation in the multi--curve environment highly depends on the availability of reliable market data of non-linear derivatives on a specific market to which one can calibrate the model. In the availability of swaptions volatilities only (without active caps and floors market), we cannot calibrate a three-factor exponential model, for example. Another limiting situation is when there are no quotes for different market derivatives based on various xIBOR tenors (i.e. 3M, 6M).

Nevertheless, we performed calibrations of several one-factor models of the Cheyette class of HJM and found that even with relativity simple specification, i.e. Hull-White with two summands, we may achieve satisfactory results in terms of quality and calculation time of the calibration. This process may easily be scheduled before more complicated states of the world's generations and simulated valuation for XVA purposes.

We have shown that in the case of less liquid financial markets, as for example the Polish PLN interest rate derivatives market, there exists a relatively simple way of simultaneously extracting discounting and forwarding curves by means of a constant multiplicative spread calculated from

[38] Cf. Green (2015).

[39] Producing quantile exposures proves to take much longer than generic expected exposures.

the furthest common observed tenor of the two market curves. Such a construction is instrumental in the XVA engine calculations and may be used together with the Hull-White with two summands in the volatility functional form to cheaply (in CPU terms) evaluate a modest portfolio of interest rate derivatives. For sure, this is a practical take-away for risk managers, traders and middle-office specialists in banks and other financial institutions that are involved in interest rate derivatives trading and are obliged to model different XVA adjustments.

The efficacy of the whole system of XVA calculation depends, however, on many other degrees of freedom rather than the one resulting from model choice. The overall performance relies on dozens of decisions regarding, among others, a granularity of time-scale for XVA integral and of horizon steps, interpolation methods used at all stages, a method of minimization used in calibration, a degree of complexity of the instruments under simulation, complexity of the rules to calculate final exposure, and the possibility of analytical vs numerical integration of volatility surface function.

References

Ametrano F.M., Bianchetti M. (2009), *Bootstrapping the illiquidity: multiple yield curves construction for market coherent forward rates estimation*, chapter in: *Modelling Interest Rates: Latest Advances for Derivatives Pricing*, Risk Books.

Ametrano F.M., Bianchetti M. (2013), *Everything you always wanted to know about multiple interest rate curve bootstrapping but were afraid to ask*, SSRN Electronic Journal, DOI: 10.2139/ssrn.1334356.

Andersen L., Piterbarg V. (2010), *Interest Rate Modelling – Volume II: Term Structure Models,* 2, Atlantic Financial Press.

Beyna I. (2013), *Interest Rate Derivatives: Valuation, calibration and Sensitivity Analysis*, Springer Science & Business Media.

Beyna I., Wystup U. (2010), *On the calibration of the Cheyette interest rate model*, CPQF Working Paper Series, 25.

Beyna I., Chiarella C., Kang B. (2012), *Pricing Interest Rate Derivatives in a Multifactor HJM Model with Time*.

Bianchetti M. (2009), *Two curves, one price: pricing & hedging interest rate derivatives decoupling forwarding and discounting yield curves*, SSRN Electronic Journal, DOI: 10.2139/ssrn.1334356.

Bianchetti M. (2010), Two curves, one price, *Risk,* 23(8), 66.

Bianchetti M., Carlicchi M. (2011), *Interest rates after the credit crunch: multiple curve vanilla derivatives and SABR*, SSRN Electronic Journal, DOI: 10.2139/ssrn.1783070.

Bianchetti M., Carlicchi M. (2013), *Markets evolution after the credit crunch*, chapter in: *Interest Models After The Financial Crisis*, Risk Books.

Brace A., Musiela M. (1994), A multifactor Gauss Markov implementation of Heath, Jarrow, and Morton, *Mathematical Finance*, 4(3), 259–83.

Brace A., Musiela M. (1997), Swap derivatives in a Gaussian HJM framework, in: *Mathematics of Derivatives Securities*, Cambridge University Press.

Brigo D., Morini M., Pallavicini A. (2013), *Counterparty Credit Risk, Collateral and Funding: With Pricing Cases for All Asset Classes*, John Wiley & Sons.

Cheyette O. (1995), *Markov representation of the Heath-Jarrow-Morton model,* SSRN Electronic Journal, DOI: 10.2139/ssrn.6073.

Chibane M., Selvaraj J., Sheldon G. (2009), *Building curves on a good basis*, SSRN Electronic Journal, DOI: 10.2139/ssrn.1394267.

Filipovic D. (2009), *Term-Structure Models. A Graduate Course*, Springer.

Green A. (2015), *XVA: Credit, Funding and Capital Valuation Adjustments*, John Wiley & Sons.

Gregory J. (2012), *Counterparty Credit Risk and Credit Value Adjustment: A Continuing Challenge for Global Financial Markets*, John Wiley & Sons.

Hagan P.S., West G. (2006), Interpolation methods for curve construction, *Applied Mathematical Finance,* 13(2), 89–129.

Heath D., Jarrow R., Morton A. (1992), Bond pricing and the term structure of interest rates: a new methodology for claims valuation, *Econometrica,* 60(1), 77–105.

Henrard M. (2003), Explicit bond option formula in Heath-Jarrow-Morton one factor model, *International Journal of Theoretical and Applied Finance*, 6(01), 57–72.

Henrard M. (2014), *Interest Rate Modelling in the Multi-Curve Framework: Foundations, Evolution and Implementation*, Springer.

Hirsa A. (2012), *Computational Methods in Finance*, CRC Press.

Jamshidian F. (1991), Bond and option evaluation in the Gaussian interest rate model, *Research in Finance*, 9, 131–170.

Kienitz J. (2013), *Transforming volatility-multi curve cap and swaption volatilities*, SSRN Electronic Journal, DOI: 10.2139/ssrn.2204702.

Kienitz J., Caspers P. (2017), *Interest Rate Derivatives Explained*, Volume 2: *Term Structure and Volatility Modelling*, Springer.

Lichters R., Stamm R., Gallagher D. (2015), *Modern Derivatives Pricing and Credit Exposure Analysis: Theory and Practice of CSA and CVA Pricing, Exposure Simulation and Backtesting*, Springer.

Lu D. (2015), *The Xva of Financial Derivatives: CVA, DVA and FVA Explained*, Springer.

Moreni N., Pallavicini A. (2014), Parsimonious HJM modelling for multiple yield curve dynamics, *Quantitative Finance*, 14(2), 199–210.

Musiela M., Rutkowski M. (2005), *Martingale methods in financial modelling*, Springer-Verlag.

Ritchken P. Sankarasubramanian L. (1995), Volatility structures of forward rates and the dynamics of the term structure, *Mathematical Finance*, 5(1), 55–72.

Appendix

Table 1

Objective function values for different models calibrated

Date	Mod1	Mod2	Mod3	Mod4	Mod5
2017-10-09	0.017	0.017	0.017	0.016	0.005
2017-07-21	0.087	0.087	0.087	0.086	0.041
2017-05-05	0.078	0.069	0.078	0.061	0.014
2017-02-21	0.051	0.042	0.026	0.037	0.033
2016-12-06	0.057	0.045	0.057	0.040	0.059
2016-09-30	0.140	0.140	0.140	0.140	0.025
2016-07-27	0.170	0.160	0.170	0.160	0.019
2016-05-20	0.130	0.130	0.130	0.130	0.022
2016-03-15	0.093	0.093	0.093	0.091	0.016
2015-12-24	0.035	0.028	0.034	0.028	0.026
2015-10-20	0.040	0.020	0.040	0.018	0.028
2015-08-11	0.064	0.022	0.050	0.014	0.045
2015-05-28	0.057	0.020	0.057	0.013	0.034
2015-03-12	0.044	0.027	0.039	0.025	0.035
2014-12-24	0.045	0.031	0.038	0.029	0.038
2014-10-13	0.063	0.056	0.063	0.055	0.065
2014-08-06	0.020	0.015	0.020	0.015	0.016
2014-06-03	0.019	0.014	0.019	0.014	0.015
2014-03-12	0.012	0.006	0.012	0.006	0.006
2014-01-06	0.023	0.011	0.023	0.009	0.014

Notes: own calculations using Thomson Reuters and Tullett Prebon through Thomson Reuters data with Python *scipy.optimize* package. Models used: Mod1: Simple Hull-White (2 parameters), Mod2: Hull-White with $\beta(t)$ of Poly(1) form, Mod3: 2 summands of a simple Hull-White (2x2 parameters), Mod4: Beyna-Wystup with $\beta(t)$ of Poly(1) form and piecewise constant $\kappa(\cdot)$ function with one internal node $n_1 = 4$ (5 parameters), Mod5: same as Mod4 but with another summand – constant d (6 parameters).

Table 2
Results of quality checks of Beyna Wystup (Mod5) volatility surface calibration

Date	Cost function value	M1	M2	M3	M4	M5	M6	M7	M8	Outcome
2017-10-09	0.0046	0	0	0	0	0	0	0.96	0.92	**good**
2017-07-21	0.0414	0	0	0	0	0	0	0.96	0.92	**good**
2017-05-05	0.0137	0	0	0	0	0	0	0.92	0.92	**good**
2017-02-21	0.0334	0.13	0.13	0	0	0	1	0.96	0.96	**passed**
2016-12-06	0.0588	0.27	0.27	0	0	0	0	0.92	0.92	**good**
2016-09-30	0.0254	0	0	0	0	1	0	0.92	0.92	**passed**
2016-07-27	0.0191	0	0	0	0	0	0	1	0.96	**good**
2016-05-20	0.0221	0	0	0	0	1	0	0.92	0.92	**passed**
2016-03-15	0.0164	0	0	0	0	0	1	0.96	0.92	**passed**
2015-12-24	0.0257	0	0	0	0	0	0	0.92	0.88	**passed**
2015-10-20	0.028	0	0	0	0	0	0	0.96	0.88	**passed**
2015-08-11	0.0447	0.13	0.13	0	0	0	0	0.96	0.88	**passed**
2015-05-28	0.0341	0.13	0.13	0	0	1	0	0.96	0.92	**passed**
2015-03-12	0.0346	0.13	0.13	0	0	1	0	0.96	0.92	**passed**
2014-12-24	0.0378	0.13	0.13	0	0	1	0	0.96	0.92	**passed**
2014-10-13	0.0654	0.13	0.13	0	0	1	0	0.96	0.92	**passed**
2014-08-06	0.0165	0	0	0	0	0	0	0.96	0.92	**good**
2014-06-03	0.0152	0.13	0.13	0	0	0	0	0.96	0.92	**good**
2014-03-12	0.0058	0	0	0	0	0	0	0.92	0.92	**good**
2014-01-06	0.0143	0	0	0	0	1	0	0.96	0.88	**failed**

Notes: market data thanks to Reuters Thomson and Tullett Prebon through Thomson Reuters. Optimization method: Brute force with Nelder and Mead finish implemented in *scipy.optimize*, M1–M4: core conditions, M5-M8 secondary conditions. Model specification: $\sigma(t,T)=c+(at+b)\exp\left(-\int_t^T \underline{\kappa^{node}}(u)\,du\right)$ with one inner node at $n_1 = 4$ to Polish swaptions market on selected dates in the period 2014–2017.

Table 3
Results of quality checks of Hull-White model with two summands (Mod3) volatility surface calibration

Date	Cost function value	M1	M2	M3	M4	M5	M6	M7	M8	Outcome
2017-10-09	0.017389879	0.13	0.13	0	0	0	0	0.92	0.96	**good**
2017-07-21	0.087350276	0.13	0.13	0	0	0	0	0.96	0.96	**good**
2017-05-05	0.07836386	0.13	0.27	0	0	0	0	0.96	0.92	**good**
2017-02-21	0.026364114	0	0	0	0	0	0	0.96	0.92	**good**
2016-12-06	0.056664465	0.13	0.13	0	0	0	0	0.96	0.96	**good**
2016-09-30	0.137922479	0.27	0.27	0	0	0	0	0.96	0.92	**good**
2016-07-27	0.165262659	0.27	0.27	0	0	0	0	0.96	0.92	**good**
2016-05-20	0.130544034	0.13	0.13	0	0	0	0	0.96	0.96	**good**
2016-03-15	0.092947296	0.13	0.13	0	0	0	0	0.96	0.88	**passed**
2015-12-24	0.033947198	0	0	0	0	0	0	0.96	0.96	**good**
2015-10-20	0.040125248	0	0	0	0	0	0	0.92	0.96	**good**
2015-08-11	0.049816333	0	0	0	0	0	0	0.96	0.92	**good**
2015-05-28	0.057149841	0	0	0	0	0	0	0.92	0.92	**good**
2015-03-12	0.03858575	0	0	0	0	0	0	0.96	0.96	**good**
2014-12-24	0.037683165	0	0	0	0	0	0	0.96	0.92	**good**
2014-10-13	0.06301622	0.13	0.13	0	0	0	0	0.96	0.92	**good**
2014-08-06	0.020307428	0	0	0	0	0	0	0.96	0.96	**good**
2014-06-03	0.019179645	0	0	0	0	0	0	0.96	0.96	**good**
2014-03-12	0.011737015	0	0	0	0	0	0	0.96	0.96	**good**
2014-01-06	0.02282191	0	0	0	0	0	0	0.96	0.92	**good**

Notes: market data thanks to Reuters Thomson and Tullett Prebon through Thomson Reuters. Optimization method: Brute force with Nelder and Mead finish implemented in *scipy.optimize*, M1–M4: core conditions, M5–M8 secondary conditions, Model specification: $\sigma(t,T)=c_1\exp\left(-a_1\left(T-t\right)\right)+c_2\exp\left(-a_2\left(T-t\right)\right)$ to Polish swaptions market on selected dates in the period 2014–2017.

Table 4

Some optimisation results for three arbitrary dates for selected models

Mod1	**c**	**b**	**a**
2017-10-09	0.006704	(0.000114)	(0.038693)
2015-12-24	0.009073	(0.000336)	(0.013276)
2014-03-12	0.010737	(0.000570)	(0.004178)

Mod2	**c**	**a**
2017-10-09	0.006699	(0.031548)
2015-12-24	0.009323	0.010996
2014-03-12	0.010822	0.025863

Mod3	c_1	a_2	c_2	a_2
2017-10-09	0.000096	(0.030811)	0.006606	(0.031492)
2015-12-24	0.000185	0.005430	0.008914	0.005438
2014-03-12	0.002216	0.017683	0.008686	0.041846

Mod4	**a**	**c**	κ_1	κ_1
2017-10-09	(0.000220)	0.006789	(0.064704)	(0.026300)
2015-12-24	(0.000298)	0.009018	(0.001196)	(0.020264)
2014-03-12	(0.000547)	0.010698	0.005905	(0.010461)

Mod5	**d**	**a**	**c**	κ_1	κ_1
2017-10-09	(0.008419)	(0.004761)	0.009447	11.528936	0.843829
2015-12-24	0.008341	(0.000095)	0.000481	(0.418278)	0.002055
2014-03-12	0.004894	(0.000535)	0.005970	0.000071	0.000462

Table 5
Mean and standard deviation of objective function values and quality check final results for 20 selected dates in 2014–2017 for Polish swaptions, using different models

	Mod1	**Mod2**	**Mod3**	**Mod4**	**Mod5**
Avg OF	0.0517	0.0620	0.0594	0.0493	0.0279
SD OF	0.0458	0.0416	0.0426	0.0463	0.0158
Good	70%	90%	95%	50%	40%
Passed	25%	10%	5%	45%	55%
Failed	5%	0%	0%	5%	5%

Table 6
Simulation times for expected exposures of different portfolios (in sec)

Paths	**Number of contracts**		
	1	**10**	**50**
10	2.07	2.89	9.24
100	19.03	31.86	89.67
1000	182.95	310.72	913.47
10000	1926.66	3232.00	9327.12

Figure 1

Market Black-Scholes and calibrated implied volatility in Hull-White with two summands as of 20 November 2017

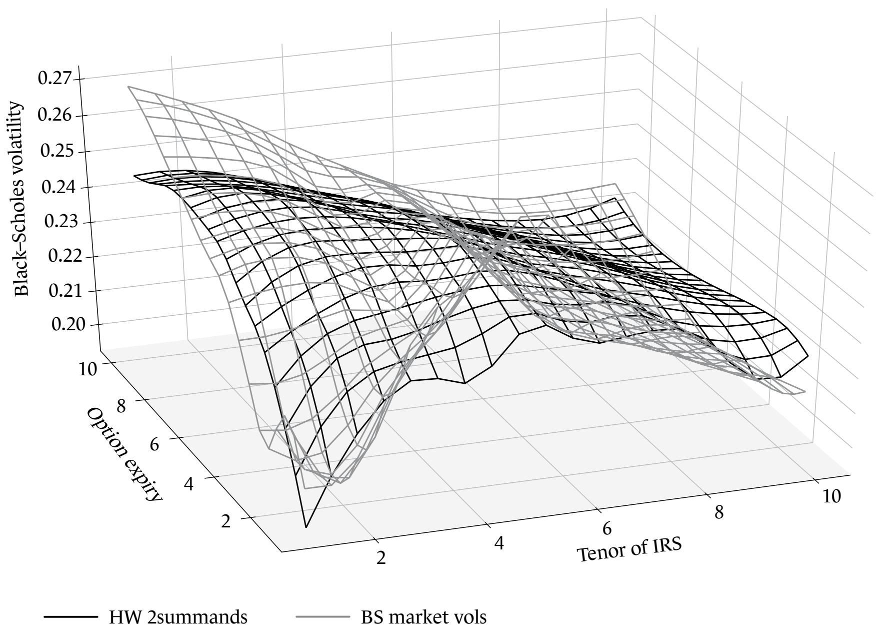
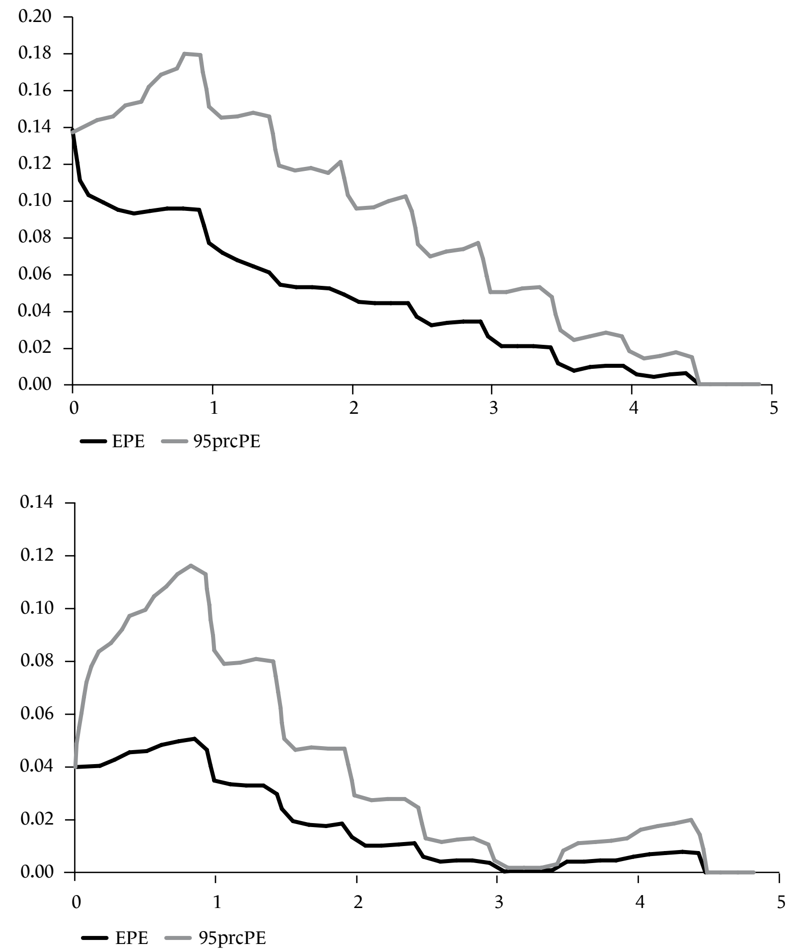
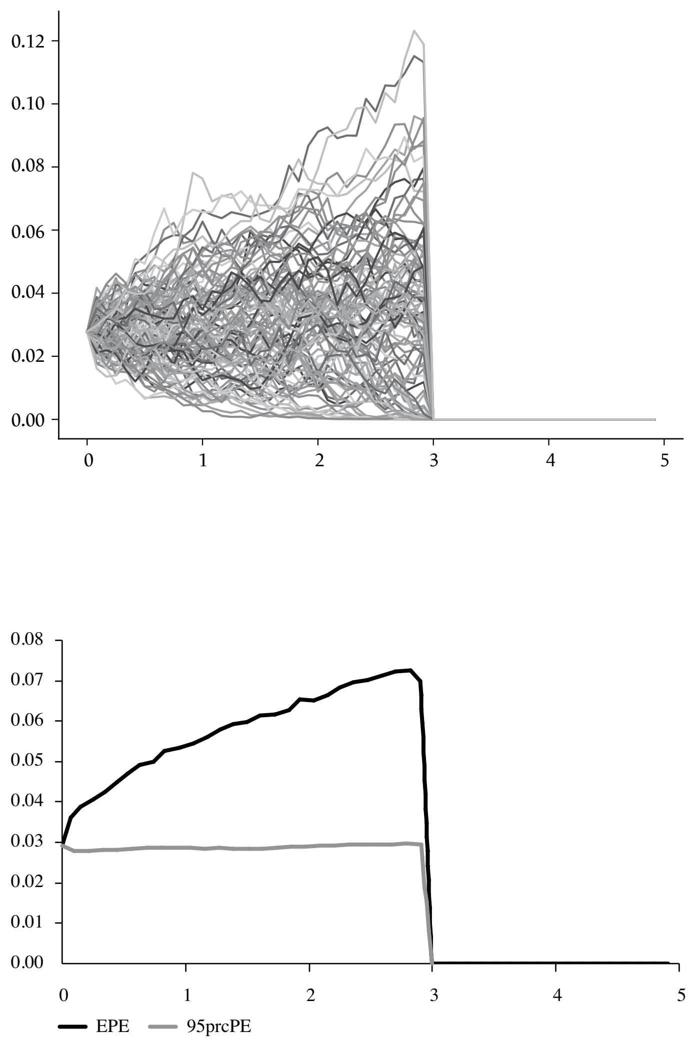


Figure 2
Expected exposure and quantile potential future exposure profiles with and without netting rules applied for a portfolio of 10 instruments



Notes: calculations for market data in PLN as of 2017-10-09 provided by Thomson Reuters and Tullett Prebon through Thomson Reuters, time step $\Delta = 1/12$, number of simulations $J = 2000$. Semi-annual coupons on both legs, notionals = 1. Netting rules applied in the lower pane.

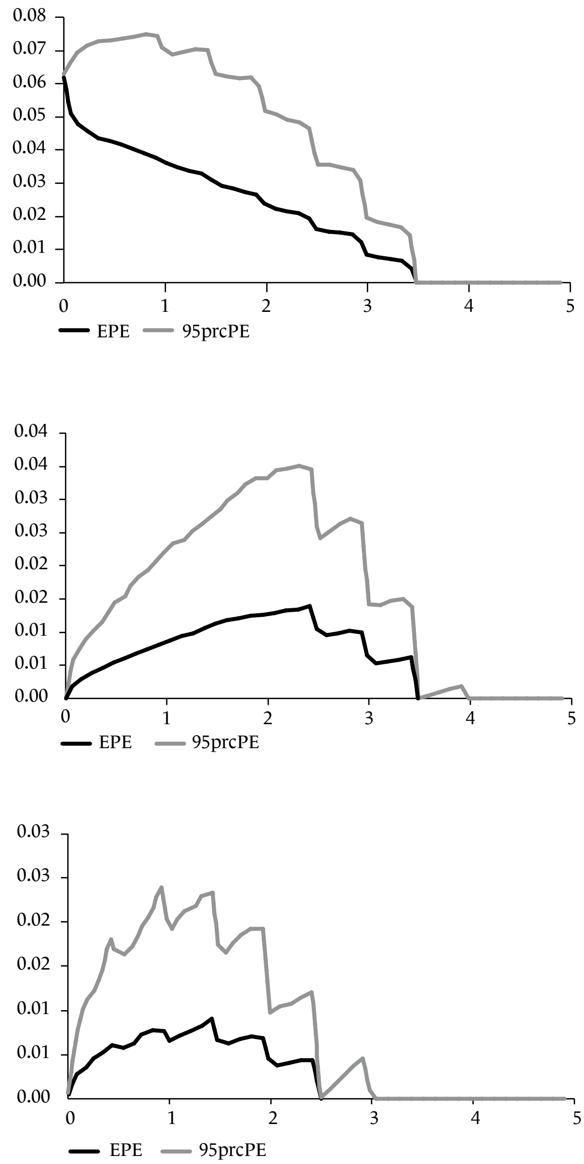
Figure 3
Expected exposure and quantile potential future exposure profiles and some sample paths for receiver swaption contract



Notes: receiver swaption fixed 0.044 2Y IRS in 3Y expiry time, cash equivalent settlement. All notionals are equal to 1. Sample paths (upper pane) and EE profiles (lower pane).

Figure 4

Expected exposure and quantile potential future exposure profiles for three different IRS contracts



Notes:

Calculations for market data in PLN as of 2017-10-09 provided by Thomson Reuters and Tullett Prebon through Thomson Reuters, calibrated parameters of HW2s: $c_1 = 0.00009629, a_1 = -0.0308108, c_2 = 0.00660564, a_2 = -0.0314923889$, time step $\Delta = 1/12$, number of simulations $J = 2000$. Instruments: 4Y IRS receiving fixed 0.03 starting at 0 (upper pane), 2Y IRS receiving fixed 0.024 staring in 2Y time (central pane), 3Y IRS receiving fixed 0.015 starting at 0 (lower pane). All IRSes with semi-annual coupons on both legs. All notionals are equal to 1.