Bermudan Swaption Valuation Under The Hull White Model

Audric Dongfack Kengfack

February 2020

1 Introduction

Bermudan swaptions are compounded options. At each exercise date you can or enter into a swap or keep your right up to the next exercise date. A popular way to value (Bermudan) swaption in a Hull-White (see equation (1)) or extended Vasicek model is to use a tree approach. That is what we intend to do in this report.

$$dr_t = (\theta(t) - ar_t)dt + \sigma(t)dW_t \tag{1}$$

The report follows specific order of questions from the interest model project given at Centrale Marseille by Mr Abderrahim Ben Jazia. This report is also follows by a code written in python (3.7).

2 The Hull White model

1. Explicit form of the process r_t :

We want to show that $r_t = r_s e^{-a(t-s)} + \alpha(t) - \alpha(s)e^{-a(t-s)} + \sigma \int_0^1 e^{-a(t-u)} dW_u$ where α is define as $\alpha(t) = f^M(0,t) + \frac{\sigma^2}{2a^2}(1-e^{-at})^2$ and f^M the instantaneous forward rate.

Let us consider the process H_t as we can write r_t on this form.

$$r_t = H_t e^{-at}$$

The Ito Lemma[1] give us :

$$dr_t = -ae^{-at}H_t dt + e^{-at}dH_t = -ar_t dt + e^{-at}dH_t$$

so, by replacing dr_t by its form give by equation (1) (see introduction) we have $dH_t = e^{-at}\theta(t)dt + \sigma e^{-at}dW_t$. By integrating dH_t between s and t , we have :

$$H_t = H_s + \int_s^t e^{au}\theta(u)du + \sigma \int_s^t e^{au}dW_u = r_s e^{as} + \int_s^t e^{au}\theta(u)du + \sigma \int_s^t e^{au}dW_u$$

[1] One of the main result of stochastic calculus theory developped by Kiyoshi Ito in 1940. Look Wikipedia for more details

So, we get this first form for r_t.

$$r_t = r_s e^{-a(t-s)} + \int_s^t e^{-a(t-u)}\theta(u)du + \sigma \int_s^t e^{-a(t-u)} dW_u$$

Since $\theta(u) = \frac{\partial f^M(0,u)}{\partial u} + af^M(0,u) + \frac{\sigma^2}{2a}(1-e^{-2au})$. An integration by part of the second term of the second member of the equality give us:

$$\int_s^t e^{-a(t-u)}\theta(u)du = e^{-at}(\int_s^t e^{au}\frac{\partial f^M(0,u)}{\partial u}du + \int_s^t ae^{au}f^M(0,u)du + \int_s^t \frac{\sigma^2}{2a}(e^{au}-e^{-au})du)$$

$$\int_s^t e^{-a(t-u)}\theta(u)du = f^M(0,t) - e^{-a(t-s)}f^M(0,s) + \frac{\sigma^2}{a^2}[Cosh(au)]_s^t e^{-at}$$

Or we have $[Cosh(au)]_s^t = \frac{1}{2}(1-e^{-at})^2e^{at} - \frac{1}{2}(1-e^{-as})^2e^{as}$, that is because $\frac{1}{2}(1-e^{-at})^2e^{at}$ is one primitive of the function $Cosh(au)$.
Using this and replacing the final result obtained in the first form of the process r_t, we get:

$$r_t = r_s e^{-a(t-s)} + \alpha(t) - \alpha(s)e^{-a(t-s)} + \sigma \int_0^1 e^{-a(t-u)} dW_u \quad (2)$$

$$\alpha(t) = f^M(0,t) + \frac{\sigma^2}{2a^2}(1-e^{-at})^2 \quad (3)$$

2. Calibration of the Nelson Siegel Model :
The model of Nelson and Siegel (1987) and its extension by Svensson (1994) are widely used by central banks and other market participants as a model for the term structure of interest rates . Nelson Siegel states that yield for maturity τ is given by :

$$y(\tau) = \beta_0 + \beta_1(\frac{1-e^{\frac{-\tau}{\lambda}}}{\frac{\tau}{\lambda}}) + \beta_2(\frac{1-e^{\frac{-\tau}{\lambda}}}{\frac{\tau}{\lambda}} - e^{\frac{-\tau}{\lambda}}) \quad (4)$$

The traditional procedures used to estimate the NS model parameters can be summarized as follows: i) minimizing the sum of squared errors (SSE) using OLS over a grid of pre-specified values of ls (Nelson and Siegel, 1987); ii) minimizing SSE using linear regression conditional on a chosen fixed shape parameter l (Diebold and Li, 2006; de Pooter, 2007; and Fabozzi et al., 2005) and iii) using nonlinear optimization techniques (Cairns and Pritchard, 2001). Annaert et al. (2013) propose combining a grid search to determine the value of the optimal shape parameter with a ridge regression in order to solve some of the estimation problems resulting from the traditional linear and nonlinear estimation of the model. We will use in this report the first calibration method.

In this approach, we fixed the value λ and we estimate the remaining parameters by OLS. In this method, the objective function is :

$$min_{\theta} \sum_{i=1}^{N} (f(\theta_i, \tau_i) - y_i)^2$$

Where $f(\theta_i, \tau_i)$ is the estimated yield in the model and y_i is the observed yield (take in data).N is the length of maturities vector . The parameter set is $\theta = \{\beta_0, \beta_1, \beta_2\}$.
In python, there exist a *nelson_siegel_svenson* module. This module contained the function *calibrate_ns_ols* which calibrated the curve using OLS approach discuss previously.In addition, this optimizer determines the optimal λ using grid search (value between $[0 - 100]$ according to the developer of the module, which is convenient with literature, see Diebold and Li (2006)) by minimizing the error at fixed λ.The objective function seem like :

$$min_{\lambda} min_{\theta} \sum_{i=1}^{N} (f(\theta_i, \tau_i) - y_i)^2$$

By applying the *calibrate_ns_ols* to our dataset (see the python code), we obtain the follow parameters:

Parameters values			
β_0	β_1	β_2	λ
0.047	-0.046	-0.039	1.239

Now let's plot our estimated Zero coupond yield curve and see how his fit well the yield curve.

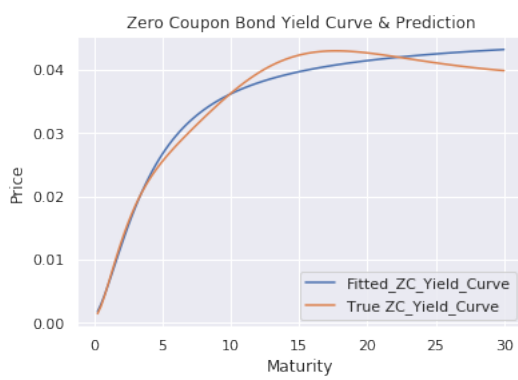

Figure 1: Comparaison graph between the ZC_Yield_Curve and the $Fitted\_ZC\_Yield\_Curve$

3. Expression of the forward rate under the Nelson Siegel Characterization :

The definition of the instantaneous forward rate in term of the yield curve is giving by :

$$y(\tau) = \int_0^\tau \frac{f(s)}{\tau} ds \tag{5}$$

so the instantaneous forward rate is giving by :

$$f(\tau) = ((.)y(.))'(\tau)$$

By replacing y by its Nelson Sielgel definition, we get:

$$f(\tau) = ((.)y(.))'(\tau) = (\tau\beta_0 - \lambda\beta_1(1 - e^{\frac{-\tau}{\lambda}}) + \lambda\beta_2(1 - e^{\frac{-\tau}{\lambda}}) + \tau\beta_2 e^{\frac{-\tau}{\lambda}})'$$

$$f(\tau) = \beta_0 + \beta_1 e^{\frac{-\tau}{\lambda}} + \beta_2 e^{\frac{-\tau}{\lambda}} - \beta_2 e^{\frac{-\tau}{\lambda}} + \frac{\tau}{\lambda}\beta_2 e^{\frac{-\tau}{\lambda}} = \beta_0 + \beta_1 e^{\frac{-\tau}{\lambda}} + \beta_2(\frac{\tau}{\lambda})e^{\frac{-\tau}{\lambda}}$$

Let's try to draw the instantaneous forward rate curve.

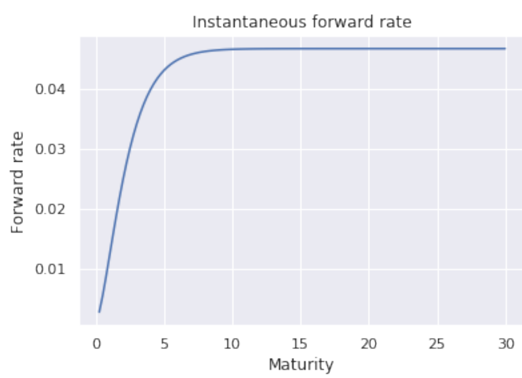

Figure 2: Instantaneous Forward Rate

We can notice that after 10 years of maturity, the instantaneous forward rate becomes constant. This is clearly explained by the fact that the term $\beta_1 e^{\frac{-\tau}{\lambda}} + \beta_2(\frac{\tau}{\lambda})e^{\frac{-\tau}{\lambda}}$ becomes zero (decreasing exponential).

4. Determination of $\theta(t)$:

The parameters θ in the Hull and White model is give by:

$$\theta(\tau) = \frac{\partial f(\tau)}{\partial \tau} + af(\tau) + \frac{\sigma^2}{2a}(1 - e^{-2a\tau})$$

Using the previous expression of the instantaneous forward rate, we have :

$$\frac{\partial f(\tau)}{\partial \tau} = \frac{(\beta_2 - \beta_1)}{\lambda} e^{\frac{-\tau}{\lambda}} - \frac{\tau}{\lambda^2}\beta_2 e^{\frac{-\tau}{\lambda}}$$

So, by replacing f and its first derivative by their expression we have :

$$\theta(\tau) = a\beta_0 + (a - \frac{1}{\lambda})\beta_1 e^{\frac{-\tau}{\lambda}} + (\frac{1}{\lambda} + \frac{\tau}{\lambda}(a - \frac{1}{\lambda}))\beta_2 e^{\frac{-\tau}{\lambda}} + \frac{\sigma^2}{2a}(1 - e^{-2a\tau})$$

Let's try to draw the parameter θ curve.

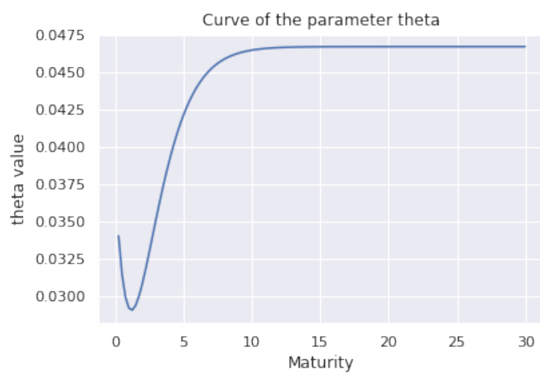

Figure 3: Graph of the parameter θ with $a = 1$ and $\sigma = 0.01$

3 Trinomial tree

Hull and White (1994) proposed a robust two-stage procedure for constructing trinomial trees to represent a wide range of one-factor models. We suppose that the time step in the tree Δt remains constant over time and we assume that the Δt rate, R, follows the same process as r:

$$dR = (\theta - aR)dt + \sigma dW \tag{6}$$

In this part of the repport, we will progressively simulated the two-stage.

3.1 The first stage

5. Construction of the auxiliary tree for a period of 10 years:
Simulated the auxilliary tree $R^*_{i,j}$ consists to constructing the parsimonious matrix which column represent the time spanning (represent here by i) and row represent the spatial spanning (represent here by $j = j_{min,i}.......j_{max,i}$). A simple way to do that is to determine at each time i the corresponding nodes j. Here is a preview of our auxilliary tree.

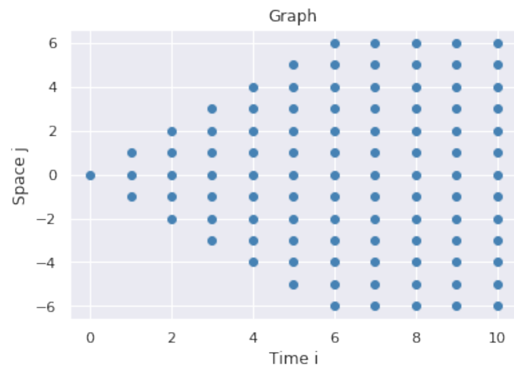

Figure 4: Representation of the auxilliary tree for a period of 10 years with $a = 0.1$ and $\sigma = 0.01$

The branching is symmetrical with respect to the time-axis ($j_{min,i} = j_{max,i}$). And here we have :

$$R^*_{i,j} = j\Delta R^*$$

From date 6 years , the nodes are identical. This reflect the mean reversion property of the process. That means, when the short rate is heigh , means reversion tends to cause the interest rates to decrease and vice verse (When is low, mean reversion tends to cause then to increase).

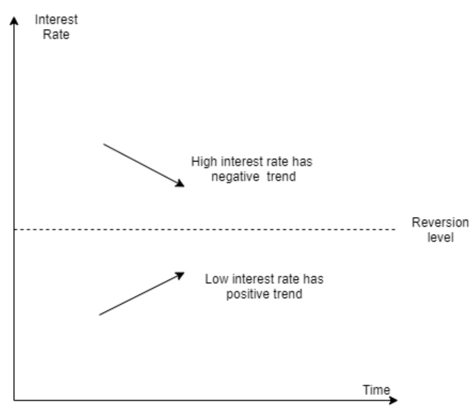

Figure 5: Visual explanation of the mean reversion phenomenon

3.2 The second stage

6. Construction of the tree for the process $Q_{i,j}$:
$Q_{i,j}$ is the value at the initial time of an asset pays 1 if the node (i,j) is reached and zero otherwise. The recursive definition of $Q_{i,j}$ is follow :

$$Q_{0,0} = 1$$

$$Q_{i+1,j} = \sum_{h} Q_{i,h} q(h,j) e^{-(\theta(i)+h\Delta R)\Delta T}$$

Here, $q(h,j)$ represent the transition probability from the node (i,h) to the node $(i+1,j)$. So, constructing the tree for $Q_{i,j}$ process consists first of all to define the transition probability $q(h,j)$ (see python code). And after them, we can easily define recursively the value $Q_{i,j}$,

Note that according to the branching tree above $(Fig-4)$, the node (i,j) is an attainable if and only if this two conditions are verified:

$$j \in [-i, i]$$

$$j \in [-j_{max}, j_{max}]$$

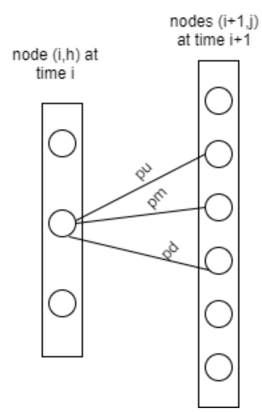

Figure 6: Illustration of the logic behind the algorithm defining the transition probability. We reason in the time intervall i and $i+1$, we also add the constraint define above on j.

7. Value of $\theta(i)$:
We have :

$$\theta(i) = -\frac{1}{\Delta T} ln\left(\frac{\sum_{j=j_{min}}^{j_{max}} Q_{i+1,j}}{\sum_{j=j_{min}}^{j_{max}} \sum_{h=j_{min}}^{j_{max}} Q_{i,h} q(h,j) e^{-h\Delta R\Delta T}}\right) \tag{7}$$

Proof : In fact, since we have:

$$Q_{i+1,j} = \sum_h Q_{i,h} q(h,j) e^{-(\theta(i)+h\Delta R)\Delta T}$$

By summing each member of the equality between $j = j_{min}$ and $j = j_{max}$. And also remmenber that $j_{min} = -j_{max}$, we can write:

$$\sum_{j=j_{min}}^{j_{max}} Q_{i+1,j} = \sum_{j=j_{min}}^{j_{max}} \sum_h Q_{i,h} q(h,j) e^{-(\theta(i)+h\Delta R)\Delta T}$$

$$\sum_{j=j_{min}}^{j_{max}} Q_{i+1,j} = e^{-(\theta(i)\Delta T)} \sum_{j=j_{min}}^{j_{max}} \sum_h Q_{i,h} q(h,j) e^{-(h\Delta R)\Delta T}$$

$$\theta(i) = -\frac{1}{\Delta T} ln(\frac{\sum_{j=j_{min}}^{j_{max}} Q_{i+1,j}}{\sum_{j=j_{min}}^{j_{max}} \sum_h Q_{i,h} q(h,j) e^{-h\Delta R\Delta T}})$$

let's draw an overview of the $\theta(i)$ function.

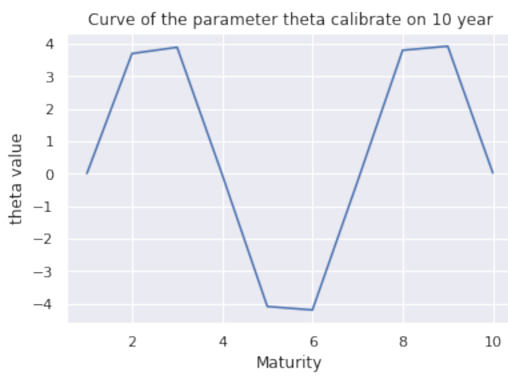

Figure 7: Curve of the parameter θ calibrate on 10 year

8. Tree for the Libor rate $L(T_i, T_{i+1})$ **:**
Knowing that the libor rate respects this relationship (equation 8) with the tree of $R_{i,j}$.

$$P(T_i, T_{i+1}) = e^{-R_i \Delta T} = \frac{1}{1 + L(T_i, T_{i+1})\Delta T} \tag{8}$$

Thus,

$$L(T_i, T_{i+1}) = \frac{1}{\Delta T}(e^{R_i \Delta T} - 1)$$

Constructing the libor tree is therefore equivalent to constructing at each node the value of the corresponding libor (see notebook).

9. Valuation procedure of the Swap between $(T_0, T_{10}) = (0, 10)$ with fixed leg 2%:

If we consider one leg of an interest rate swap in which a floating rate of interest is exchanged (LIBOR) for a specified fixed rate of interest, K. The start date is T_i and the end date is T_{i+1}. On the start date we observe the rate that applies between T_i and T_{i+1}. There is a payement of the end date equal to $N(L(T_i, T_{i+1}) - K)(T_{i+1} - T_i)$ where $L(T_i, T_{i+1})$ is the Libor for the periode between T_i and T_{i+1}. N is the notional.
It is show that in the world where interest rates are stochastic, we can use $P(0, T_{i+1})$ as the discount factor providing. So the value of the leg is :

$$S_i(K) = N(L(T_i, T_{i+1}) - K)(T_{i+1} - T_i)P(0, T_{i+1})$$

Thus, for the interest Swap between $(T_0, T_{10}) = (0, 10)$ with fixed leg $K = 2\%$, the value is giving by:

$$S(T_0, T_{10}, K) = N\sum_{i=0}^{9}(L(T_i, T_{i+1}) - K)(T_{i+1} - T_i)P(0, T_{i+1})$$

Since we have the libor tree, we can simply update the sums up to the initial node T by back-propagation (see notebook).

10. Verification of the Swap Price: The closed formula for the Swap price between $(T_0, T_{10}) = (0, 10)$ with fixed leg $K = 2\%$, is giving by :

$$S(T_0, T_{10}, K) = NP(T_0, T_{10})$$

If we compute this, we find the almost the same price as a technique describe above.
11. Valuation of the caplet

We recall that a caplet is a financial product that delivers a rate equal to the difference between an interest rate and an exercise rate, only if this difference is positive. Our caplet has a T_9-setting date and a T_{10}-exercise date.

In general, the rate given by the caplet depends on the fraction of time between the setting and exercise date, and also a nominal value, but often these values can be omitted. This means that in general we will generally consider the value of a caplet with an exercise rate K, setting date T_9 and maturity T_{10} given by the formula,

$$V^{Caplet}(T_9, T_{10}) = NP(T_9, T_{10})(L(T_9, T_{10}) - K)^+$$

For the code, we compute at each j-node the Payoff by using the formula above. So, each node will give the possible price at date T_9
12. Valuation of the European Swaption

We recall that the European Swaption can be viewed as an option on a coupon-bearing bond ([4]). Indeed, for a receiver swaption with strike K, maturity T_0 and nominal N, which gives the holder the right to enter at time T_0 an interest rate swap with payment times $T = (T_1, ..., T_{10})$, where he receives at the fixed rate K and pays LIBOR. The swaption price at time T_0 is giving by :

$$V^{EuropeanSwaption} = N \sum_{i=1}^{10} \Delta T K ZBC(T_0, T_i, K)$$

Where ZBC is giving by the following formula :

$$ZBC(T_0, T_i, K) = P(T_0, T_i)\Phi(h) - KP(T_0, T_i)\Phi(h - \sigma_p)$$

$$\sigma_p = \frac{1}{a}\sqrt{\frac{1}{2a}}$$

$$h = \frac{1}{\sigma_p} ln \frac{P(T_0, T_i)}{K} + \frac{\sigma_p}{2}$$

Φ denoting the standard normal cumulative distribution function.

13. Valuation of the Bermudan Swaption between T_0 and T_{10}

Idea behind the valuation : the algorithm due to lack of time we didn't implement this part. However we unroll the algorithm.

Hypothesis

We suppose we hold the contract.

We can either wait the final maturity T_9 and last payment T_{10}. Or exercice at any earlier time T_l with $T_0 <= T_l < T_9$ and then enter the Interest Rate Swap with first reset date T_l and last payment T_{10}.

For each l, we have predefined partition define as follows: $[T_l, T_9]$. $P_{i,j}(T_s)$ is the bond price $P(T_i, T_s)$ for the maturity T_s in the j-node of the tree at time i.

1. $P_{i,j}(T_{10}) = 1$ for all j
2. (Backward propagation inside T_{10} and T_0). While going backward from time T_{i+1} to T_i in the tree. The vector of the bond price is compute as : $P_{i,j}(T_s) = e^{R_{i,j}}[p_u P_{i+1,k+1}(T_s) + p_m P_{i+1,k}(T_s) + p_d P_{i+1,k-1}(T_s)]$
3. if $i > l$ - decrease i by one and go back to the preceding point
4. if i = l , we reached T_l. If $l > 0$ decrease l by one and go back to the point 1
5. since l = 0, by going backwards we have reached the last point in the time where the swaption can be exerciced
6. (Checking the exercise opportunity in each node of the current time). Compute for each level j in the current column of the three, the time-T_l

value of the underlying IRS (Interest Rates Swap) with first reset date T_l and last payment in T_{10}, based on the backward-propaged bond prices.

$$P_{1,j}(T_{l+1}), P_{1,j}(T_{l+2}), ..., P_{1,j}(T_{10})$$

i.e

$$IRS_j^l = 1 - P_{1,j}(T_{10}) - \sum_{s=l+1}^{10} KP_{1,j}(T_s)$$

7. If l = 0 then define the backwardly-Cumulated value from Continuation (CC) of the Bermudan swaption as this IRS value in each node j of the current time level in the tree,

$$CC_{0,j} = IRS_j^l$$

8. Else if $l > 0$, check the exercice opportunity as follows: if the underlying IRS is larger than the backwardly-cumulated value from continuation, then set the CC value equal to the IRS value. In our notation, for each node j in the current column of the tree If $IRS_j^l > CC_{0,j}$, then $CC_{0,j} = IRS_j^l$ Store such CC values in the corresponding nodes of the tree. Decrease by one and move to the next step.

So we can easily use the valuation of the financial instruments in questions 10 , 11 and 12.

14. Hull and white model Calibration (a and σ) :

Volatility parameters are estimated from market data on liquid options (swaption etc.). The first step is to look for parameters that minimize this objective function:

$$\sum_{i=1}^{n}(U_i - V_i)^2$$

where U_i is the market price of the financial instrument i and V_i is the compute price give by the model. n represent the number of the liquid asset using for the calibration.
The levenberg Marquardt[2] procedure is used to minimize the fitting error. When both parameters are functions of time, a penalty function is added to the fitting error to ensure that the functions obtained are regular.

The assets used for calibration should be chosen as close as possible to the assets that will be valued by the model. For our 10-year Bermudan swaption option that can be exercised at any payment date. The most appropriate calibration instruments are the following European swaptions : (1 * 9) , (2 * 8) , (3 * 7),.... (9 * 1). k and l in $(k * l)$ is respectively the start and the lenght of

[2]For more details, see W.H Press, B.P. Flannery,S.A. Teukolsky and al, Numerical Recipes in C: The Art of scientific computing, Cambrige University Press, Cambridge, 2007

the european swaption.

For the coding of this part, we will use a famous python library called **Quantlib** and the JamshidianSwaptionEngine pricer contained in it (to go quickly and also since our model above is not very accurate - but is pretty the same pricing procedure)

Parameters a and σ after Calibration	
a	σ
0.000011	0.004233

References

[1] K. Chourdakis. Financial engineering: A brief introduction using the matlab system. Working Book, 8:1–18, 2008

[2] John Hull, Alan W.(1994). Numerical procedures for implementing term structure models I: Single-Factor Models

[3] John Hull, Alan W.(2013). OIS Discounting, interest rates derivatives, and the modeling of stochastic interest rate spread.

[4] D. Brigo and F. Mercurio. Interest rate models-theory and practice: with smile, inflation and credit. Springer Science & Business Media, 2007.

[5] J. C. Hull. Options futures and other derivatives. Pearson Education India , 2010.