ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
Corso di Laurea Magistrale in Matematica
- Curriculum Applicativo -

The Libor Market Model: from theory to calibration

Tesi di Laurea in Equazioni Differenziali Stocastiche in Finanza Matematica

Relatore:
Chiar.mo Prof.
Andrea Pascucci

Presentata da:
Candia Riga

I Sessione di Laurea: 24 Giugno 2011
Anno Accademico 2010/2011

Acknowledgements

I am heartily thankful to my supervisor, Andrea Pascucci, who gave me the right directions for studying the subject and, above all, guided me throughout my studies and my dissertation. Then, it is a pleasure to thank the others who have played an important role by accompanying me in the studies that are object of my thesis, in particular Paolo Foschi and Giacomo Ricco. Thanks also to Unipol UGF Banca e Assicurazioni.

Furthermore, I would like to thank my university colleagues and my friends, who shared with me these two years, and my partner, who has always supported me. Last but not least, I will be forever grateful to my family, that made all this possible.

Abstract

This thesis is focused on the financial model for interest rates called the LIBOR Market Model, which belongs to the family of market models and it has as main objects the forward LIBOR rates. We will see it from its theoretical approach to its calibration to data provided by the market. In the appendixes, we provide the theoretical tools needed to understand the mathematical manipulations of the model, largely deriving from the theory of stochastic differential equations. In the inner chapters, firstly, we define the main interest rates and financial instruments concerning with the interest rate models. Then, we set the LIBOR market model, demonstrate its existence, derive the dynamics of forward LIBOR rates and justify the pricing of caps according to the Black's formula. Then, we also present the model acting as a counterpart to the LIBOR market model, that is the Swap Market Model, which models the forward swap rates instead of the LIBOR ones. Again, even this model is justified by a theoretical demonstration and the resulting formula to price the swaptions coincides with the one used by traders in the market, i.e. the Black's formula for swaptions. However, the two models are not compatible from a theoretical point, because the dynamics that would be obtained for the swap rate, by starting from the dynamics of the LIBOR market model, is not log-normal as instead is in the swap market model. Took note of this inconsistency, we select the LIBOR market model and derive various analytical approximating formulae to price the swaptions. It will also be explained how to perform a Monte Carlo algorithm to calculate the expectation of any payoff involving such rates by a simulation. Finally, it will be presented the calibration of the LIBOR market model to the markets of both caps and swaptions, together with various examples of application to the historical correlation matrix and the cascade calibration of the forward volatilities to the matrix of implied swaption volatilities provided by the market.

Sommario

Questa tesi è incentrata su un modello di mercato per i tassi d'intersse detto LIBOR Market Model, il quale modellizza i tassi forward LIBOR, a partire dalla sua impostazione teorica fino alla sua calibrazione ai dati forniti dal mercato. Nelle appendici vengono forniti gli strumenti teorici necessari per la gestione matematica del modello, derivanti in gran parte dalla teoria delle equazioni differenziali stocastiche. Nei capitoli interni, innanzitutto vengono definiti i principali tassi di interesse e gli strumenti finanziari alla base dei modelli di mercato sui tassi d'interesse. Poi viene impostato il LIBOR market model, dimostrata la sua esistenza, ricavate le dinamiche dei tassi forward LIBOR e i prezzi dei cap, fedeli alla formula di Black. Viene poi presentato anche il modello che funge da controparte al LIBOR market model, ovvero lo Swap Market Model, che modellizza i tassi forward swap anziché i LIBOR. Anche in questo modello si giustifica, mediante una dimostrazione teorica, una formula usata dai traders sul mercato per prezzare le swaption, ovvero la formula di Black per le swaption. Tuttavia, i due modelli non sono compatibili dal punto di vista teorico, in quanto la dinamica che si otterrebbe per il tasso swap a partire dalle dinamiche del LIBOR market model non è log-normale come invece è nello swap market model. Preso atto di questa inconsistenza, viene scelto il LIBOR market model e vengono derivate diverse formule analitiche approssimate per prezzare le swaption. Inoltre è spiegato come realizzare l'algoritmo di Monte Carlo per calcolare tali prezzi mediante una simulazione. Infine viene presentata la calibrazione del suddetto modello al mercato dei cap e a quello delle swaption, con diversi esempi di applicazione alla calibrazione della matrice di correlazione storica e alla calibrazione a cascata delle volatilità forward alla matrice delle volatilità implicite di mercato delle swaption.

Abbreviations and Notations

w.r.t. = with respect to

s.p. = stochastic process

s.t. = such that

a.s. = almost surely

SDE = stochastic differential equation

B.m. = Brownian motion

EMM = equivalent martingale measure

e.g. = exempli gratia ≡ example given

i.e. = id est ≡ that is

IRS = Interest Rate Swap

PFS = Payer Interest Rate Swap

RFS = Receiver Interest Rate Swap

PFS = Payer Interest Rate Swap

LMM = LIBOR Market Model

SMM = Swap Market Model

c.d.f. = cumulative distribution function

r.v. = random variable(s)

E = expectation

Std = standard deviation

i.i.d. = independent identically distributed

Contents

List of Figures

List of Tables

Introduction

In this thesis we present the most promising family of interest rate models, that are the market models. The two main representatives of this family are the LIBOR Market Model (LMM), which models the forward LIBOR rates as the primary objects in an arbitrage free way instead of deriving it from the term structure of instantaneous rates, and the Swap Market Model (SMM), which models the dynamics of the forward swap rates. The advantages of them is that, by choosing a deterministic volatility structure for the dynamics of the rates modeled, the first one prices caps according to the Black's cap formula, whereas the second one prices swaptions according to the Black's swaption formula. Indeed, these Black's formulae are the standard ones used respectively in the cap and swaption markets, which are the two main markets in the interest-rate derivatives world. Despite the good premises, the desirable compatibility between the two market formulae is not theoretically confirmed.

The LIBOR Market Model came out at the end of the 90s, in particular it was rigorously introduced in 1997 by Brace, Gatarek and Musiela, "The Market Model of Interest Rate Dynamics", then other significant contributes came by Jamshidian, "LIBOR and swap market models and measures", and Miltersen, Sandmann and Sondermann, "Closed Form Solutions for Term Structure Derivates with Log-Normal Interest Rates", all in 1997. At the same time, Jamshidian introduced the Swap Market Model, in 1997.

The point of this work is to analyze in detail the LIBOR market model, from its theoretical setting and mathematical results to its financial and practical use, together with some practical applications referring to the current market data. This thesis is structured as follows.

Appendixes. We give a synthetic presentation of the mathematical definitions and the main theoretical results concerning the theory of the stochastic processes and the stochastic differential equations. Moreover, we show the details of one of the most important tool in the mathematical studying of financial markets, that is the change of numeraire.

Chapter 1. We define the different types of interest rates and introduce the basic financial instruments and the main derivatives we are dealing with, i.e. caps and swaptions, together with their practically used Black's formulae.

Chapter 2. We introduce the LMM, derive the dynamics of the forward LIBOR rate modeled and prove its existence. Then we introduce the Black volatilities implied by the cap market and show that the risk-neutral valuation formula of caps gives the same prices as the Black's cap formula.

Chapter 3. We introduce the SMM, prove that the pricing formula for swaptions coincides with the Black's swaption formula and show the inconsistency of the dynamics assumed by the SMM with the ones given by the LMM.

Chapter 4. We show in detail the different approaches to price swaptions under the framework of the LMM, from the Monte Carlo simulation to various analytical approximating formulae.

Chapter 5. We introduce the important and rich subject of the instantaneous correlation modeling, dealing with the modeling and parametrization of the correlations between the Brownian motions driving the dynamics of forward LIBOR rates.

Chapter 6. We introduce the calibration of the LMM to the cap market and we present various plausible parameterizations for the forward volatility structure. Then we consider, for each of them, the evolution in time of the term structure of volatility.

Chapter 7. We introduce the calibration of the LMM to the swaption market, in particular we show an exact cascade calibration. Moreover, we introduce the use of a historical correlation matrix, along with its computation and parametric calibration.

Chapter 1

Interest rates and basic instruments

Definition 1.1. A *zero-coupon bond* (also known as pure discount bond) with maturity date T, briefly called T-bond, is a contract which guarantees the holder to be paid 1 unit of currency at time T, with no intermediate payments. The contract value at time $t < T$ is denoted by $p(t, T)$.

We must make some assumptions:

- there exist a (frictionless) market for T-bond for every $T > 0$;
- $p(T, T) = 1$ holds for all $T > 0$ (it avoids arbitrage);
- for all fixed $t < T$ the application $T \mapsto p(t, T)$ is differentiable w.r.t. maturity time.

The graph of the function $T \mapsto p(t, T)\,, \quad T > t\,$, called *zero-bond curve*, is decreasing starting from $p(t, t) = 1$ and will be typically very smooth. Whereas, for each fixed maturity T, $p(t, T)$ is a scalar stochastic process whose trajectory will be typically very irregular (determined by a Brownian motion).

The amount of time from the present date t and the maturity date T, called the time to maturity, is calculated in different ways, according to the market convention (*day-count convention*). Once this last is made clear, the measure of time to maturity, denoted by $\tau(\mathbf{t}, \mathbf{T})$, is referred to as the *year-fraction* between the dates t and T and it's usually expressed in years. The most frequently used day-count convention are:

- Actual/365 $\longrightarrow$ a year is 365 days long and the year-fraction between two dates is the actual number of days between them divided by 365;

- Actual/360 $\longrightarrow$ a year is 360 days long and the year-fraction between two dates is the actual number of days between them divided by 360;

- 30/360 $\longrightarrow$ months are 30 days long, a year is 360 days long and the year-fraction between two dates (d_1, m_1, y_1) and (d_2, m_2, y_2) is given by the ratio

$$\frac{\max(30 - d_1, 0) + \min(d_2, 30) + 360 \cdot (y_2 - y_1) + 30 \cdot (m_2 - m_1 - 1)}{360}.$$

In all the conventions adjustments may be included to leave out holidays.

Zero-coupon bonds are fundamental objects in the interest rate theory, in fact all interest rates can be also defined in terms of zero-coupon bond prices.

Talking about interest rates we need to distinguish the two main categories:

- government rates, related to bonds issued by governments;

- interbank rates, at which deposits are exchanged between banks and swap transactions between them are.

We are considering the interbank sector of the market, however the mathematical modeling of the resulting rates would be analogous in the two sectors.

Actually, interest rates are what is usually quoted in the (interbank) financial markets, whereas zero-coupon bonds are theoretically instruments not directly observable.

Definition 1.2. The *continuously-compounded spot interest rate* at time t for the maturity T is the constant rate at which an investment of $p(t,T)$ unit of currency at time t accrues continuously to yield 1 unit of currency at time T.
It is denoted by $R(t,T)$ and is defined by:

$$R(t,T) := -\frac{\ln p(t,T)}{\tau(t,T)} .$$

Equivalently:

$$e^{R(t,T)\tau(t,T)} p(t,T) = 1 ,$$

from which we get the zero-coupon bond prices:

$$p(t,T) = e^{-R(t,T)\tau(t,T)} .$$

Definition 1.3. The *simply-compounded spot interest rate* at time t for the maturity T is the constant rate at which an investment of $p(t,T)$ unit of currency at time t accrues proportionally to the investment time to produce 1 unit of currency at time T.
It is denoted by $L(t,T)$ and is defined by:

$$L(t,T) := \frac{1 - p(t,T)}{\tau(t,T)p(t,T)} .$$

The most important interbank rate, as a reference for contracts, is the *LIBOR* (London InterBank Offered Rate) rate, fixing daily at 12 o' clock in London. This is a simply-compounded spot interest rate, from which derive the notation L for this last, and is typically linked to zero-coupon bond prices by the "Actual/360" day-count convention. Generally the term "LIBOR" refers also to analogous rates fixing in other markets, e.g. the EURIBOR rate (fixing in Bruxelles).
The bond prices in terms of LIBOR rate is:

$$p(t,T) = \frac{1}{1 + L(t,T)\tau(t,T)} .$$

Definition 1.4. The *annually-compounded spot interest rate* at time t for the maturity T is the constant rate at which an investment of $p(t,T)$ unit of currency at time t has to be reinvested once a year to produce 1 unit of currency at time T.
It is denoted by $Y(t,T)$ and is defined by:

$$Y(t,T) := \frac{1}{p(t,T)^{\frac{1}{\tau(t,T)}}} - 1\,.$$

The day-count convention typically associated to the annual compounding is the "Actual/365".
The bond prices in terms of these rates is:

$$p(t,T) = \frac{1}{(1+Y(t,T))^{\tau(t,T)}}\,.$$

An extension of the annual compounding case is the following.

Definition 1.5. The *k-times-per-year compounded spot interest rate* at time t for the maturity T is the constant rate at which an investment of $p(t,T)$ unit of currency at time t has to be reinvested k times a year to produce 1 unit of currency at time T.
It is denoted by $Y^k(t,T)$ and is defined by:

$$Y^k(t,T) := \frac{k}{p(t,T)^{\frac{1}{k\tau(t,T)}}} - k\,.$$

All the above spot interest rates are equivalent in infinitesimal time intervals. For this reason, we can define the short rate in the following way.

Definition 1.6. The *instantaneous short interest rate* at time t is the limit of each of the different spot rates between times t and T with $T \to t^+$.
It is denoted by $r(t)$ and is defined by:

$$\begin{aligned} r(t) &= \lim_{T\to t^+} R(t,T) \\ &= \lim_{T\to t^+} L(t,T) \\ &= \lim_{T\to t^+} Y(t,T) \\ &= \lim_{T\to t^+} Y^k(t,T) \quad \forall k\,. \end{aligned}$$

Based on this, we can define the mathematical representation of a bank account, which accrues continuously according to the instantaneous rate.

Definition 1.7. The *bank account* (or *money market account*) at time $t \geq 0$ is the value (the t-value) of a bank account with a unitary investment at initial time 0.
It is denoted by $B(t)$ and its dynamics is given by:

$$dB(t) = r(t)B(t)dt\,, \quad B(0) = 1\,,$$

solved by:

$$B(t) = \exp(\int_0^t r(s)ds)\,,$$

as shown in (C.2). Notice that, according to the market setting (C.1)-(C.2) and the theorem of change of drift with correlation, B is the only asset in the market which is not modified when moving to a risk-adjusted probability measure, in fact its instantaneous variation is not affected by a change of measure.
The bank account is a stochastic process that provide us with a model of the time value of money and allows us to build a discount factor of that value.

Definition 1.8. The *(stochastic) discount factor* between two time instants t and $T \geq t$ is the amount of money at time t equivalent (according to the dynamic of $B(t)$) at 1 unit of currency at time T.
It is denoted by $D(t,T)$ and is defined by:

$$D(t,T) = \frac{B(t)}{B(T)} = \exp\left(-\int_t^T r(s)ds\right)\,.$$

In some financial areas where the short rate r is considered a deterministic function of time (i.e. in markets where the variability of the rate is negligible with respect to movements of the underlying assets of options to be priced), both the bank account and the discount factor become deterministic processes and we have $D(t,T) = p(t,T)$ for all (t,T). However, when dealing with interest rate derivatives, the variability of primary importance is just that of the rates themselves. In this case r is modeled as a stochastic process and consequently $D(t,T)$ is a random value at time t which depends on the future evolution of r up to T, whereas $p(t,T)$ is the t-value (known) of a contract with maturity date T.

1.1 Forward Rates

Now we move to define the forward rates, that are characterized by three time instants: the present time t, at which they are locked in, and two points in future time, the *expiry* time T and the *maturity* time S, with $t \leq T \leq S$. The forward rates can be defined in two different ways.

First approach to define forward rates

We start defining a contract at the current time t which allows us to make an investment of 1 unit of currency at time T and to have a deterministic rate of return (determined at t) over the period $[T,S]$. Then we compute the relevant interest rate involved by solving an equation that avoids arbitrage. The corresponding financial strategy is the following:

Time	Operations	Portfolio value
t	sell one T-bond and use	
	the income $p(t,T)$ to buy $\frac{p(t,T)}{p(t,S)}$ S-bonds	0
T	pay out 1	-1
S	receive the amount $\frac{p(t,T)}{p(t,S)}$	$-1+\frac{p(t,T)}{p(t,S)}$

The net effect of all this, based on a contract made at time t, is that an investment of 1 unit of currency at time T has yielded $\frac{p(t,T)}{p(t,S)}$ at time S. Thus it guarantees a riskless rate of interest over the future interval $[T,S]$. Such an interest rate is what is called a *forward rate* and it can be characterized depending on the compounding type, as follows.

Definition 1.9. The *simply-compounded forward interest rate* at time t for the expiry $T>t$ and maturity $S>T$ is denoted by $F(t;T,S)$ and is defined by:

$$F(t;T,S) := \frac{1}{\tau(T,S)}\left(\frac{p(t,T)}{p(t,S)} - 1\right). \tag{1.1}$$

It is the solution to the equation

$$1 + \tau(T,S)F = \frac{p(t,T)}{p(t,S)}.$$

Definition 1.10. The *continuously-compounded forward interest rate* at time t for the expiry $T > t$ and maturity $S > T$ is denoted by $R(t;T,S)$ and is defined by:

$$R(t;T,S) := -\frac{\ln p(t,S) - \ln p(t,T)}{\tau(T,S)}.$$

It is the solution to the equation

$$e^{R\tau(T,S)} = \frac{p(t,T)}{p(t,S)}.$$

The simple rate notation is the one used in the market, whereas the continuous one is used in theoretical contexts and does not regard the model we are exposing.

Second approach to define forward rates

Definition 1.11. A *forward rate agreement* (*FRA*) is a contract stipulated at the current time t that gives its holder an interest rate payment for the period of time between the expiry $T > t$ and the maturity $S > T$. Precisely, at time S he receives a fixed payment based on a fixed rate K and pays a floating amount based on the spot rate $L(T,S)$ resetting in T.
Its S-payoff is thus:

$$N\tau(T,S)(K - L(T,S)),$$

where N is the contract nominal value.

This payoff can be rewritten by substituting the LIBOR rate with its expression:

$$N\left(\tau(T,S)K - \frac{1}{p(T,S)} + 1\right).$$

Now we use the principle of no arbitrage to calculate the value of the above FRA at time t. The amount $\frac{1}{p(T,S)}$ at time S equals to have $p(T,S)\frac{1}{p(T,S)} = 1$ unit of currency at time T, in turn this one equals to have $p(t,T)$ units of currency at time t. Instead, the amount $\tau(T,S)K + 1$ at time S equals to have $p(t,S)(\tau(T,S)K+1)$ units of currency at time t. Therefore, the t-value of the contract is

$$FRA(t,T,S,\tau(T,S),N,K) = N\left[p(t,S)\tau(T,S)K - p(t,T) + p(t,S)\right]). \tag{1.2}$$

Finally we can define the simply-compounded forward interest rate as the unique value of the fixed rate K which renders the FRA with expiry T and maturity S a fair contract at time t, i.e. such that the t-price (1.2) is 0, to achieve the same Definition 1.9 again. Then, the value of the FRA can be rewritten in terms of the forward rate as

$$\begin{aligned} FRA(t,T,S,\tau(T,S),N,K) &= N\left[p(t,S)\tau(T,S)K + p(t,S)\left(1 - \frac{p(t,T)}{p(t,S)}\right)\right] \\ &= N\left[p(t,S)\tau(T,S)K + \right. \\ &\qquad \left. + p(t,S)(-\tau(T,S)F(t;T,S))\right] \\ &= N\tau(T,S)p(t,S)\left(K - F(t;T,S)\right). \end{aligned} \tag{1.3}$$

Comparing the payoff and the price of the above FRA, we can view the forward rate $F(t;T,S)$ as a kind of estimate of the future spot rate $L(T,S)$, which is random at time t.

The last type of interest rate worthy to be mentioned is the analogous of the instantaneous short interest rate, in the future.

Definition 1.12. The *instantaneous forward interest rate* at time t for the maturity T is the limit of the forward rates expiring in T when collapsing towards their expiry.

It is denoted by $f(t)$ and is defined by:

$$\begin{aligned} f(t) &= \lim_{S \to T^+} F(t;T,S) \\ &= \lim_{S \to T^+} R(t;T,S) \\ &= -\frac{\partial \ln p(t,T)}{\partial T} . \end{aligned}$$

Interest Rate Swaps

Definition 1.13. Given the tenor structure $\mathcal{T} = \{T_\alpha, \dots, T_\beta\}$ with the corresponding set of year fractions $\tau = \{\tau_\alpha, \dots, \tau_\beta\}$, an *Interest Rate Swap* (*IRS*) with tenor $T_\beta - T_\alpha$ is a contract which, at every time $T_i \in \{T_{\alpha+1}, \dots, T_\beta\}$ exchanges the floating leg payment

$$N\tau_i L(T_{i-1}, T_i)$$

with the fixed leg payment

$$N\tau_i K ,$$

where N is the nominal value and K a fixed interest rate.

It can be of two types: the holder of a *Payer* IRS, denoted by *PFS*, receives the floating leg and pays the fixed leg, whereas the holder of a *Receiver* IRS, denoted by *RFS*, the opposite. The discounted payoff at a time $t < T_\alpha$ of a PFS is thus:

$$\sum_{i=\alpha+1}^{\beta} D(t,T_i) N\tau_i (L(T_{i-1},T_i) - K) ,$$

whereas a RFS has the opposite payoff.

The arbitrage-free value at $t < T_\alpha$ of a PFS with a unit notional amount is:

$$
\begin{aligned}
PFS(t,\mathcal{T},K) &= E\left[\sum_{i=\alpha+1}^{\beta} D(t,T_i)\,\tau_i\,(F_i(T_{i-1})-K)\,|\,\mathcal{F}_t\right] \\
&= \sum_{i=\alpha+1}^{\beta} p(t,T_i)\,\tau_i\,E^i\left[L(T_{i-1},T_i)-K\,|\,\mathcal{F}_t\right] \\
&= p(t,T_\alpha)-p(t,T_\beta)-K\sum_{i=\alpha+1}^{\beta}\tau_i\,p(t,T_i) \qquad (1.4)\\
&= \sum_{i=\alpha+1}^{\beta} p(t,T_i)\,\tau_i\,(F_i(t)-K)\,, \qquad (1.5)
\end{aligned}
$$

where denoting $F_j(t) := F(t;T_{j-1},T_j)$.

Proof. The formula (1.4) can be obtained analogously to the price of a FRA: the floating payment set in T_{i-1} and payed in T_i can be rewritten as

$$\frac{1}{p(T_{i-1},T_i)}-1\,,$$

that equals to have

$$p(t,T_{i-1})-p(t,T_i) \qquad (1.6)$$

at time $t < T_\alpha$, so that the arbitrage-free value at t of the whole floating side is

$$\sum_{i=\alpha+1}^{\beta}\left(p(t,T_{i-1})-p(t,T_i)\right)=p(t,T_\alpha)-p(t,T_\beta)\,;$$

on the other hand the amount $\tau_i K$ at time T_i equals to have $p(t,T_i)\tau_i K$ units of currency at time t, so that the arbitrage-free value at t of the whole fixed side is

$$\sum_{i=\alpha+1}^{\beta} p(t,T_i)\tau_i K = K\sum_{i=\alpha+1}^{\beta} p(t,T_i)\tau_i\,.$$

Equivalently we could have obtained (1.6) by the risk neutral pricing formula (D.6):

$$\begin{aligned}
&E^Q\left[D(t,T_i)\left(\frac{1}{p(T_{i-1},T_i)}-1\right) \mid \mathcal{F}_t^W\right] = \\
=\ & E^Q\left[e^{-\int_t^{T_i} r_s ds}\frac{1}{p(T_{i-1},T_i)} \mid \mathcal{F}_t^W\right] - E^Q\left[D(t,T_i)\mid \mathcal{F}_t^W\right] \\
=\ & E^Q\left[e^{-\int_t^{T_i} r_s ds} E^Q\left[e^{\int_{T_{i-1}}^{T_i} r_s ds} \mid \mathcal{F}_{T_{i-1}}^W\right] \mid \mathcal{F}_t^W\right] - p(t,T_i) \\
=\ & E^Q\left[e^{-\int_t^{T_{i-1}} r_s ds} \mid \mathcal{F}_t^W\right] - p(t,T_i) \ = \ p(t,T_{i-1}) - p(t,T_i)\,.
\end{aligned}$$

Then, the formula (1.5) can be obtained by substituting the expression for forward rates (1.1), as we made in (1.3). □

From another point of view, a RFS can be view as a portfolio of FRAs, valued through formulas (1.2) or (1.3), leading to a price opposite to that of a PFS.

Definition 1.14. The *forward swap rate* (or *par swap rate*) of a $T_\alpha \times (T_\beta - T_\alpha)$ Interest Rate Swap is denoted by $S_{\alpha,\beta}(t)$ at time t and is defined as that value of the fixed rate that makes the IRS a fair contract at the current time t. It is obtained by equating to zero the t-value of the contract in (1.4):

$$S_{\alpha,\beta}(t) := \frac{p(t,T_\alpha) - p(t,T_\beta)}{\sum\limits_{i=\alpha+1}^{\beta} \tau_i\, p(t,T_i)}\,. \tag{1.7}$$

Remark 1. It can be rewritten as a nonlinear function of the forward LIBOR rates as

$$S_{\alpha,\beta}(t) = \frac{1 - \prod\limits_{j=\alpha+1}^{\beta} \frac{1}{1+\tau_j F_j(t)}}{\sum\limits_{i=\alpha+1}^{\beta} \tau_i \prod\limits_{j=\alpha+1}^{i} \frac{1}{1+\tau_j F_j(t)}}\,. \tag{1.8}$$

Proof.

$$S_{\alpha,\beta}(t) \ = \ \frac{p(t,T_\alpha) - p(t,T_\beta)}{\sum\limits_{i=\alpha+1}^{\beta} \tau_i\, p(t,T_i)} \ = \ \frac{\frac{p(t,T_\alpha)-p(t,T_\beta)}{p(t,T_\alpha)}}{\sum\limits_{i=\alpha+1}^{\beta} \tau_i \frac{p(t,T_i)}{p(t,T_\alpha)}}\,,$$

then, from the definition (1.1) of forward rates and by a simple algebraic relation, we have:

$$\frac{p(t,T_i)}{p(t,T_\alpha)} = \prod_{j=\alpha+1}^{i} \frac{p(t,T_j)}{p(t,T_{j-1})} = \prod_{j=\alpha+1}^{i} \frac{1}{1+\tau_j F_j(t)} \quad \forall i > \alpha\,.$$

□

1.2 Main derivatives

In this chapter we present the two main interest rate derivatives, that are caps/floors and swaptions.

Interest rate Caps/Floors

Definition 1.15. An interest rate *cap* is a financial insurance contract equivalent to a payer interest rate swap where each exchange payment is executed if and only if it has positive value. The discounted payoff at time t of the cap associated to the tenor structure $\mathcal{T} = \{T_\alpha, \dots, T_\beta\}$, with the corresponding set of year fractions $\tau = \{\tau_\alpha, \dots, \tau_\beta\}$, and working on a principal amount of money N, is

$$\sum_{i=\alpha+1}^{\beta} D(t,T_i) N \tau_i (L(T_{i-1},T_i) - K)^+\,. \tag{1.9}$$

Analogously, a *floor* is a contract equivalent to a receiver interest rate swap where each exchange payment is executed if and only if it has positive value, with t-discounted payoff

$$\sum_{i=\alpha+1}^{\beta} D(t,T_i) N \tau_i (K - L(T_{i-1},T_i))^+\,.$$

A cap has the task of protecting the holder, indebted with a loan at a floating rate of interest, from having to pay more than a prespecified rate K, called the *cap rate*. On the other hand, a floor guarantees that the interest paid on a floating rate loan will never be below a predetermined *floor rate*.

A cap consists in a portfolio of a number of more basic contracts, named *caplets*: the i-th caplet is determined at time T_{i-1} but not paid out until time T_i and has the t-discounted payoff

$$D(t,T_i)N\tau_i(L(T_{i-1},T_i)-K)^+ .$$

Analogously are defined the *floorlet* contracts.

The market practice is to price caps by using the Black's formula for caps, an extension of the Black and Scholes formula, dating back to 1976 when Black had to price the payoff of commodity options. The price at time t of the cap with tenor $\mathcal{T}$ and unit notional amount is

$$\mathrm{Cap}^{\mathrm{Black}}(t,\mathcal{T},\tau,K,v) = \sum_{i=\alpha+1}^{\beta} \tau_i\, p(t,T_i)\,\mathrm{Bl}\left(K, F(t;T_{i-1},T_i), v_i)\right) , \qquad (1.10)$$

where

$$\begin{aligned}\mathrm{Bl}(K,F(t;T_{i-1},T_i),v_i)) :=\ & F(t;T_{i-1},T_i)\,\Phi\left(d_1(K,F(t;T_{i-1},T_i),v_i)\right) + \\ & -K\Phi\left(d_2(K,F(t;T_{i-1},T_i),v_i)\right) ,\end{aligned}$$

$$\begin{aligned} d_1(K,F,u) &:= \frac{\ln\left(\frac{F}{K}\right)+\frac{u^2}{2}}{u} , \\ d_2(K,F,u) &:= \frac{\ln\left(\frac{F}{K}\right)-\frac{u^2}{2}}{u} , \\ v_i &:= v\sqrt{T_{i-1}} , \end{aligned}$$

with the common volatility parameter v that is retrieved from market quotes. Analogously, the Black's formula for caplets is

$$\mathrm{Capl}^{\mathrm{Black}}(t,T_{i-1},T_i,K,v_i) = \tau_i\, p(t,T_i)\,\mathrm{Bl}\left(K,F(t;T_{i-1},T_i),v_i)\right) , \qquad (1.11)$$

for all $i=\alpha+1,\ldots,\beta$.

Swaptions

Definition 1.16. A European $T_\alpha \times (T_\alpha - T_\beta)$ *Payer Swaption* (*PS*) with swaption strike K is a contract that gives the right (but not the obligation) to enter a PFS with tenor $T_\beta - T_\alpha$ and fixed rate K at the future time T_α, i.e. the swaption maturity.

Namely, a swaption is an option of an IRS. The payer swaption payoff at its first reset date T_α, which is also the swaption maturity, is:

$$(PFS(T_\alpha, \mathcal{T}, K))^+ = \left(\sum_{i=\alpha+1}^{\beta} p(T_\alpha, T_i)\, \tau_i \,(F(T_\alpha; T_{i-1}, T_i) - K) \right)^+ \quad (1.12)$$

$$= (S_{\alpha,\beta}(T_\alpha) - K)^+ \sum_{i=\alpha+1}^{\beta} \tau_i \, p(T_\alpha, T_i) \quad (1.13)$$

respectively in terms of forward rates and of the relevant forward swap rate.

Proof. The formula (1.12) is obtained simply by taking the positive part of the T_α-price of a PFS in the form (1.5), whilst the formula (1.13) is obtained from the other version (1.4) of the same price as follows:

$$(PFS(T_\alpha, \mathcal{T}, K))^+ = \left(p(T_\alpha, T_\alpha) - p(T_\alpha, T_\beta) - K \sum_{i=\alpha+1}^{\beta} \tau_i \, p(T_\alpha, T_i) \right)^+$$

$$= (S_{\alpha,\beta}(T_\alpha) - K)^+ \sum_{i=\alpha+1}^{\beta} \tau_i \, p(T_\alpha, T_i)\,.$$

□

The market practice is to price swaptions by using the Black's formula for swaptions: the price at time t of the above $T_\alpha \times (T_\alpha - T_\beta)$ payer swaption is

$$\mathrm{PS}^{\mathrm{Black}}(t, T_\alpha, \mathcal{T}, K, v_{\alpha,\beta}) = \sum_{i=\alpha+1}^{\beta} \tau_i \, p(t, T_i)\, \mathrm{Bl}\left(K, S_{\alpha,\beta}(t), v_{\alpha,\beta}\sqrt{T_\alpha - t}\right)\,. \quad (1.14)$$

A *Receiver Swaption* is defined analogously as an option on a RFS.

Chapter 2

The LIBOR Market Model (LMM)

For a very long time, namely since the early '80 to 1996, the market practice has been to value caps, floors and swaptions by using a formal extension of the Black (1976) model. However, this formula was applied in a completely heuristic way, under some simplifying and inexact assumptions. Indeed, interest rate derivatives were priced by using short rate models, based on modeling the instantaneous short interest rate; at one point this was assumed to be deterministic, so that the discount factor was identified with the corresponding bond price, that could be factorized out of the Q-expectation in the risk-neutral pricing formula; then, inconsistently with the previous assumption, the forward LIBOR rates were modeled as driftless geometric Brownian motions under Q, hence stochastic; finally the expectation could be view as the price of a call option in a market with zero risk-free rate, therefore it was obtained through the Black's formula. This is logically inconsistent.

Then, at the end of the '90, after the coming of the theory of the change of numeraire, a promising family of (arbitrage-free) interest rate models was introduced: the Market Models. This breakthrough came at the hands of *Miltersen et al* (1997), *Brace et al* (1997) and *Jamshidian* (1998). The principal idea of these approaches is to choose a different numeraire than the

risk-free bank account.
The interest rate market is radically different from the others, e.g. commodities or equities, thus needs a own kind of modeling. There are three possible choices in interest rate modeling: short rate models, that model one single variable, instantaneous forward rate models, that model all infinite points of the term structure and Market Models.
These recent ones have the following characteristics:

- instead of modeling instantaneous interest rates, they model a selection of discrete real world rates (quoted in the market) spanning the term structure;

- under a suitable change of numeraire these market rates can be modeled log-normally;

- they produce pricing formulas for caps, floors and swaptions of the Black-76 type;

- they are easy to calibrate to market data and are then used to price more exotic products.

The model we are introducing is best known generally as "LIBOR Market Model" (LMM), or else "Log-normal Forward LIBOR Model" or "Brace-Gatarek-Musiela 1997 Model" (BGM model), from the names of the authors of the first published papers that rigorously described it.

Setting the model:

- $t = 0$ is the current time;

- the set $\{T_0, T_1, \dots, T_M\}$ of expiry-maturity dates (expressed in years) is the tenor structure, with the corresponding year fractions $\{\tau_0, \tau_1, \dots, \tau_M\}$, i.e. τ_i is the one associated with the expiry-maturity pair (T_{i-1}, T_i), for all $i > 0$, and τ_0 from now to T_0;

- set $T_{-1} := 0$;

- the simply-compounded forward interest rate resetting at its expiry date T_{i-1} and with maturity T_i is denoted by $F_i(t) := F(t; T_{i-1}, T_i)$ and is alive up to time T_{i-1}, where it coincides with the spot LIBOR rate $F_i(T_{i-1}) = L(T_{i-1}, T_i)$, for $i = 1, \ldots, M$;
- there exists an arbitrage-free market bond, where an EMM Q exists and the bond prices $p(\cdot, T_i)$ are Q-prices, for $i = 1, \ldots, M$;
- Q^i is the EMM associated with the numeraire $p(\cdot, T_i)$, i.e. the T_i-*forward measure*;
- Z^i is the M-dimensional correlated Brownian motion under Q^i, with instantaneous correlation matrix ρ.

Lemma 2.0.1. *For every $i = 1, \ldots, M$ the forward LIBOR process F_i is a martingale under the corresponding T_i-forward measure, on the interval $[0, T_{i-1}]$.*

Proof. From the definition of forward rates we have

$$F_i(t)p(t, T_i) = \frac{p(t, T_{i-1}) - p(t, T_i)}{\tau_i}.$$

Since $p(t, T_{i-1})$ and $p(t, T_i)$ are tradable assets, hence Q-prices, $F_i(t)$ is a Q-price too. Thus, when normalizing it by the numeraire $p(\cdot, T_i)$, it has to be a martingale under Q^i on the interval $[0, T_{i-1}]$. □

Modeling the F's as diffusion processes, it follows that F_i has a driftless dynamics under Q^i.

Definition 2.1. A discrete tenor *LIBOR market model* assumes that the forward rates have the following dynamics under their associated forward measures:

$$dF_i(t) = \sigma_i(t)F_i(t)dZ_i^i(t)\,, \quad t \leq T_{i-1}\,, \quad \text{for } i = 1, \ldots, M \tag{2.1}$$

where the percentage instantaneous volatility process of F_i, σ_i, is assumed to be deterministic and scalar, whereas dZ_i^i is the i-th component of the Q^i-Brownian motion, hence is a standard B.m. (Observation 13 in Appendix C).

There exist also extensions of this model where the scalar volatility $\sigma_i(t)$ are positive stochastic processes.

Notice that, if σ_i is bounded, the SDE (2.1) has a unique strong solution, since it describes a geometric Brownian motion. Indeed, by Itô's formula,

$$\begin{aligned} & d\ln F_i(t) = \sigma_i(t)dZ_i^i(t) - \tfrac{\sigma_i(t)^2}{2}dt \\ \Rightarrow \quad & \ln F_i(T) = \ln F_i(t) + \textstyle\int_t^T \sigma_i(s)dZ_i^i(s) - \int_t^T \tfrac{\sigma_i(s)^2}{2}ds \\ \Rightarrow \quad & F_i(T) = F_i(t)\, e^{\int_t^T \sigma_i(s)dZ_i^i(s) - \frac{1}{2}\int_t^T \sigma_i(s)^2 ds}\,, \quad 0 \le t \le T \le T_{i-1}\,. \end{aligned}$$

Proposition 2.0.2 (Forward measure dynamics in the LMM)**.** *Under the assumptions of the LIBOR market model, the dynamics of each F_k, for $k = 1, \ldots, M$, under the forward measure Q^i with $i \in \{1, \ldots, M\}$, is:*

$$\begin{aligned} k < i: \quad & dF_k(t) = -\sigma_k(t)F_k(t)\sum_{j=k+1}^{i} \tfrac{\rho_{k,j}\tau_j\sigma_j(t)F_j(t)}{1+\tau_j F_j(t)}dt + \sigma_k(t)F_k(t)dZ_k^i(t)\,, \\ k = i: \quad & dF_k(t) = \sigma_k(t)F_k(t)dZ_k^i(t)\,, \\ k > i: \quad & dF_k(t) = \sigma_k(t)F_k(t)\sum_{j=i+1}^{k} \tfrac{\rho_{k,j}\tau_j\sigma_j(t)F_j(t)}{1+\tau_j F_j(t)}dt + \sigma_k(t)F_k(t)dZ_k^i(t)\,, \end{aligned} \tag{2.2}$$

for $t \le \min\{T_{k-1}, T_i\}$.

Proof. By assumption, there exist a LIBOR market model satisfying (2.1). We try to determine the deterministic functions $\mu_k^i(t, F(t))$, where:
$F(t) = (F_1(t), \ldots, F_M(t))'$, that satisfies

$$dF_k(t) = \mu_k^i(t, F(t))F_k(t)dt + \sigma_k(t)F_k(t)dZ_k^i(t)\,, \quad k \neq i\,. \tag{2.3}$$

In order to find $\mu_k^i(t, F(t))$, i.e. the percentage drift of dF_k under Q^i, we're going to apply the change of measure from Q^i to Q^k, then impose that the Q^k-resulting drift is null. From Corollary D.0.14, the Radon-Nikodym derivative of Q^{i-1} w.r.t. Q^i at time t is

$$\frac{dQ^{i-1}}{dQ^i}|\mathcal{F}_t^W = \frac{p(t, T_{i-1})p(0, T_i)}{p(0, T_{i-1})p(t, T_i)} =: \gamma_t^i$$

and, by (1.9),

$$\gamma_t^i = \frac{p(0,T_i)}{p(0,T_{i-1})}(1+F_i(t)\tau_i)\,.$$

Therefore, assuming (2.1), the dynamics of γ_i under Q^i is

$$\begin{aligned} d\gamma_t^i &= \frac{p(0,T_i)}{p(0,T_{i-1})}dF_i(t)\tau_i = \frac{p(0,T_i)}{p(0,T_{i-1})}\tau_i\sigma_i(t)F_i(t)dZ_i^i(t) \\ &= \frac{\gamma_t^i}{1+F_i(t)\tau_i}\tau_i\sigma_i(t)F_i(t)dZ_i^i(t)\,. \end{aligned}$$

Thus, the density γ_i is in the form of an exponential martingale with associated process λ that is the d-dimensional null vector apart from the i-th component,

$$\lambda = \begin{pmatrix} 0 & \cdots & -\frac{\tau_i\sigma_i F_i}{1+F_i\tau_i} & \cdots & 0 \end{pmatrix}'\,, \tag{2.4}$$

so that we can apply the formula (C.3) of the change of drift with correlation:

$$dZ^i(t) = dZ^{i-1}(t) - \rho\lambda dt\,,$$

namely with components

$$dZ_j^i(t) = dZ_j^{i-1}(t) + \rho^{ji}\frac{\tau_i\sigma_i(t)F_i(t)}{1+F_i(t)\tau_i}dt\,.$$

Applying this inductively we obtain:

$$\begin{aligned} k<i:\quad & dZ_j^i(t) = dZ_j^k(t) + \sum_{h=k+1}^{i}\rho^{jh}\tfrac{\tau_h\sigma_h(t)F_h(t)}{1+F_h(t)\tau_h}dt\,; \\ k>i:\quad & dZ_j^i(t) = dZ_j^k(t) - \sum_{h=i+1}^{k}\rho^{jh}\tfrac{\tau_h\sigma_h(t)F_h(t)}{1+F_h(t)\tau_h}dt\,. \end{aligned}$$

Then, inserting these into (2.3) and equating the Q^k-drift to zero, we have:

$$\begin{aligned} k<i:\quad & F_k(t)\left(\mu_k^i(t,F(t)) + \sigma_i(t)\sum_{h=k+1}^{i}\rho^{jh}\tfrac{\tau_h\sigma_h(t)F_h(t)}{1+F_h(t)\tau_h}\right)dt = 0 \\ & \Rightarrow\quad \mu_k^i(t,F(t)) = -\sigma_i(t)\sum_{h=k+1}^{i}\rho^{jh}\tfrac{\tau_h\sigma_h(t)F_h(t)}{1+F_h(t)\tau_h}\,; \\ k>i:\quad & F_k(t)\left(\mu_k^i(t,F(t)) - \sigma_i(t)\sum_{h=i+1}^{k}\rho^{jh}\tfrac{\tau_h\sigma_h(t)F_h(t)}{1+F_h(t)\tau_h}\right)dt = 0 \\ & \Rightarrow\quad \mu_k^i(t,F(t)) = \sigma_i(t)\sum_{h=i+1}^{k}\rho^{jh}\tfrac{\tau_h\sigma_h(t)F_h(t)}{1+F_h(t)\tau_h}\,. \end{aligned}$$

□

At this point, we can turn around the argument to have the following existence result.

Proposition 2.0.3. *Consider a given volatility structure* $\sigma_1, \ldots, \sigma_M$*, where each* σ_i *is bounded, and the terminal measure* Q^M *with associated* d*-dimensional correlated B.m.* W^M*. If we define the processes* $F_1, \ldots, F_M$ *by*

$$dF_i(t) = -\sigma_i(t)F_i(t)\sum_{j=i+1}^{M}\frac{\rho_{i,j}\tau_j\sigma_j(t)F_j(t)}{1+\tau_jF_j(t)}dt + \sigma_i(t)F_i(t)dZ_i^M(t)\,, \qquad (2.5)$$

for $i = 1, \ldots, M$*, then the* Q^i*-dynamics of* F_i *is given by (2.1), i.e. there exists a LIBOR model with the given volatility structure.*

Proof. First, we have to prove the existence of a solution of (2.5). For $i = M$ we simply have

$$dF_M = \sigma_M(t)F_M(t)dZ_M^M(t)\,,$$

which is just an exponential martingale, where σ_M is bounded, thus a solution does exist. Now we proceed by induction: assume that (2.5) admits a solution for $i+1, \ldots, M$, then we write the i-th dynamics as

$$dF_i(t) = \mu_i(t, F_{i+1}(t), \ldots, F_M(t))F_i dt + \sigma_i(t)F_i(t)dZ_i^M(t)\,,$$

where the crucial fact is that μ_i depends only on F_K for $k = i+1, \ldots, M$. Thus, denoting $F_{i+1}^M := (F_{i+1}, \ldots, F_M)'$, we can solve explicitly the above SDE by applying the Itô formula:

$$\begin{aligned} d\ln F_i(t) &= \frac{dF_i(t)}{F_i(t)} - \frac{1}{2F_i(t)^2}\sigma_i(t)^2F_i(t)^2dt \\ &= \mu_i(t, F_{i+1}^M(t))dt + \sigma_i(t)dZ_i^M(t) - \frac{1}{2}\sigma_i(t)^2dt \end{aligned}$$

$$\Rightarrow \quad \ln F_i(t) = \ln F_i(0) + \int_0^t \left(\mu_i(s, F_{i+1}^M(s)) - \tfrac{\sigma_i(s)^2}{2}\right)ds + \int_0^t \sigma_i(s)dZ_i^i(s)$$

$$\Rightarrow \quad F_i(t) = F_i(t)\exp\left[\int_0^t \left(\mu_i(s, F_{i+1}^M(s)) - \tfrac{\sigma_i(s)^2}{2}\right)ds\right]\exp\left[\int_0^t \sigma_i(s)dZ_i^i(s)\right]\,,$$

for $0 \leq t \leq T_{i-1}$. This proves existence.

Then, we have to prove that the process λ defined in (2.4) satisfies the

Novikov condition (B.1), in which case the density process γ^i is a Q_i-martingale and consequently we can apply the Girsanov Theorem, retracing the same steps as in the proof of Proposition 2.0.2. In this regard, given an initial positive LIBOR term structure, as it is $F(0) = (F_1(0), \dots, F_M(0))'$, notice that all LIBOR rate processes will be always positive, thus the process λ in (2.4) is bounded and consequently satisfies the Novikov condition. □

The "log-normal forward LIBOR model" takes his name from the log-normal distribution of each forward rate under the related forward measure and we find it in the literature with several names. Anyway, the most common terminology remains that of "LIBOR Market Model".

2.1 Pricing Caps in the LMM

In the market, cap prices are not quoted in monetary terms, but rather in terms of the so-called implied Black volatilities. Typically, caps whose implied volatilities are quoted have resettlement dates $T_\alpha, \dots, T_\beta$ with either $\alpha = 0$, $T_0 = 3$ months and all other T_i's equally three-months spaced, or $\alpha = 0$, $T_0 = 6$ months and all other T_i's equally six-months spaced.

Definition 2.2. Given market price data for caps with tenor structure as above mentioned, denoted by $\mathrm{Cap}^{\mathrm{m}}(t, \mathcal{T}_j, K)$ where $\mathcal{T}_j = \{T_0, \dots, T_j\}$, the *implied Black volatilities* are defined as follows:

- the implied *flat* volatilities are the solutions $v_{T_1-cap}, \dots, v_{T_M-cap}$ of the equations

$$\mathrm{Cap}^{\mathrm{m}}(t, \mathcal{T}_j, K) = \sum_{i=1}^{j} \mathrm{Capl}^{\mathrm{Black}}(t, T_{i-1}, T_i, K, v_{T_j-cap}),$$

  $j = 1, \dots, M$;

- the implied *spot* volatilities are the solutions $v_{T_1-capl}, \dots, v_{T_M-capl}$ of the equations

$$\mathrm{Capl}^{\mathrm{m}}(t, T_{i-1}, T_i, K) = \mathrm{Capl}^{\mathrm{Black}}(t, T_{i-1}, T_i, K, v_{T_i-capl}),$$

$i = 1, \dots, M$, where

$$\mathrm{Capl}^{\mathrm{m}}(t, T_{i-1}, T_i, K) = \mathrm{Cap}^{\mathrm{m}}(t, \mathcal{T}_i, K) - \mathrm{Cap}^{\mathrm{m}}(t, \mathcal{T}_{i-1}, K)\,.$$

Notice that the flat volatility v_{T_j-cap} is that implied by the Black formula by putting the same average volatility in all caplets up to T_j, whereas the spot volatility v_{T_i-capl} is just the implied average volatility from caplet over $[T_{i-1}, T_i]$.

Remark 2. There seems to be one kind of inconsistency in the cap volatility system. Indeed, when considering a set of caplets all concurring to different caps, their average volatilities change moving from a cap to another, if computed as implied flat volatilities. Therefore, to recover correctly cap prices according to the LMM dynamics, we need to have

$$\begin{aligned} &\sum_{i=1}^{j} \tau_i\, p(t, T_i)\, \mathrm{Bl}\left(K, F(t; T_{i-1}, T_i), \sqrt{T_{i-1}} v_{T_j-cap})\right) = \\ &= \sum_{i=1}^{j} \tau_i\, p(t, T_i)\, \mathrm{Bl}\left(K, F(t; T_{i-1}, T_i), \sqrt{T_{i-1}} v_{T_i-capl})\right)\,, \end{aligned} \tag{2.6}$$

for all $j = 1, \dots, M$.

Recalling the t-discounted payoff (1.9) for a cap with tenor $\mathcal{T} = \{T_\alpha, \dots, T_\beta\}$, year fractions τ, cap rate K and unit notional amount, we have that its price, given by the risk neutral valuation formula, is

$$\begin{aligned} &E^Q\left[\sum_{i=\alpha+1}^{\beta} D(t, T_i)\tau_i (L(T_{i-1}, T_i) - K)^+ \,|\, \mathcal{F}_t\right] = \\ &= \sum_{i=\alpha+1}^{\beta} \tau_i E^Q\left[D(t, T_i)(L(T_{i-1}, T_i) - K)^+ \,|\, \mathcal{F}_t\right]\,, \end{aligned} \tag{2.7}$$

but moving from the probability measure Q with numeraire B to the T_i-forward measure in each i-th summand, as in (D.7), we have

$$\sum_{i=\alpha+1}^{\beta} \tau_i p(t, T_i) E^i\left[(L(T_{i-1}, T_i) - K)^+ \,|\, \mathcal{F}_t\right]\,.$$

Notice that the joint dynamics of forward rates is not involved in the pricing of a cap, because its payoff is reduced to a sum of payoffs of the caplets involved. Consequently, marginal distributions of forward rates are enough to compute the expectation and the correlation between them does not matter. The above expectation in computed easily, remembering the log-normal distribution of the F_i's under the related Q_i's.

Proposition 2.1.1 (Equivalence between LMM and Black's caplet prices)**.** *The price of the i-th caplet implied by the LIBOR market model coincides with that given by the corresponding Black caplet formula:*

$$\begin{aligned}\mathrm{Capl}^{\mathrm{LMM}}(0,T_{i-1},T_i,K) &= \mathrm{Capl}^{\mathrm{Black}}(0,T_{i-1},T_i,K,v_i)\\ &= \tau_i\, p(0,T_i)\,\mathrm{Bl}\left(K,F(0;T_{i-1},T_i),v_i\sqrt{T_{i-1}})\right)\ ,\end{aligned}$$

where

$$(v_i)^2 = \frac{1}{T_{i-1}}\int_0^{T_{i-1}} \sigma_i(t)^2 dt\,. \tag{2.8}$$

Proof.

$$\mathrm{Capl}^{\mathrm{LMM}}(t,T_{i-1},T_i,K) = \tau_i p(t,T_i) E^i\left[(L(T_{i-1},T_i)-K)^+ \mid \mathcal{F}_t\right]\ ,$$

where $L(T_{i-1},T_i) = F(T_{i-1};T_{i-1},T_i) = F_i(T_{i-1})$ and, for $T < T_{i-1}$, $F_i(T)$ is log-normal distributed under the forward measure Q^i, indeed

$$F_i(T) = F_i(t)e^{\int_t^T \sigma_i(s)dZ_i^i(s) - \frac{1}{2}\int_t^T \sigma_i(s)^2 ds}\,,\quad 0\le t\le T\le T_{i-1}\,,$$

with σ_i assumed to be deterministic. Let

$$Y_i(t,T) := \int_t^T \sigma_i(s)dZ_i^i(s) - \frac{1}{2}\int_t^T \sigma_i(s)^2 ds\,,$$

we have

$$F_i(T) = F_i(t)e^{Y_i(t,T)}\,,\quad Y_i(t,T)\sim\mathcal{N}(m_i(t,T),\Sigma_i^2(t,T))\,,$$

where

$$m_i(t,T) = -\frac{1}{2}\int_t^T \sigma_i(s)^2 ds\,,\quad \Sigma_i^2(t,T) = \int_t^T \sigma_i(s)^2 ds$$

and $F_i(t)\in\mathbb{R}$ at time t. Thus the proof follows from (A.5). □

The quantity v_i in (2.8), denote the average (standardized with respect to time) instantaneous percentage variance of the forward rate $F_i(t)$ for $t\in[0,T_{i-1}]$, that is its average volatility.

Chapter 3

The Swap Market Model (SMM)

We are going to illustrate the counterpart of the LIBOR market model among the market models, i.e. the "Swap Market Model" (SMM), which models the evolution of the forward swap rates instead of the one of the forward LIBOR rates, these two kind of rates being the bases of the two main markets in the interest rate derivatives world. The SMM is also referred to as "Log-Normal Forward Swap Model" or "Jamshidian 1997 Market Model".

The settings of this model are still the same of the LMM.

From the formula (1.13) for the T_α-price of a $T_\alpha \times (T_\beta - T_\alpha)$ payer swaption, it comes clearly that the natural choice of numeraire to model the dynamics of the forward swap rate is the process

$$C_{\alpha,\beta}(t) := \sum_{i=\alpha+1}^{\beta} \tau_i \, p(t, T_i) \, ,$$

which is referred to as the *accrual factor* or the *present value of a basis point*, given $\alpha, \beta \in \{0, \ldots, M\}$, $\alpha < \beta$. Moreover the accrual factor has the representation of the value at a time t of a traded asset that is a buy-and-hold portfolio consisting, for each i, of τ_i units of the zero coupon bond maturing at T_i, thus it is a plausible numeraire.

Lemma 3.0.2. *Denoted by $Q^{\alpha,\beta}$ the EMM associated with the numeraire $C_{\alpha,\beta}$, the forward swap rate process $S_{\alpha,\beta}$ is a martingale under $Q^{\alpha,\beta}$, on the interval $[0, T_\alpha]$.*

Proof. This follows immediately from the definition (1.7) of the forward swap rate, in fact the product

$$C_{\alpha,\beta}(t)S_{\alpha,\beta}(t) = p(t, T_\alpha) - p(t, T_\beta)$$

gives the t-price of a tradable asset, whose discounted process by the numeraire $C_{\alpha,\beta}$ has to be a $Q^{\alpha,\beta}$-martingale, by Theorem D.0.13 and the properties of an EMM. □

The probability measure $Q^{\alpha,\beta}$ is called the (*forward*) *swap measure* related to α, β. We may note that the accrual factors play for the swap rate the same role as the the zero coupon bond prices did for the forward rates in the LIBOR market model. The model we are defining is founded on this basis.

Definition 3.1. Consider a fixed a subset $\mathcal{T}^{pairs}$ of all pairs (α, β) of integer indexes such that $0 \leq \alpha < \beta \leq M$ of the resettlement dates in the tenor structure $\{T_0, T_1, \ldots, T_M\}$ and consider for each pair a deterministic function of time $t \mapsto \sigma_{\alpha,\beta}(t)$. A *swap market model* (*SMM*) with volatilities $\sigma_{\alpha,\beta}$ assumes that the forward swap rates have the following dynamics under their associated swap measures:

$$dS_{\alpha,\beta}(t) = \sigma_{\alpha,\beta}(t)S_{\alpha,\beta}(t)dW^{\alpha,\beta}(t)\,, \quad t \leq T_\alpha\,, \tag{3.1}$$

for $(\alpha, \beta) \in \mathcal{T}^{pairs}$, where $W^{\alpha,\beta}$ is a scalar standard $Q^{\alpha,\beta}$-Brownian motion.

We can also allows for correlation between the different Brownian motions, however, this will not affect the swaption prices but only the pricing of more complicated products.

Remark 3. In a model with $M + 1$ resettlement dates it is possible to model only M swap rates as independent. The two typical choices of possible $\mathcal{T}^{pairs}$ identify the following substructures:

- the *regular SMM*, which models the swap rates $S_{0,M}, S_{1,M}, \ldots, S_{M-1,M}$, i.e.
$$\mathcal{T}^{pairs} = \{(0,M),(1,M),\ldots,(M-1,M)\}\,;$$
- the *reverse SMM*, which models the swap rates $S_{0,1}, S_{0,2}, \ldots, S_{0,M}$, i.e.
$$\mathcal{T}^{pairs} = \{(0,1),(0,2),\ldots,(0,M)\}\,.$$

3.1 Pricing Swaptions in the SMM

In a swap market model, the pricing of swaptions result trivial and exactly analogous to the pricing of caplets in the LMM.

Proposition 3.1.1 (Equivalence between SMM and Black's swaption prices)**.** *The price of a $T_\alpha \times (T_\beta - T_\alpha)$ payer swaption implied by the swap market model coincides with that given by the corresponding Black swaptions formula:*

$$\begin{aligned}\mathrm{PS}^{\mathrm{SMM}}(0,T_\alpha,\{T_\alpha,\ldots,T_\beta\},K) &= \mathrm{PS}^{\mathrm{Black}}(0,T_\alpha,\{T_\alpha,\ldots,T_\beta\},K,v_{\alpha,\beta}(T_\alpha))\\ &= C_{\alpha,\beta}(0)\,\mathrm{Bl}\Big(K,S_{\alpha,\beta}(0),\sqrt{T_\alpha}v_{\alpha,\beta}(T_\alpha))\Big)\,,\end{aligned}$$

where
$$v_{\alpha,\beta}(T)^2 = \frac{1}{T_\alpha}\int_0^T \sigma_{\alpha,\beta}(t)^2 dt\,. \tag{3.2}$$

Proof. From (1.13), the risk neutral valuation formula at time t for the price of the above swaption is

$$\begin{aligned}&\mathrm{PS}^{\mathrm{SMM}}(t,T_\alpha,\{T_\alpha,\ldots,T_\beta\},K) =\\ &= E^Q\left[D(t,T_\alpha)\,(S_{\alpha,\beta}(T_\alpha)-K)^+\,C_{\alpha,\beta}(T_\alpha)\,|\,\mathcal{F}_t\right]\\ &= E^{Q^{\alpha,\beta}}\left[\tfrac{C_{\alpha,\beta}(t)}{C_{\alpha,\beta}(T_\alpha)}\,(S_{\alpha,\beta}(T_\alpha)-K)^+\,C_{\alpha,\beta}(T_\alpha)\,|\,\mathcal{F}_t\right]\\ &= C_{\alpha,\beta}(t)\,E^{Q^{\alpha,\beta}}\left[(S_{\alpha,\beta}(T_\alpha)-K)^+\,|\,\mathcal{F}_t\right]\,,\end{aligned}$$

by moving from the probability measure Q with numeraire B to the forward swap measure $Q^{\alpha,\beta}$, as in (D.7). Since $S_{\alpha,\beta}$ is log-normal distributed under the swap measure $Q^{\alpha,\beta}$, precisely

$$S_{\alpha,\beta}(T) = S_{\alpha,\beta}(t)e^{\int_t^T \sigma_{\alpha,\beta}(s)dW^{\alpha,\beta}(s)-\frac{1}{2}\int_t^T \sigma_{\alpha,\beta}(s)^2 ds}\,,\quad 0\le t\le T\le T_\alpha\,,$$

where $\sigma_{\alpha,\beta}$ is assumed to be deterministic, we can rewrite it consistently with the assumptions of Corollary A.0.4:

$$S_{\alpha,\beta}(T) = S_{\alpha,\beta}(t)e^{Y_{\alpha,\beta}(t,T)}\,,$$

where

$$Y_{\alpha,\beta}(t,T) := \int_t^T \sigma_{\alpha,\beta}(s)dW^{\alpha,\beta}(s) - \frac{1}{2}\int_t^T \sigma_{\alpha,\beta}(s)^2 ds\,.$$

Hence

$$Y_{\alpha,\beta}(t,T) \sim \mathcal{N}(m_{\alpha,\beta}(t,T), \Sigma^2_{\alpha,\beta}(t,T))\,, \quad \text{where}$$

$$m_{\alpha,\beta}(t,T) = -\frac{1}{2}\int_t^T \sigma_{\alpha,\beta}(s)^2 ds\,, \quad \Sigma^2_{\alpha,\beta}(t,T) = \int_t^T \sigma_{\alpha,\beta}(s)^2 ds$$

and $S_{\alpha,\beta}(t) \in \mathbb{R}$ at time t.

Thus, considering the actual price of the swaption, i.e. $t = 0$, the proof follows directly from (A.5). □

3.2 Theoretical incompatibility between LMM and SMM

At this point, a crucial question rises: Are the two main market models, the LMM and the SMM, theoretically consistent? That is, can the assumptions of log-normality of both LIBOR forward rates and forward swap rates coexist? In order to give an answer we can proceed as follows:

1. assume the hypothesis of the LMM, namely that each forward rate F_i is log-normal under its related forward measure Q^i, $i = 1, \ldots, M$, as in (2.1);

2. apply the change of measure to obtain their dynamics under the swap measure $Q^{\alpha,\beta}$, for a choice of $(\alpha, \beta) \in \mathcal{T}^{pairs}$;

3. apply the Itô's formula to obtain the resulting dynamics for the swap rate $S_{\alpha,\beta}$ under $Q^{\alpha,\beta}$;

4. check if this distribution is log-normal, as it is under the hypothesis of the SMM.

Unfortunately, the answer is negative. However, from a practical point of view, this incompatibility seems to weaken considerably. Indeed, simulating a large number of realizations of $S_{\alpha,\beta}(T_\alpha)$ with the dynamics induced by the LMM one can compute its empirical (numerical) density and compare it with the log-normal density. Consequently, it has been argued (Brace-Dun-Burton 1998 and Morini 2001-2006) that, in normal market conditions, the two distributions are hardly distinguishable.

Once ascertained the mathematical inconsistency of these two models, we must admit that the SMM is particularly convenient when pricing a swaption, because it yields the practice Black's formula for swaptions. However, for different products, even those involving the swap rate, there is no analytical formula in general. The problem left is choosing either of the two models for the whole market. After that choice, the half market consistent with the model is calibrated almost automatically, thanks to Black's formulae, but the remaining half is more intricate to calibrate.

Since the LIBOR forward rates, rather than swap rates, are more natural and representative coordinates of the yield curve usually considered, besides being mathematically more manageable, the better choice of modeling may be to assume as framework the LIBOR market model. Thus, hereafter, we are working under the hypothesis of the LMM.

Chapter 4

Pricing of Swaptions in the LMM

The LMM, unfortunately, does not feature a known distribution for the joint dynamics of forward rates, hence to evaluate swaptions, as well as other payoffs involving that joint dynamics, we have to resort to Monte Carlo simulation, under a chosen numeraire among the $T_1, \ldots, T_M$-zero coupon bonds, or to some analytical approximation.

4.1 Monte Carlo Pricing of Swaptions

The Monte Carlo method is a numerical and probabilistic method which consists in a computational algorithm relying on repeated independent random sampling to compute approximations of theoretical results, especially when it is infeasible to compute an exact result with a deterministic algorithm.

In general, Monte Carlo intends to estimate an expectation value through an arithmetic mean of realizations of i.i.d. random variables and it proceed as follows: let X be the r.v., with known distribution, on which the expectation we need to estimate depends; a pseudorandom number generator provides a sequence of realizations $X^{(k)}$ of theoretical independent r.v. $X_1, X_2, \ldots$

distributed as X; then, the desired expectation is approximated by

$$E\left[\varphi(X)\right] \cong \frac{1}{m}\sum_{k=1}^{m}\varphi(X^{(k)})\,.$$

Indeed, by the "Law of large numbers", the sample average converges to the expected value, under the hypothesis that $X_1, X_2, \ldots$ is an infinite sequence of i.i.d. random variables with finite expected value.

The most general way to price a swaption, as well as any other option with underlying forward rates, is through the Monte Carlo simulation. In order to simulate all the processes involved in the payoff, their joint dynamics is discretized with a numerical scheme for SDEs, e.g. the Euler scheme.

Recall the price of a $T_\alpha \times (T_\beta - T_\alpha)$ payer swaption:

$$\begin{aligned} &E^Q\left[D(0,T_\alpha)\,(S_{\alpha,\beta}(T_\alpha)-K)^+\sum_{i=\alpha+1}^{\beta}\tau_i\,p(T_\alpha,T_i)\right] = \\ &= p(0,T_\alpha)E^\alpha\left[(S_{\alpha,\beta}(T_\alpha)-K)^+\sum_{i=\alpha+1}^{\beta}\tau_i\,p(T_\alpha,T_i)\right]\,, \end{aligned}$$

by considering this time the T_α-bond $p(\cdot,T_\alpha)$ as numeraire.
Then, by keeping in mind that $S_{\alpha,\beta}$ has an expression in terms of the relevant spanning forward rates, given by (1.8), notice that the expectation above depends on the joint distribution of the same F's.

The dynamics of the k-th forward rate, for each $k=\alpha+1,\ldots,\beta$, under Q^α is

$$dF_k(t) = \sigma_k(t)F_k(t)\sum_{j=\alpha+1}^{k}\frac{\rho_{k,j}\tau_j\sigma_j(t)F_j(t)}{1+\tau_jF_j(t)}dt + \sigma_k(t)F_k(t)dZ_k^\alpha(t)\,, \quad t<T_\alpha\,, \tag{4.1}$$

and, in order to evaluate the payoff

$$(S_{\alpha,\beta}(T_\alpha)-K)^+\sum_{i=\alpha+1}^{\beta}\tau_i\,p(T_\alpha,T_i)\,, \tag{4.2}$$

we have to generate a number of realization, say m, of

$$F_{\alpha+1}(T_\alpha),\ldots,F_\beta(T_\alpha)\,,$$

according to the dynamics (4.1). Finally the Monte Carlo price of the swaption is given by the mean of the m evaluations of the payoff (4.2).

To simulate the dynamics in the SDE (4.1), which has neither analytical solution nor known distribution, we discretize it by using the Euler scheme applied to the natural logarithm-version of the same equation. The choice to discretize the ln-version of (4.1) is based on the advantage of having both a deterministic diffusion coefficients and a better numerical stability. We have, by the Itô's formula, the ln-dynamics

$$d\ln F_k(t) = \left(\sigma_k(t)\sum_{j=\alpha+1}^{k}\frac{\rho_{k,j}\tau_j\sigma_j(t)F_j(t)}{1+\tau_jF_j(t)} - \frac{\sigma_k(t)^2}{2}\right)dt + \sigma_k(t)dZ_k^\alpha(t)\,. \quad (4.3)$$

We introduce a time grid with a sufficiently (but not too) small step $\Delta t = \frac{T_\alpha}{N}$ and consider the discrete scheme

$$\begin{aligned}\ln F_k(t_{i+1}) = \;& \ln F_k(t_i)\left(\sigma_k(t_i) + \sum_{j=\alpha+1}^{k}\frac{\rho_{k,j}\tau_j\sigma_j(t_i)F_j(t_i)}{1+\tau_jF_j(t_i)} - \frac{\sigma_k(t_i)^2}{2}\right)\Delta t + \\ & +\sigma_k(t_i)(Z_k^\alpha(t_{i+1}) - Z_k^\alpha(t_i))\,,\end{aligned} \quad (4.4)$$

with $t_i = i\Delta t,\ i = 0,\dots,N-1$. This provides us with m approximated realizations $F_k^{(1)}(T_\alpha),\dots,F_k^{(m)}(T_\alpha)$ of the true process $F_k(T_\alpha)$, such that

$$\exists\delta_0 > 0 : \quad E^\alpha\left[|F_k^{(i)}(T_\alpha) - F_k(T_\alpha)|\right] \le c(T_\alpha)\,\Delta t \text{ for } \Delta t \le \delta_0\,,$$

where $c(T_\alpha)$ is a positive constant. Hence the convergence is of order 1.

Remark 4. We may consider a more refined scheme coming from (4.4) by the following substitution, in the vector version:

$$\Sigma(t_i)(Z^\alpha(t_{i+1}) - Z^\alpha(t_i)) \longmapsto \Delta\zeta(t_i)\,,$$

where

$$\Sigma(t) := \begin{pmatrix} \sigma_{\alpha+1} & 0 & \cdots & 0 \\ 0 & \sigma_{\alpha+2} & 0\cdots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_\beta \end{pmatrix}, \quad Z^\alpha = \begin{pmatrix} Z^\alpha_{\alpha+1} \\ Z^\alpha_{\alpha+2} \\ \vdots \\ Z^\alpha_\beta \end{pmatrix}.$$

$$\Delta\zeta(t) := \int_t^{t+\Delta t} \Sigma(s) dZ^\alpha(s) \sim \mathcal{N}(0, Cov(t))\,,$$

with the covariance $n \times n$ matrix, $n := \beta - \alpha$, having the elements

$$Cov_{i,j}(t) := \int_t^{t+\Delta t} (\Sigma\rho\Sigma')_{i,j}\, ds\,.$$

Indeed, integrating the ln-dynamics (4.3) in the vector version between t and $t + \Delta t$, the resulting stochastic integral in it is just $\Delta\zeta(t_i)$. By means of this substitution, we can simulate more accurate random shocks with gaussian distribution

$$\mathcal{N}(0, Cov(t))$$

instead of

$$\mathcal{N}(0, \Delta t \Sigma\rho\Sigma')\,.$$

Monte Carlo Variance Reduction

Before introducing the variance reduction technique for Monte Carlo simulation, we need to give some general notions and results.

Consider a general payoff at time T, $\Pi(T)$, depending on a vector of spanning forward LIBOR rates $F(t)$, for $t \in [0, T]$, where typically T is smaller than or equal to the expiry of the first forward rate. For instance (4.2) is a particular case of $\Pi(T) = \Pi(T_\alpha)$. We simulate various scenarios of $\Pi(T)$ through a scheme as (4.4) under the T-forward measure. Let m be the number of simulated paths, the Monte Carlo price of the payoff is

$$E^Q\left[D(0,T)\Pi(T)\right] = p(0,T)E^T\left[\Pi(T)\right] \;\approx\; p(0,T)\,\frac{\sum_{j=1}^{m} \Pi^{(j)}(T)}{m}\,.$$

Since $\Pi^{(1)}(T), \Pi^{(2)}(T), \ldots$ constitute a sequence of realizations of independent identically distributed random variables distributed as $\Pi(T)$, under the hypothesis that the r.v. are summable, i.e. $\Pi(T) \in \mathrm{L}^1(\Omega)$, we can apply the Central Limit Theorem to have the convergence

$$\frac{\sum_{j=1}^{m} (\Pi^{(j)}(T) - E^T[\Pi(T)]}{\sqrt{m} Std(\Pi(T))} \stackrel{\text{in law}}{\longrightarrow} \mathcal{N}(0,1)$$

for $m \to +\infty$. Thus, for large m, the following approximation yields:

$$\frac{\sum_{j=1}^{m} \Pi^{(j)}(T)}{m} - E^T[\Pi(T)] \overset{Q^T}{\sim} \frac{Std(\Pi(T))}{\sqrt{m}} Z\,, \quad Z \overset{Q^T}{\sim} \mathcal{N}(0,1)\,.$$

It follows that the probability that the Monte Carlo estimate is not farther than ϵ from the true price is

$$\begin{aligned} Q^T \left\{ \left| \frac{\sum_{j=1}^{m} \Pi^{(j)}(T)}{m} - E^T[\Pi(T)] \right| < \epsilon \right\} &= Q^T \left\{ |Z| < \epsilon \frac{\sqrt{m}}{Std(\Pi(T))} \right\} \\ &= 2\,\Phi\left(\epsilon \frac{\sqrt{m}}{Std(\Pi(T))}\right) - 1\,, \end{aligned}$$

where as usual Φ denotes the c.d.f. of the standard gaussian distribution. Once we have chosen a desired value for such a probability, we find the corresponding value for ϵ. For a typical choice of accuracy of 98%, we solve in ϵ the equation

$$2\,\Phi\left(\epsilon \frac{\sqrt{m}}{Std(\Pi(T))}\right) - 1 = 0.98$$

by

$$2\,\Phi(z) - 1 = 0.98 \;\Leftrightarrow\; \Phi(z) = 0.99 \;\Leftrightarrow\; z \approx 2.33 \;\Leftrightarrow\; \epsilon \approx 2.33 \frac{Std(\Pi(T))}{\sqrt{m}}\,.$$

The resulting confidence interval at level 98% for the true value $E[\Pi(T)]$ is

$$\left[\frac{\sum_{j=1}^{m} \Pi^{(j)}(T)}{m} - 2.33 \frac{Std(\Pi(T))}{\sqrt{m}}\,,\; \frac{\sum_{j=1}^{m} \Pi^{(j)}(T)}{m} + 2.33 \frac{Std(\Pi(T))}{\sqrt{m}} \right].$$

As m increases, the window shrinks as $\frac{1}{\sqrt{m}}$.

Moreover, the standard deviation of the payoff is usually unknown, thus it is typically replaced by the sample standard deviation, with square

$$\widehat{Std}(\Pi; m)^2 := \frac{\sum_{j=1}^{m} (\Pi^{(j)}(T))^2}{m} - \left(\frac{\sum_{j=1}^{m} \Pi^{(j)}(T)}{m} \right)^2$$

and the actual Monte Carlo window is

$$\left[\frac{\sum_{j=1}^{m}\Pi^{(j)}(T)}{m} - 2.33\frac{\widehat{Std}(\Pi;m)}{\sqrt{m}}\,,\ \frac{\sum_{j=1}^{m}\Pi^{(j)}(T)}{m} + 2.33\frac{\widehat{Std}(\Pi;m)}{\sqrt{m}}\right]. \tag{4.5}$$

In some cases, to have a small enough window, we need to simulate a huge number of scenarios, being thus too time-consuming. A way to tackle this problem is given by the *control variate technique*, which allows to reduce the sample variance, so as to narrow the window in (4.5), without increasing m. Omitting for simplicity the time T in the notations, the method is to proceed as follows:

I. Consider another payoff Π_{an} which we can evaluate analytically, whose expectation is denoted by

$$E[\Pi_{an}] = \pi_{an}\,,$$

and simulate it together with Π under the same scenarios for F.

II. Define an unbiased control-variate estimator $\widehat{\Pi}_c(\gamma;m)$ for $E[\Pi]$ as

$$\widehat{\Pi}_c(\gamma;m) := \frac{\sum_{j=1}^{m}\Pi^{(j)}}{m} + \gamma\left(\frac{\sum_{j=1}^{m}\Pi_{an}^{(j)}}{m} - \pi_{an}\right),$$

which is also the sample mean of the r.v.

$$\Pi_c(\gamma) := \Pi + \gamma(\Pi_{an} - \pi_{an})\,.$$

Hence $\Pi_c(\gamma)$ has expectation $E[\Pi]$ and variance

$$Var(\Pi_c(\gamma)) = Var(\Pi) + \gamma^2 Var(\Pi_{an}) + 2\gamma Cov(\Pi,\Pi_{an})\,,$$

this last minimized by

$$\arg\min\{Var(\Pi_c(\gamma))\} =: \gamma^* = -\frac{Cov(\Pi,\Pi_{an})}{Var(\Pi_{an})} = -\frac{Corr(\Pi,\Pi_{an})Std(\Pi)}{Std(\Pi_{an})}.$$

III. The minimum variance of $\Pi_c(\gamma)$ is computed as

$$\begin{aligned} Var(\Pi_c(\gamma^*)) = \; & Var(\Pi) + Corr(\Pi,\Pi_{an})^2 \tfrac{Var(\Pi)}{Var(\Pi_{an})} Var(\Pi_{an}) + \\ & -2Corr(\Pi,\Pi_{an})^2 Var(\Pi) \\ = \; & Var(\Pi)\left(1 - Corr(\Pi,\Pi_{an})^2\right) , \end{aligned}$$

that is smaller than the variance of Π; moreover, the larger the distance between the two variances the larger (in absolute value) the correlation between the two r.v.

IV. Moving to simulated quantities, we have

$$\widehat{Std}(\Pi_c(\gamma^*);m) = \widehat{Std}(\Pi;m)\sqrt{\left(1 - \widehat{Corr}(\Pi,\Pi_{an};m)^2\right)} ,$$

where the sample correlation is

$$\widehat{Corr}(\Pi,\Pi_{an};m) := \frac{\widehat{Cov}(\Pi,\Pi_{an};m)}{\widehat{Std}(\Pi;m)\widehat{Std}(\Pi_{an};m)}$$

and the sample covariance is

$$\widehat{Cov}(\Pi,\Pi_{an};m) := \frac{\sum_{j=1}^{m} \Pi^{(j)}\Pi_{an}^{(j)}}{m} - \frac{1}{m^2}\left(\sum_{j=1}^{m} \Pi^{(j)}\right)\left(\sum_{j=1}^{m} \Pi_{an}^{(j)}\right) .$$

V. Concording with the observation at point III., the variance reduction will be relevant if Π and Π_{an} are as correlated as possible. Now the window (4.5) can be substituted by

$$\left[\widehat{\Pi}_c(\gamma;m) - 2.33\frac{\widehat{Std}(\Pi_c(\gamma^*);m)}{\sqrt{m}} \; , \; \widehat{\Pi}_c(\gamma;m) + 2.33\frac{\widehat{Std}(\Pi_c(\gamma^*);m)}{\sqrt{m}}\right] , \tag{4.6}$$

which is narrower than (4.5) by a factor $\sqrt{\left(1 - \widehat{Corr}(\Pi,\Pi_{an};m)^2\right)}$.

This technique is quite general and the choice of Π_{an} is theoretically free.

In the case of the pricing of swaptions in the LMM, we select as Π_{an} one of the simplest payoff with underlying rates $F_{\alpha+1},\ldots,F_\beta$, as may be a portfolio of FRA contracts at time T_α on each single time interval $(T_{i-1}, T_i]$, such that are all fair contracts at time 0. By recalling the price in (1.3) and reversing the two roles involved, we consider the T_α-price

$$\sum_{i=\alpha+1}^{\beta} p(T_\alpha, T_i)\tau_i(F_i(T_\alpha) - K)\,,$$

where the fair value at time 0 of K is equal to $F_i(0)$. We take such contract as our T_α-payoff and rewrite it by a change of measure under its expectation as follows:

$$\begin{aligned} & E^Q\left[D(0,T_\alpha)\sum_{i=\alpha+1}^{\beta} p(T_\alpha, T_i)\tau_i(F_i(T_\alpha) - F_i(0))\right] = \\ &= E^j\left[\frac{p(0,T_j)}{p(T_\alpha,T_j)}\sum_{i=\alpha+1}^{\beta} p(T_\alpha, T_i)\tau_i(F_i(T_\alpha) - F_i(0))\right] = \\ &= p(0,T_j)E^j\left[\frac{\sum_{i=\alpha+1}^{\beta} p(T_\alpha,T_i)\tau_i(F_i(T_\alpha)-F_i(0))}{p(T_\alpha,T_j)}\right]. \end{aligned} \tag{4.7}$$

Thus we can set

$$\Pi_{an}(T_\alpha) = \frac{\sum_{i=\alpha+1}^{\beta} p(T_\alpha, T_i)\tau_i(F_i(T_\alpha) - F_i(0))}{p(T_\alpha, T_j)}\,,$$

whose price at time 0 is

$$\pi_{an} = 0\,.$$

Indeed, the payoff $\Pi_{an}(\cdot)$ is a sum of traded assets divided by $p(\cdot, T_j)$, hence it is a martingale under the T_j- forward measure Q^j, which implies that

$$E^j\left[\Pi_{an}(T_\alpha)\right] = E^j\left[\Pi_{an}(0)\right] = E^j\left[0\right] = 0\,.$$

4.2 Approximated Analytical Swaption Prices

4.2.1 Rank-One analytical Swaption prices

This approximation is due to Brace (1996).

By assuming the tenor structure $\mathcal{T} = \{T_\alpha, \dots, T_\beta\}$ and the year fractions $\tau = \{\tau_\alpha, \dots, \tau_\beta\}$, the swaption price can be also written as follows:

$$\begin{aligned} PS(0, T_\alpha, \mathcal{T}, K) &= E\left[D(0,T_\alpha)\left(\sum_{i=\alpha+1}^{\beta} p(T_\alpha, T_i)\,\tau_i\,(F_i(T_\alpha) - K)\right)^+\right] \\ &= p(0,T_\alpha)\,E^\alpha\left[\left(\sum_{i=\alpha+1}^{\beta} p(T_\alpha, T_i)\,\tau_i\,(F_i(T_\alpha) - K)\right)^+\right] \\ &= \sum_{i=\alpha+1}^{\beta} p(0,T_i)\,\tau_i\,E^i\left[(F_i(T_\alpha) - K)\,\mathbb{1}_A\right], \end{aligned}$$

where

$$A := \left\{\left(\sum_{i=\alpha+1}^{\beta} p(T_\alpha, T_i)\,\tau_i\,(F_i(T_\alpha) - K)\right) > 0\right\} = \{S_{\alpha,\beta}(T_\alpha) > K\}\,.$$

The problem is to find approximated analytical formulae for the amount

$$E^i\left[(F_i(T_\alpha) - K)\,\mathbb{1}_A\right]. \tag{4.8}$$

Steps

- Choose a forward measure Q^γ, with $\gamma \in \{\alpha, \dots, \beta\}$, under which consider the LMM forward rate dynamics given by Proposition 2.0.2.
- As first approximation replace the stochastic percentage drift with a deterministic one:

$$\begin{aligned} k < \gamma:&\quad -\sum_{j=k+1}^{\gamma} \tfrac{\rho_{k,j}\tau_j\sigma_j(t)F_j(t)}{1+\tau_j F_j(t)} \approx -\sum_{j=k+1}^{\gamma} \tfrac{\rho_{k,j}\tau_j\sigma_j(t)F_j(0)}{1+\tau_j F_j(0)} =: \mu_{\gamma,k}(t) \\ k = \gamma:&\quad 0 =: \mu_{\gamma,\gamma}(t) \\ k > \gamma:&\quad \sum_{j=\gamma+1}^{k} \tfrac{\rho_{k,j}\tau_j\sigma_j(t)F_j(t)}{1+\tau_j F_j(t)} \approx \sum_{j=\gamma+1}^{k} \tfrac{\rho_{k,j}\tau_j\sigma_j(t)F_j(0)}{1+\tau_j F_j(0)} =: \mu_{\gamma,k}(t) \end{aligned}$$

  so that each dynamics follows now a geometric Brownian motion:

$$dF_k(t) = \sigma_k(t)\mu_{\gamma,k}(t)F_k(t)dt + \sigma_k(t)F_k(t)dZ_k^\gamma(t)\,,\quad k = \alpha+1, \dots, \beta\,.$$

Thus:

$$F_k(T_\alpha) = F(0)\exp\left(\int_0^{T_\alpha}\left(\sigma_k(t)\mu_{\gamma,k}(t) - \frac{1}{2}\sigma_k(t)^2\right)dt + \int_0^{T_\alpha}\sigma_k(t)dZ_k^\gamma(t)\right),$$

or equivalently, in the vector form:

$$F(T_\alpha) = F(0)\exp\left(\int_0^{T_\alpha}\sigma(t)\mu_\gamma(t)dt - \frac{1}{2}\int_0^{T_\alpha}\sigma(t)^2dt\right)\exp\left(X^\gamma\right), \tag{4.9}$$

where

$$\sigma(t) := \begin{bmatrix}\sigma_{\alpha+1}(t)\\ \vdots \\ \sigma_\beta(t)\end{bmatrix}, \quad \mu_\gamma(t) = \begin{bmatrix}\mu_{\gamma,\alpha+1}(t)\\ \vdots \\ \mu_{\gamma,\beta}(t)\end{bmatrix}, \quad X^\gamma := \int_0^{T_\alpha}\sigma(t)dZ^\gamma(t)$$

and all the products act componentwise.
Notice that

$$X^\gamma \overset{Q^\gamma}{\sim} \mathcal{N}(0,V), \quad V_{i,j} := \int_0^{T_\alpha}\sigma_i(t)\sigma_j(t)\rho_{i,j}dt\,.$$

Remark 5.

$$X^\gamma = X^\alpha - \int_0^{T_\alpha}\sigma(t)(\mu_\gamma(t) - \mu_\alpha(t))dt\,.$$

Proof. Choosing as forward measure Q^α, we analogously obtain the forward rate dynamics in (4.9) as long as you replace $\mu_\gamma(t)$ with $\mu_\alpha(t)$ and $dZ^\gamma(t)$ with $dZ^\alpha(t)$. Then we equal this with (4.9) and we get X^γ in terms of X^α. □

- Since $\rho_{i,j} > 0\ \forall i,j$, we have $V_{i,j} > 0\ \forall i,j$ and V is irreducible.
  Thus, by the Perron-Frobenius Theorem, V admits a unique dominant eigenvalue $\lambda_1(V) > 0$ whose associated eigenvector is $e_1(V) > 0$.
  Approximate V with a rank-one matrix:

  $$V \approx V^1 := \Gamma\Gamma', \quad \text{where} \quad \Gamma := \sqrt{\lambda_1(V)}\,e_1(V)\,.$$

  The previous X^γ is then substituted by the one with components

  $$X_i^\gamma = \Gamma_i U^\gamma, \quad \text{where} \quad U^\gamma \overset{Q^\gamma}{\sim} \mathcal{N}(0,1),$$

and

$$\begin{aligned} k<\gamma: &\quad \int_0^{T_\alpha} \sigma_k(t)\mu_{\gamma,k}(t)dt = -\Gamma_k \sum_{j=k+1}^{\gamma} \tfrac{\tau_j\Gamma_j(t)F_j(0)}{1+\tau_jF_j(0)} =: \Gamma_k q_{\gamma,k}\ , \\ k=\gamma: &\quad \int_0^{T_\alpha} \sigma_k(t)\mu_{\gamma,\gamma}(t)dt = -\Gamma_\gamma 0 =: \Gamma_\gamma q_{\gamma,\gamma}\ , \\ k>\gamma: &\quad \int_0^{T_\alpha} \sigma_k(t)\mu_{\gamma,k}(t)dt = \Gamma_k \sum_{j=\gamma+1}^{k} \tfrac{\tau_j\Gamma_jF_j(0)}{1+\tau_jF_j(0)} =: \Gamma_k q_{\gamma,k}\ . \end{aligned} \tag{4.10}$$

Proof of (4.10).

$$\begin{aligned} k<\gamma: \quad \int_0^{T_\alpha} \sigma_k(t)\mu_{\gamma,k}(t)dt &= -\int_0^{T_\alpha} \sigma_k(t) \sum_{j=k+1}^{\gamma} \tfrac{\rho_{k,j}\tau_j\sigma_j(t)F_j(0)}{1+\tau_jF_j(0)}dt \\ &= -\sum_{j=k+1}^{\gamma} \tfrac{\tau_jF_j(0)}{1+\tau_jF_j(0)} \int_0^{T_\alpha} \sigma_k(t)\sigma_j(t)\rho_{k,j}dt\,, \end{aligned}$$

where

$$\int_0^{T_\alpha} \sigma_k(t)\sigma_j(t)\rho_{k,j}dt \approx \Gamma_k\Gamma_j\,.$$

Analogously the other cases. □

The forward-rate dynamics becomes

$$F(T_\alpha) = F(0)\exp\left(\Gamma q_{\gamma,\cdot} - \frac{1}{2}Diag(\Gamma\Gamma')dt\right)\exp(X^\gamma)\,,$$

where the product $\Gamma q_{\gamma,\cdot}$ acts componentwise.

- Set

$$p := q_{\alpha,\cdot} \quad \text{and} \quad U := U^\alpha\,,$$

then express

$$X^\gamma = X^\alpha - \Gamma(q_{\gamma,\cdot} - q_{\alpha,\cdot}) = \Gamma(U + p - q_{\gamma,\cdot})$$

and notice that

$$q_{\gamma,k} = p_k - p_\gamma\,.$$

Indeed, in the case $k<\gamma$:

$$q_{\gamma,k} = -\sum_{j=k+1}^{\gamma} \frac{\tau_j\Gamma_j(t)F_j(0)}{1+\tau_jF_j(0)} = \sum_{j=\alpha+1}^{k} \frac{\tau_j\Gamma_j(t)F_j(0)}{1+\tau_jF_j(0)} - \sum_{j=\alpha+1}^{\gamma} \frac{\tau_j\Gamma_j(t)F_j(0)}{1+\tau_jF_j(0)}$$

and analogously in the other cases.

Thanks to the fact that $q_{\gamma,\gamma} = 0$, we obtain the following expression for the γ^{th} forward-rate dynamics:

$$\begin{aligned} F_\gamma(T_\alpha; U^\gamma) := \ & F_\gamma(0) \exp\left(-\tfrac{1}{2}\Gamma_\gamma^2\right) \exp(\Gamma_\gamma U^\gamma) = \\ & = F_\gamma(0) \exp\left(-\tfrac{1}{2}\Gamma_\gamma^2\right) \exp(\Gamma_\gamma (U + p_\gamma)) \quad =: F_\gamma(T_\alpha; U)\,, \end{aligned}$$

from which: $\quad U^\gamma = U + p_\gamma\,, \quad \gamma = \alpha, \ldots, \beta\,.$

- Denote

$$p(T_\alpha, T_i) = \prod_{k=\alpha+1}^{i} \frac{p(T_\alpha, T_k)}{p(T_\alpha, T_{k-1})} = \prod_{k=\alpha+1}^{i} \frac{1}{1 + \tau_k F_k(T_\alpha; U)} =: p(T_\alpha, T_i; U)$$

and rewrite

$$\begin{aligned} A \ &= \ \left\{ \left(\sum_{i=\alpha+1}^{\beta} p(T_\alpha, T_i; U)\, \tau_i \left(F_i(T_\alpha; U) - K\right) \right) > 0 \right\} \qquad (4.11) \\ &= \ \left\{ \left(\sum_{i=\alpha+1}^{\beta} p(T_\alpha, T_i; U) \left(\frac{p(T_\alpha, T_{i-1}; U)}{p(T_\alpha, T_i; U)} - 1 - \tau_i K \right) \right) > 0 \right\} \\ &= \ \left\{ \left(1 - p(T_\alpha, T_\beta; U) - \sum_{i=\alpha+1}^{\beta} \tau_i K\, p(T_\alpha, T_i; U) \right) > 0 \right\}. \end{aligned}$$

Remark 6. The equation

$$\sum_{i=\alpha+1}^{\beta} \prod_{k=\alpha+1}^{i} \frac{1}{1 + \tau_k F_k(T_\alpha; U)}\, \tau_i \left(F_i(T_\alpha; U) - K\right) = 0 \qquad (4.12)$$

has a unique solution $U = U_*$.

Proof. We aim to prove the monotony of the function in the left-hand side of (4.12) and we do that by proving the positivity of its derivative, i.e.

$$\begin{aligned} & \frac{\partial}{\partial U} \left(\sum_{i=\alpha+1}^{\beta} \prod_{k=\alpha+1}^{i} \frac{1}{1+\tau_k F_k(T_\alpha;U)}\, \tau_i \left(F_i(T_\alpha; U) - K\right) \right) = \\ {\scriptstyle from\ (4.11)}= \ & \frac{\partial}{\partial U} \left(1 - p(T_\alpha, T_\beta; U) - \sum_{i=\alpha+1}^{\beta} \tau_i K\, p(T_\alpha, T_i; U) \right) > 0 \end{aligned}$$

A sufficient condition for this is

$$\begin{array}{rll} & \frac{\partial}{\partial U} p(T_\alpha, T_i; U) < 0 \quad \forall i \in \{\alpha+1, \dots, \beta\} & \Leftrightarrow \\ \Leftrightarrow & \frac{\partial}{\partial U} \ln p(T_\alpha, T_i; U) < 0 \quad \forall i \in \{\alpha+1, \dots, \beta\} & \Leftrightarrow \\ \Leftrightarrow & \frac{\partial}{\partial U} \sum_{k=\alpha+1}^{i} \ln(1+\tau_k F_k(T_\alpha; U)) > 0 \quad \forall i \in \{\alpha+1, \dots, \beta\} & . \end{array}$$

A sufficient condition for this is

$$\begin{array}{rl} & \frac{\partial}{\partial U} \ln(1+\tau_k F_k(T_\alpha; U)) > 0 \quad \forall k \in \{\alpha+1, \dots, \beta\} \quad \Leftrightarrow \\ \Leftrightarrow & \frac{\partial}{\partial U} F_k(T_\alpha; U) > 0 \quad \forall k \in \{\alpha+1, \dots, \beta\}\,. \end{array}$$

This is satisfied because

$$\frac{\partial}{\partial U} F_k(T_\alpha; U) = \Gamma_k F_k(T_\alpha; U) > 0\,, \text{ since } \Gamma > 0\,.$$

Thus the left-hand side of equation (4.12) is an increasing function of U, hence the proof. □

From the last remark, the inequality in (4.11) is equivalent to

$$U > U_*\,, \quad \text{as well as} \quad U^i > U_* + p_i\,.$$

Compute the expectation (4.8) as

$$\begin{aligned} & E^i\left[(F_i(T_\alpha) - K)\,\mathbb{1}_A\right] = \\ = \; & E^i\left[(F_i(T_\alpha) - K)\,\mathbb{1}_{\{U^i > U_* + p_i\}}\right] \\ = \; & E^i\left[\left(F_i(0)\exp\left(-\frac{1}{2}\Gamma_i^2 + \Gamma_i U^i\right) - K\right)\mathbb{1}_{\{U^i > U_* + p_i\}}\right] \\ = \; & E^i\left[\left(F_i(0)\exp\left(-\frac{1}{2}\Gamma_i^2 + \Gamma_i U^i\right) - K\right)\mathbb{1}_{\{F_i(0)\exp\left(-\frac{1}{2}\Gamma_i^2 + \Gamma_i U^i\right) > F_i^*\}}\right] \\ = \; & E^i\left[\left(F_i(0)\exp\left(-\frac{1}{2}\Gamma_i^2 + \Gamma_i U^i\right) - F_i^* + F_i^* + \right.\right. \\ & \left.\left. -K\right)\mathbb{1}_{\{F_i(0)\exp\left(-\frac{1}{2}\Gamma_\gamma^2 + \Gamma_i U^i\right) > F_i^*\}}\right] \\ = \; & E^i\left[F_i(0)\exp\left(-\frac{1}{2}\Gamma_i^2 + \Gamma_i U^i\right) - F_i^*\right]^+ + \qquad (4.13) \\ & \quad + (F_i^* - K)\,E^i\left[\mathbb{1}_{\{F_i(0)\exp\left(-\frac{1}{2}\Gamma_i^2 + \Gamma_i U^i\right) > F_i^*\}}\right] \qquad (4.14) \\ = \; & F_i(0)\,\Phi(d_1(F_i^*, F_i(0), \Gamma_i)) - K\,\Phi(d_2(F_i^*, F_i(0), \Gamma_i))\,, \end{aligned}$$

where

$$F_i^* := F_i(0)\exp\left(-\frac{1}{2}\Gamma_i^2 + \Gamma_i U_* + p_i\right)$$

and the last formula is obtained from the previous as:

$$\begin{aligned}(4.13) &= \int_D \left(F_i(0)\exp\left(-\frac{1}{2}\Gamma_i^2 + \Gamma_i x\right) - F_i^*\right)\frac{\exp\left(-\frac{x^2}{2}\right)}{\sqrt{2\pi}}dx \\ &= \int_D \left(F_i(0)\exp\left(-\frac{(x-\Gamma_i)^2}{2}\right) - F_i^*\exp\left(-\frac{x^2}{2}\right)\right)\frac{1}{\sqrt{2\pi}}dx\,,\end{aligned}$$

$$\begin{aligned}\text{with } D :=& \ \{F_i(0)\exp\left(-\tfrac{1}{2}\Gamma_i^2 + \Gamma_i x\right) > F_i^*\} \\ =& \ \{\ln F_i(0) - \tfrac{1}{2}\Gamma_i^2 + \Gamma_i x > \ln F_i^*\} \\ =& \ \{x > \tfrac{\ln\frac{F_i^*}{F_i(0)} + \frac{1}{2}\Gamma_i^2}{\Gamma_i}\} = \{z > \tfrac{\ln\frac{F_i^*}{F_i(0)} - \frac{1}{2}\Gamma_i^2}{\Gamma_i}\}\end{aligned}$$

$$\Rightarrow \quad (4.13) = F_i(0)\,\Phi\left(\frac{\ln\frac{F_i(0)}{F_i^*} + \frac{1}{2}\Gamma_i^2}{\Gamma_i}\right) - F_i^*\,\Phi\left(\frac{\ln\frac{F_i(0)}{F_i^*} - \frac{1}{2}\Gamma_i^2}{\Gamma_i}\right);$$

$$\begin{aligned}(4.14) &= (F_i^* - K)\int_D \exp\left(-\frac{x^2}{2}\right)\frac{1}{\sqrt{2\pi}}dx \\ &= (F_i^* - K)\Phi\left(\frac{\ln\frac{F_i(0)}{F_i^*} - \frac{1}{2}\Gamma_i^2}{\Gamma_i}\right).\end{aligned}$$

Eventually we have an analytical formula that requires only a root-searching procedure to find U_* .

Proposition 4.2.1 (Brace's rank-one formula)**.** *The approximated price of a $T_\alpha \times (T_\beta - T_\alpha)$ payer swaption is given by*

$$\sum_{i=\alpha+1}^{\beta} \tau_i\, p(0,T_i)\ [F_i(0)\,\Phi(\Gamma_i - U_* - p_i) - K\,\Phi(-U_* - p_i)]\ . \tag{4.15}$$

Proof.

$$\begin{aligned}d_1(F_i^*, F_i(0), \Gamma_i) &= \frac{\ln\frac{F_i(0)}{F_i^*} + \frac{1}{2}\Gamma_i^2}{\Gamma_i} = \frac{-\ln\left(\exp\left(-\frac{1}{2}\Gamma_i^2 + \Gamma_i(U_* + p_i)\right)\right) + \frac{1}{2}\Gamma_i^2}{\Gamma_i} \\ &= \frac{\frac{1}{2}\Gamma_i^2 - \Gamma_i(U_* + p_i) + \frac{1}{2}\Gamma_i^2}{\Gamma_i} = \Gamma_i - U_* - p_i\,.\end{aligned}$$

Analogously

$$d_2(F_i^*, F_i(0), \Gamma_i) \;=\; U_* - p_i\,.$$

□

4.2.2 Rank-r analytical Swaption prices

The approach is the same as in the rank-one case and the first two steps still hold in the extended rank-r case.
Then define a rank-r approximation of V as follows: $V \approx V^r := \Gamma\Gamma'$,

$$\text{where} \quad \Gamma := \left[\sqrt{\lambda_1(V)}e_1(V) \quad \sqrt{\lambda_2(V)}e_2(V) \quad \cdots \quad \sqrt{\lambda_r(V)}e_r(V)\right].$$

Thus

$$X^\gamma \overset{Q^\gamma}{\sim} \mathcal{N}(0, V^r)$$

with components given by

$$X_i^\gamma = \Gamma_{(i)}U^\gamma\,, \quad \text{where} \quad U^\gamma \overset{Q^\gamma}{\sim} \mathcal{N}(0, I_r)$$

and $\Gamma_{(i)}$ is the i^{th} row of Γ.
Typically it suffices $r << \beta - \alpha$.
Formulae and definitions in (4.10) still hold by replacing $\Gamma_k\Gamma_j$ with $\Gamma_{(k)}{\Gamma_{(j)}}'$.
Under the same adjustments, the following formulae still hold up to the observation (4.12), but here U, the $q_{\gamma,k}$'s and the p_γ's are r-vectors.

Now the equation (4.12) was found (Brace,1997) to describe with a fair approximation a hyperplane:

$$U_1 = s_1 + \sum_{k=2}^{r} s_k U_k\,.$$

Remark 7. The set A can be rewritten as follows:

$$A = \left\{ U \in \mathbb{R}^r \;:\; U_1 > s_1 + \sum_{k=2}^{r} s_k U_k \right\}. \tag{4.16}$$

Indeed, exactly as in the rank-one case, the partial derivative with respect to U_1 of the left-hand side of the equation corresponding to (4.12) is shown to be positive.

Remark 8. The set A can be expressed in terms of U^i as follows:

$$A = \left\{ U^i \in \mathbb{R}^r \,:\, wU^i > s_i^* \right\} , \tag{4.17}$$

where

$$w := \begin{bmatrix} 1 & -s_2 & \cdots & -s_r \end{bmatrix} , \quad s_i^* := s_1 + (p_i)_1 - \sum_{k=2}^{r} s_k (p_i)_k .$$

Proof of (4.17). From (4.16) and $U^i = U + p_i$, we have

$$A = \left\{ U^i \in \mathbb{R}^r \,:\, U_1^i - \sum_{k=2}^{r} s_k U_k^i > s_1 + (p_i)_1 - \sum_{k=2}^{r} s_k (p_i)_k \right\} .$$

□

Eventually, compute the expectation (4.8) as:

$$\begin{aligned}
& E^i \left[(F_i(T_\alpha) - K)\, \mathbb{1}_A \right] = \\
= \; & E^i \left[(F_i(T_\alpha) - K)\, \mathbb{1}_{\{wU^i > s_i^*\}} \right] \\
= \; & E^i \left[\left(F_i(0) \exp\left(\Gamma_{(i)} U^i - \frac{1}{2} |\Gamma_{(i)}|^2 \right) - K \right) \mathbb{1}_{\{wU^i > s_i^*\}} \right] \\
= \; & F_i(0)\, E^i \left[\exp\left(\Gamma_{(i)} U^i - \frac{1}{2} |\Gamma_{(i)}|^2 \right) \mathbb{1}_{\{wU^i > s_i^*\}} \right] + K\, E^i \left[\mathbb{1}_{\{wU^i > s_i^*\}} \right] \\
= \; & F_i(0)\, E^i \left[\mathbb{1}_{\{w\left(U^i + \Gamma_{(i)}^*\right) > s_i^*\}} \right] - K\, \Phi\left(-\frac{s_i^*}{|w|} \right) \qquad (4.18) \\
= \; & F_i(0)\, \Phi\left(-\frac{s_i^* - w\Gamma_{(i)}^*}{|w|} \right) - K\, \Phi\left(-\frac{s_i^*}{|w|} \right) . \qquad (4.19)
\end{aligned}$$

Here, (4.18) follows from the first piece of the previous sum by applying the property

$$E^i \left[\exp\left(b'U^i - \frac{1}{2} b'b \right) g(U^i) \right] = E^i \left[g(U^i + b) \right] ,$$

which holds because

$$\begin{aligned}
E^i \left[\exp\left(b'U^i - \tfrac{1}{2} b'b \right) g(U^i) \right] &= \int_{\mathbb{R}^r} e^{b'u - \frac{1}{2} b'b} g(u) e^{-\frac{1}{2} u'u} \frac{1}{\sqrt{(2\pi)^r}} du \\
&= \int_{\mathbb{R}^r} e^{-\frac{1}{2}(u-b)'(u-b)} \frac{g(u)}{\sqrt{(2\pi)^r}} du \\
\text{(c.o.v. } z = u - b) \qquad &= \int_{\mathbb{R}^r} e^{z'z} g(z+b) \frac{1}{\sqrt{(2\pi)^r}} dz \\
&= E^i \left[g(U^i + b) \right] .
\end{aligned}$$

Instead, (4.19) follows from the second summand by exploiting the fact that

$$U^i \overset{Q^i}{\sim} \mathcal{N}(0, I_r) \quad \Leftrightarrow \quad (U^i)_j \overset{Q^i}{\sim} \mathcal{N}(0,1) \;\; \forall j \in \{1,\dots,r\} \text{ and } (U^i)_{j=1,\dots,r}$$

are stochastically independent,

$$\text{thus} \qquad Z = \frac{wU^i}{|w|} = \frac{\sum_{j=1}^{r} w_j (U^i)_j}{\sqrt{\sum_{j=1}^{r} w_j^2}} \quad \Longrightarrow \quad Z \overset{Q^i}{\sim} \mathcal{N}(0,1)\,.$$

Proposition 4.2.2 (Brace's rank-r formula)**.** *The approximated price of a $T_\alpha \times (T_\beta - T_\alpha)$ payer swaption is given by*

$$\sum_{i=\alpha+1}^{\beta} \tau_i\, p(0,T_i) \left[F_i(0)\, \Phi\left(-\frac{s_i^* - w\Gamma(i)'}{|w|} \right) - K\, \Phi\left(-\frac{s_i^*}{|w|} \right) \right]. \tag{4.20}$$

This analytical formula requires only $2r - 1$ root-searching procedures to find $s_1, \dots, s_r$, that can proceed according to the following:

- Solve numerically eq. (4.12) with $U = [\alpha_1\, 0 \cdots 0]'$, then $s_1 = \alpha_1$.
- Solve eq. (4.12) with $U = [\alpha_2^- - \frac{1}{2}\, 0 \cdots 0]'$ and with $U = [\alpha_2^+\, \frac{1}{2}\, 0 \cdots 0]'$, then $s_2 = \alpha_2^+ - \alpha_2^-$.
- For $k = 3, \dots, r$ solve eq. (4.12) with $U = [\alpha_k^-\, 0 \cdots \overset{k}{-\frac{1}{2}}\; 0 \cdots 0]'$ and with $U = [\alpha_k^+\, 0 \cdots \overset{k}{\frac{1}{2}}\; 0 \cdots 0]'$, then $s_k = \alpha_k^+ - \alpha_k^-$.

4.2.3 Rebonato's approximation

In the LSM, the forward swap rate, which is always a martingal under the measure $Q^{\alpha,\beta}$, has the dynamics

$$dS_{\alpha,\beta}(t) = \sigma_{\alpha,\beta}(t) S_{\alpha,\beta}(t) dW^{\alpha,\beta}(t)\,,$$

where the volatility process $\sigma_{\alpha,\beta}$ is a deterministic function of time.
Instead, in the LMM, $\sigma_{\alpha,\beta}$ is a stochastic process, thus we can't use this to

determine the Black swaption volatility $v^{\mathrm{B}}_{\alpha,\beta}$ as

$$(v^{\mathrm{B}}_{\alpha,\beta})^2 = \frac{1}{T_\alpha}\int_0^{T_\alpha} \sigma^2_{\alpha,\beta}(t)dt\,,.$$

The goal is to find an approximated quantity $v^{\mathrm{LMM}}_{\alpha,\beta}$ of $v^{\mathrm{B}}_{\alpha,\beta}$ in the LMM.

Steps

- The forward swap rate can be obtained also by equating to zero the expression of the swap price in the form of (1.5):

$$S_{\alpha,\beta}(t) = \sum_{i=\alpha+1}^{\beta} w_i(t)F_i(t)\,, \tag{4.21}$$

  that looks like a weighted sum with weights

$$w_i(t) := \frac{\tau_i\, p(t,T_i)}{\sum_{k=\alpha+1}^{\beta} \tau_k\, p(t,T_k)} = \frac{\tau_i \prod_{j=\alpha+1}^{i} \frac{1}{1+\tau_j F_j(t)}}{\sum_{k=\alpha+1}^{\beta} \tau_k \prod_{j=\alpha+1}^{k} \frac{1}{1+\tau_j F_j(t)}}\,. \tag{4.22}$$

  However these ones are not really weights, because they depend on the stochastic F's.
  According to empirical studies (both historically and through Monte Carlo simulations), the variability of the w_i's is small compared to the one of LIBOR rates, so that it makes sense to freeze the w_i's at their initial values $w_i(0)$.

- Thus approximate

$$S_{\alpha,\beta}(t) \approx \sum_{i=\alpha+1}^{\beta} w_i(0)F_i(t)\,.$$

  Under any forward measure Q^γ, $\gamma \in \{1,\ldots,M\}$, its dynamics is

$$dS_{\alpha,\beta}(t) \approx \sum_{i=\alpha+1}^{\beta} w_i(0)dF_i(t) = (\cdots)dt + \sum_{i=\alpha+1}^{\beta} w_i(0)\sigma_i(t)F_i(t)dZ_i^\gamma(t)$$

and its quadratic variation process is

$$\begin{aligned} d < S_{\alpha,\beta}, S_{\alpha,\beta} >_t &= dS_{\alpha,\beta}(t) dS_{\alpha,\beta}(t) \\ &\approx \sum_{i,j=\alpha+1}^{\beta} w_i(0) w_j(0) \sigma_i(t) \sigma_j(t) F_i(t) F_j(t) \rho_{i,j} dt \,. \end{aligned}$$

The approximated forward swap volatility is then given by

$$\begin{aligned} \sigma^2_{\alpha,\beta}(t) &= \frac{d < \ln S_{\alpha,\beta}, \ln S_{\alpha,\beta} >_t}{dt} = \frac{1}{dt} \left(\frac{dS_{\alpha,\beta}(t)}{S_{\alpha,\beta}(t)} \right) \left(\frac{dS_{\alpha,\beta}(t)}{S_{\alpha,\beta}(t)} \right) \\ &\approx \frac{\sum_{i,j=\alpha+1}^{\beta} w_i(0) w_j(0) \sigma_i(t) \sigma_j(t) F_i(t) F_j(t) \rho_{i,j}}{S^2_{\alpha,\beta}(t)} \,. \end{aligned}$$

- A further approximation makes the quadratic variation of $\ln S_{\alpha,\beta}$ deterministic, by freezing all the remained random variables, the F's (even inside $S_{\alpha,\beta}$), at time 0:

$$\sigma^2_{\alpha,\beta}(t) \approx \frac{\sum_{i,j=\alpha+1}^{\beta} w_i(0) w_j(0) \sigma_i(t) \sigma_j(t) F_i(0) F_j(0) \rho_{i,j}}{S^2_{\alpha,\beta}(0)} \,.$$

The forward swap rate, that isn't log-normal in the LMM, can thus be approximated with a log-normal process:

$$dS_{\alpha,\beta}(t) \approx \widetilde{\sigma}_{\alpha,\beta}(t) S_{\alpha,\beta}(t) dW^{\alpha,\beta}(t) \,,$$

where

$$\widetilde{\sigma}_{\alpha,\beta}(t) := \sqrt{\frac{\sum_{i,j=\alpha+1}^{\beta} w_i(0) w_j(0) \sigma_i(t) \sigma_j(t) F_i(0) F_j(0) \rho_{i,j}}{S^2_{\alpha,\beta}(0)}} \,.$$

- Finally compute the swaption price through the Black's formula.

Proposition 4.2.3 (Rebonato's formula)**.** *The LMM Black-like swaption volatility can be approximated by* $v^{\mathrm{LMM}}_{\alpha,\beta}$, *with*

$$(v^{\mathrm{LMM}}_{\alpha,\beta})^2 := \sum_{i,j=\alpha+1}^{\beta} \frac{w_i(0) w_j(0) F_i(0) F_j(0) \rho_{i,j}}{T_\alpha \, S^2_{\alpha,\beta}(0)} \int_0^{T_\alpha} \sigma_i(t) \sigma_j(t) dt \,.$$

4.2.4 Hull and White's approximation

The above procedure can be slightly improved by computing the dynamics of the swap rate $S_{\alpha,\beta}(t)$ before freezing the w's:

$$\begin{aligned} dS_{\alpha,\beta}(t) &= \sum_{i=\alpha+1}^{\beta} (w_i(t)dF_i(t) + F_i(t)dw_i(t)) + (\cdots)dt & (4.23) \\ &= \sum_{i,h=\alpha+1}^{\beta} (w_h(t)\delta_{i,h}(t) + F_i(t)\frac{\partial w_i(t)}{\partial F_h})dF_h(t) + (\cdots)dt\,, & (4.24) \end{aligned}$$

where (4.23) is obtained by applying the Itô's formula to $S_{\alpha,\beta} = S_{\alpha,\beta}(t, F, w)$ and (4.24) applying the Itô's formula to $w_i = w_i(t, F)$.
Then, from (4.22), compute

$$\frac{\partial w_i(t)}{\partial F_h} = \frac{w_i \tau_h \left(\sum_{k=h}^{\beta} \tau_k \prod_{j=\alpha+1}^{k} \frac{1}{1+\tau_j F_j(t)} - \mathbb{1}_{\{i \geq h\}} \right)}{(1 + \tau_h F_h(t)) \sum_{k=\alpha+1}^{\beta} \tau_k \prod_{j=\alpha+1}^{k} \frac{1}{1+\tau_j F_j(t)}} .$$

Proof.

$$\frac{\partial w_i(t)}{\partial F_h} = \begin{cases} \clubsuit, & \text{if } i < h \\ \spadesuit, & \text{if } i \geq h\,, \end{cases}$$

where

$$\begin{aligned} \clubsuit &= \frac{-\tau_i \prod_{r=\alpha+1}^{i} \frac{1}{1+\tau_r F_r(t)} \cdot \sum_{k=h}^{\beta} \tau_k \prod_{\substack{j=\alpha+1 \\ j \neq h}}^{k} \frac{1}{1+\tau_j F_j(t)} \frac{-\tau_h}{(1+\tau_h F_h(t))^2}}{\left(\sum_{k=\alpha+1}^{\beta} \tau_k \prod_{j=\alpha+1}^{k} \frac{1}{1+\tau_j F_j(t)} \right)^2} \\ &= \frac{w_i \tau_h}{(1 + \tau_h F_h(t))^2 \sum_{k=\alpha+1}^{\beta} \tau_k \prod_{j=\alpha+1}^{k} \frac{1}{1+\tau_j F_j(t)}} \cdot \sum_{k=h}^{\beta} \tau_k \prod_{\substack{j=\alpha+1 \\ j \neq h}}^{k} \frac{1}{1 + \tau_j F_j(t)} \\ &= \frac{w_i \tau_h}{(1 + \tau_h F_h(t)) \sum_{k=\alpha+1}^{\beta} \tau_k \prod_{j=\alpha+1}^{k} \frac{1}{1+\tau_j F_j(t)}} \cdot \sum_{k=h}^{\beta} \tau_k \prod_{j=\alpha+1}^{k} \frac{1}{1 + \tau_j F_j(t)}\,, \end{aligned}$$

$$\spadesuit = \frac{\tau_i \prod_{\substack{r=\alpha+1\\ r\neq h}}^{i} \frac{1}{1+\tau_r F_r(t)} \frac{-\tau_h}{(1+\tau_h F_h(t))^2} \cdot \sum_{k=\alpha+1}^{\beta} \tau_k \prod_{j=\alpha+1}^{k} \frac{1}{1+\tau_j F_j(t)}}{\left(\sum_{k=\alpha+1}^{\beta} \tau_k \prod_{j=\alpha+1}^{k} \frac{1}{1+\tau_j F_j(t)}\right)^2} + \clubsuit$$

$$= \clubsuit - \frac{w_i \tau_h}{(1+\tau_h F_h(t)) \sum_{k=\alpha+1}^{\beta} \tau_k \prod_{j=\alpha+1}^{k} \frac{1}{1+\tau_j F_j(t)}} .$$

Hence the proof. □

By denote

$$\overline{w}_h(t) := w_h(t) + \sum_{i=\alpha+1}^{\beta} F_i(t) \frac{\partial w_i(t)}{\partial F_h} ,$$

the swap rate dynamics can be rewritten as

$$dS_{\alpha,\beta}(t) = \sum_{h=\alpha+1}^{\beta} \overline{w}_h(t) dF_h(t) + (\cdots)dt .$$

Now, freeze all F's at their initial values in order to approximate

$$dS_{\alpha,\beta}(t) \approx \sum_{h=\alpha+1}^{\beta} \overline{w}_h(0) dF_h(t) .$$

Finally, as in the Rebonato's procedure, derive the volatility-like quantity $\overline{v}_{\alpha,\beta}^{LLM}$ to be entered in the Black's formula for swaptions.

Proposition 4.2.4 (Hull and White's formula)**.** *The LMM Black-like swaption volatility can be better approximated by* $\overline{v}_{\alpha,\beta}^{\mathrm{LMM}}$ *, with*

$$(\overline{v}_{\alpha,\beta}^{\mathrm{LMM}})^2 := \sum_{i,j=\alpha+1}^{\beta} \frac{\overline{w}_i(0)\overline{w}_j(0)F_i(0)F_j(0)\rho_{i,j}}{T_\alpha\, S_{\alpha,\beta}^2(0)} \int_0^{T_\alpha} \sigma_i(t)\sigma_j(t)dt .$$

4.3 Example: computational results of the different methods of swaption pricing

We now show the results we obtain by implementing in Mathematica all the different methods we presented above to price the swaptions in the LMM.

We started from an annual tenor structure

$$\{T_0 = 1y, \ldots, T_{49} = 50y\}$$

and with the market quotes for the spot-starting swap rates (i.e. where the first reset date is the day of valuation, i.e. the 26th. of April, 2011)

$$S_{-1,0}(0), S_{-1,1}(0), \ldots, S_{-1,49}(0)\,,$$

from which we recover through a bootstrapping procedure and a log-linear interpolation, the zero-bond curve shown in Figure 4.1.

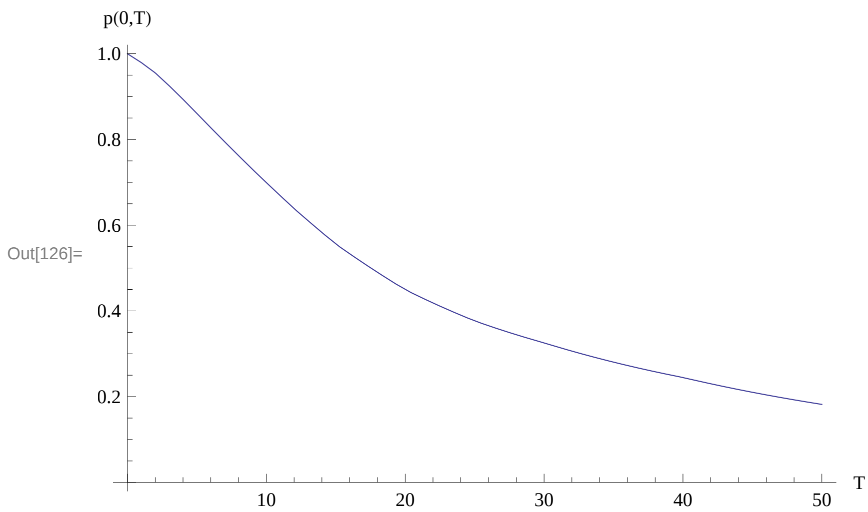

Figure 4.1: Zero-bond curve obtained from market data in April 26,2011.

Furthermore, as other inputs to our functions, we have:

- an historical correlation matrix obtained by market data spanning the year before March 29, 2001, and calibrated in the Schoenmakers and Coffey's three-parameter structure, shown later on in Figure 7.8 (we will describe the topics of correlation modeling and historical correlation respectively in Chapter 5 and in Section 7.1);

- the current quotes of LIBOR forward rates modeled, $F_1(0), \ldots, F_{49}(0)$, obtained by their definition in terms of zero-coupon bond prices from the curve in Figure 4.1;

- the forward volatilities obtained by calibrating the model to swaption data with an extended triangular cascade calibration, which we will see in Subsection 7.2.3.

We priced for example three payer swaptions with different maturities and tenors and we display the results in Table 4.1.

	Swaption with tenor $T_\alpha \times T_\beta - T_\alpha$ and fixed rate $K = 0.03$		
	$\alpha = 5,\ \beta = 8$	$\alpha = 5,\ \beta = 14$	$\alpha = 13,\ \beta = 18$
PS^{Reb}	0.0339696	0.100866	0.0467792
PS^{HW}	0.0339599	0.100712	0.0468129
PS^{Br1}	0.0339502	0.100645	0.0468045
PS^{Br3}	0.0339492	0.100643	0.0468043
PS^{MC}	0.0330096	0.100043	0.0467224
99%	[0.0326904, 0.0333288]	[0.099198, 0.100888]	[0.0462018, 0.047243]

Table 4.1: Approximated prices of three payer swaptions with different maturities and tenors, respectively by Rebonato's, Hull and White's, Brace's rank-1 and Brace's rank-3 formulae and Monte Carlo simulation with 100000 scenarios and a time grid of step $dt = \frac{1}{360}y$.

We can notice that, often, the four approximated analytical prices fall all into the 99%-confidence interval provided by the Monte Carlo simulation. Moreover, sometimes they seem to be grouped into two classes, one with the Rebonato's and Hull and White's prices and another with the Braces' prices, among which we cannot state what is in general more accurate with respect to the Monte Carlo result.

Chapter 5

Instantaneous Correlation Modeling

In the LMM setting, it remains to define the instantaneous correlation between all the forward LIBOR rates modeled. In fact, we derived the dynamics of the F_k's as each dependent on a different random source Z_k, that is instantaneously correlated with the others.

Not all interest rate derivatives have the same dependence on correlations. In the pricing of caps we have already observed that the payoff does not depend on the joint dynamics of F's, unlike as the pricing of swaptions does, depending on more than a single rate in a non-linear relation.
An appropriate correlation modeling can be important in the following step of calibration to swaption prices and becomes definitely relevant when the number of contracts to which the model has to be calibrated is large. The challenge is to choose a structure both flexible enough to express a large number of swaption prices and, at a same time, parsimonious enough to be tractable.

The M-dimensional correlated Brownian motion Z is assumed to have an associated constant correlation matrix ρ, which, recalling the Remark (14) in Appendix C, is defined as

$$\rho_{i,j} := \mathrm{Corr}(Z_i(t), Z_j(t)) = \frac{\mathrm{Cov}(Z_i(t), Z_j(t))}{\mathrm{Std}(Z_i(t))\mathrm{Std}(Z_j(t))} = \frac{< Z_i, Z_j >_t}{t}$$

and it also holds that

$$\rho\, dt = d < Z >_t = dZ(t)(dZ(t))' ,$$

from which we call ρ the *instantaneous correlation*.
As far as the LMM is concerned, if we assume the volatilities of forward rates constant on small time intervals of length Δt, then we can consider $\rho_{i,j}$ a quantity roughly summarizing the "degree of dependence" between instantaneous changes of $\ln F_i$ and $\ln F_j$:

$$< \ln F_i, \ln F_j >_t = \int_0^t \sigma_i(s)\sigma_j(s)\rho_{i,j} ds$$

$$\Rightarrow \quad \Delta < \ln F_i, \ln F_j >_t = \int_t^{t+\Delta t} \sigma_i(s)\sigma_j(s)\rho_{i,j} ds \; = \; \sigma_i(t)\sigma_j(t)\rho_{i,j}\Delta t$$

and

$$\frac{\Delta < \ln F_i, \ln F_j >_t}{\sqrt{\Delta < \ln F_i, \ln F_i >_t \Delta < \ln F_j, \ln F_j >_t}} = \frac{\sigma_i(t)\sigma_j(t)\rho_{i,j}\Delta t}{\sqrt{\sigma_i^2(t)\sigma_j^2(t)(\Delta t)^2}} = \rho_{i,j} \, ,$$

so that the correlation between different forward rates is completely determined by the correlation between different scalar Brownian motions, as far as the dynamics of the F's are given by the solutions of the SDEs in (2.2).

In the setting of the LMM, we stated that ρ is a square M-dimensional matrix. There exist also low-factor models, i.e. driven by a d-dimensional B.m. with $d < M$, which nevertheless have intrinsic problems to match correlations between forward LIBOR rates realistically, therefore we consider only M-dimensional structures.

Now, we recall the properties of a generic correlation matrix:

1. $\rho_{i,j} = \rho_{j,i} \; \forall i, j$ (symmetry);
2. $|\rho_{i,j}| \leq 1 \; \forall i, j$ (normalization);
3. $\rho_{i,i} = 1 \; \forall i$ (maximum correlation for maximum dependence);
4. $x'\rho x \geq 0 \; \forall x$ (positive semidefinite matrix).

Then, we describe the additional qualities that an instantaneous correlation matrix associated with a LIBOR market model would have:

I. $\rho_{i,j} \geq 0 \; \forall i, j$ (positive correlations);

II. $i \mapsto \rho_{i,j}$ is decreasing $\forall i \geq j$ (joint movements of far away rates are less correlated than ones of rates with close maturities);

III. $i \mapsto \rho_{i+p,i}$ is increasing $\forall i \,, \forall$ fixed $p \in \mathbb{N}$ (the larger the tenor, the more correlated changes in equally spaced forward rates become).

A full-rank correlation matrix ρ has $\frac{M(M-1}{2}$ entries, thanks to the symmetry and the 1's in the diagonal. The high number of parameters can be a problem for practical purposes, e.g. it makes correlations to be irregular when obtained by fitting swaption prices. There are two possible approaches to this problem: to use parameterizations of a full-rank matrix or to reduce the rank.

5.1 Full-rank parameterizations

A general full-rank **semi-parametric structure**, suggested by Schoenmakers and Coffey (2000) is:

$$\rho^{SC}(c)_{i,j} := \frac{\min\{c_i, c_j\}}{\max\{c_i, c_j\}} \,, \quad i, j = 1, \ldots, M \,, \tag{5.1}$$

where $c \in (\mathbb{R}^M)^+$ with components such that

$$1 = c_1 < c_2 < \ldots < c_M \quad \text{and} \quad \frac{c_1}{c_2} < \frac{c_2}{c_3} < \ldots < \frac{c_{M-1}}{c_M} \,. \tag{5.2}$$

Remark 9. $\rho^{SC}(c)$ is determined by $M-1$ parameters and satisfies all necessary properties.

Proof. 1. follows directly from the definition;

2. , 3. , I. are obviously derived;

4. is hard to prove;

II. $\forall$ fixed j , the map

$$i \mapsto \rho_{i,j} = \frac{c_j}{c_i}\,, \quad i \geq j$$

is decreasing, because the c_i's are increasing;

III. For all $i \in \{1,\dots,M-p\}$:

$$\rho^{SC}(c)_{i+p,i} = \tfrac{c_i}{c_i+p} = \prod_{k=0}^{p-1} \tfrac{c_i+k}{c_i+k+1} = \tfrac{c_i}{c_i+1} \cdot A$$

$$\text{where} \qquad A := \prod_{k=1}^{p-1} \tfrac{c_i+k}{c_i+k+1}\,,$$

$$\rho^{SC}(c)_{i+1+p,i+1} = \tfrac{c_i+1}{c_i+1+p} = \prod_{k=0}^{p-1} \tfrac{c_i+1+k}{c_i+1+k+1} = \prod_{k=1}^{p} \tfrac{c_i+k}{c_i+k+1}$$

$$= A \cdot \tfrac{c_i+p}{c_i+p+1}\,,$$

$$\text{but} \quad \tfrac{c_i}{c_i+1} < \tfrac{c_i+p}{c_i+p+1} \quad \text{from definition}\,.$$

□

For this reason this form is called *semi-parametric*, in the sense that it depends on $M-1$ parameters, i.e. $\mathcal{O}(M)$, rather than $\mathcal{O}(M^2)$ in case of a non-parametric $M \times M$ matrix, and consequently the parameter dimension increase proportionally to the model dimension.

Let's give an idea of the practical meaning of this structure, by showing that there exists a class of random vectors which have a correlation matrix satisfying the conditions (5.1)-(5.2).

Let $c_1 < c_2 < \dots < c_M$ be an arbitrary positive increasing sequence with $c_1 = 1$ and let $W_i\,, i = 1,\dots,M$, be standard normally distributed independent real r.v., then consider the random vector Y with components

$$Y_i := \sum_{k=1}^{i} a_k W_k \;\sim\; \mathcal{N}\left(0, \sum_{k=1}^{i} a_k^2\right)\,,$$

where

$$a_1 = c_1 = 1 \quad \text{and} \quad a_i := \sqrt{c_i^2 - c_{i-1}^2} \;\text{ for } i = 2,\dots,M\,.$$

We rewrite Y as

$$Y = A \cdot W \sim \mathcal{N}(0, A\,A'), \tag{5.3}$$

where

$$A = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ a_1 & a_2 & 0 \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ & & & 0 \\ a_1 & a_2 & \cdots & a_M \end{pmatrix}.$$

Then the covariance between Y_i and Y_j, for $i \leq j$, is

$$\mathrm{Cov}(Y_i, Y_j) = (A\,A')_{i,j} = \sum_{k=1}^{i} a_k^2 = c_i^2\,,$$

because $\sum_{k=1}^{i} a_k^2 = c_1^2 + \sum_{k=2}^{i} (c_k^2 - c_{k-1}^2) = c_i^2$.
Thus the correlation between them is

$$\rho_{Y_i,Y_j} = \frac{\mathrm{Cov}(Y_i, Y_j)}{\sqrt{\sum_{k=1}^{i} a_k^2}\sqrt{\sum_{k=1}^{j} a_k^2}} = \frac{c_i^2}{c_i\,c_j} = \frac{c_i}{c_j}.$$

Hence it follows that the correlation matrix of Y is given by a structure of kind (5.1)-(5.2), which consequently defines a real correlation matrix.

Theorem 5.1.1. *Every correlation structure of type (5.1)-(5.2) can be equivalently characterized in terms of a sequence of non-negative numbers,* $\Delta_2, \ldots, \Delta_M$, *in the following representation:*

$$c_i = \exp\left(\sum_{l=2}^{M} \min\{i-1, l-1\}\Delta_l\right). \tag{5.4}$$

Proof. We now prove the direction (5.1)-(5.2) $\Rightarrow$ (5.4), whereas the converse follows straightforwardly by checking (5.1)-(5.2) for the sequence $(c_i)_{i=1,\ldots,M}$ defined by (5.4).

Define $\xi_i := \ln c_i$, $1 \leq i \leq M$. Then $\xi_1 = 0$, since $c_1 = 1$, and for the sequence (ξ_i) the following constraints yield:

$$\xi_i < \xi_{i+1}\,, \qquad 1 \leq i \leq M-1\,; \tag{5.5}$$

$$\xi_{i-1} + \xi_{i+1} < 2\xi_i\,, \qquad 2 \leq i \leq M-1\,. \tag{5.6}$$

Indeed: (5.5) follows from the increasing monotony of both the sequence (c_i) and the function ln; (5.6) follows from the increasing monotony of the function $i \mapsto \frac{c_i}{c_{i+1}}$ and consequently we have $\ln c_{i-1} - \ln c_i < \ln c_i - \ln c_{i+1}$, with again ln $\nearrow$.
Now, introduce the new variables

$$\begin{aligned} \Delta_i &:= 2\xi_i - \xi_{i-1} - \xi_{i+1} = (\xi_i - \xi_{i-1}) - (\xi_{i+1} - \xi_i) \geq 0\,, \quad (5.7) \\ &\quad \text{for } 2 \leq i \leq M-1\,, \\ \Delta_M &:= \xi_2 - \sum_{j=2}^{M-1} \Delta_j\,. \quad (5.8) \end{aligned}$$

Hence, for $2 \leq i \leq M$,

$$\begin{aligned} \xi_i = \xi_i - \xi_1 &= \sum_{k=2}^{i} (\xi_k - \xi_{k-1}) = \sum_{k=2}^{i} (\xi_k - \xi_{k-1} - (\xi_2 - \xi_1) + \xi_2 - \xi_1) \\ &= \xi_2 + \sum_{k=3}^{i} \left(\sum_{l=2}^{k-1} (\xi_{l+1} - \xi_l - (\xi_l - \xi_{l-1})) + \xi_2 \right) \\ &= (i-1)\xi_2 + \sum_{k=3}^{i} \sum_{l=2}^{k-1} (\xi_{l+1} - \xi_l - (\xi_l - \xi_{l-1})) \\ &= (i-1)\xi_2 - \sum_{k=3}^{i} \sum_{l=2}^{k-1} \Delta_l = (i-1)\xi_2 - \sum_{l=2}^{i-1} \sum_{k=l+1}^{i} \Delta_l \\ &= (i-1)\xi_2 - \sum_{l=2}^{i-1} (i-l)\Delta_l\,. \quad (5.9) \end{aligned}$$

It follows that

$$\begin{aligned} \xi_{i+1} - \xi_i &= \xi_2 - \sum_{l=2}^{i} (i+1-l)\Delta_l + \sum_{l=2}^{i-1} (i-l)\Delta_l \\ &= \xi_2 - \sum_{l=2}^{i} \Delta_l - (l-l)\Delta_l \quad (5.10) \\ &= \xi_2 - \sum_{l=2}^{i} \Delta_l \quad (5.11) \end{aligned}$$

and the constraints (5.5)-(5.6) transform into

$$\Delta_i \geq 0\,, \qquad 2 \leq i \leq M\,.$$

Indeed:

$$(5.5),(5.11) \;\Leftrightarrow\; \xi_2 > \sum_{l=2}^{i} \Delta_l\,, \text{ for } 2 \le i \le M\,;$$

$$(5.6),(5.11) \;\Leftrightarrow\; \sum_{l=2}^{i} \Delta_l - \xi_2 = \xi_i - \xi_{i+1} > \xi_{i-1} - \xi_i = \sum_{l=2}^{i-1} \Delta_l - \xi_2 \text{ for } 2 \le i \le M-1 \;\Leftrightarrow\; \Delta_i > 0 \text{ for } 2 \le i \le M-1\,.$$

Then, by (5.8) and (5.9), we may express the ξ's in terms of the new coordinates Δ's:

$$\begin{aligned}
\xi_i &= (i-1)\sum_{l=2}^{M} \Delta_l - \sum_{l=2}^{i-1}(i-l)\Delta_l \\
&= i\left(\sum_{l=2}^{M} \Delta_l - \sum_{l=2}^{i-1} \Delta_l\right) - \sum_{l=2}^{M} \Delta_l + \sum_{l=2}^{i-1} l\Delta_l \\
&= (i-1)\sum_{l=i}^{M} \Delta_l - \sum_{l=2}^{i-1} \Delta_l - \sum_{l=i}^{M} \Delta_l + \sum_{l=2}^{i-1} l\Delta_l \\
&= (i-1)\sum_{l=i}^{M} \Delta_l + \sum_{l=2}^{i-1}(l-1)\Delta_l \\
&= \sum_{l=2}^{M} \min\{(i-1),(l-1)\}\Delta_l\,.
\end{aligned}$$

Finally substitute the last equation in $c_i = \exp(\xi)$. □

Remark 10. A correlation structure (5.1)-(5.2), by representation (5.4), yields

$$\rho_{i,j}^{SC} = \exp\left(-\sum_{j=i+1}^{M} \min\{l-i, j-i\}\Delta_l\right)\,, \quad i < j\,. \tag{5.12}$$

Proof.

$$\ln\frac{c_i}{c_j} = \sum_{l=2}^{M} \min\{i-1,l-1\}\Delta_l - \sum_{l=2}^{M} \min\{j-1,l-1\}\Delta_l\,,$$

from which we have three cases:

$$\min\{i-1,l-1\} - \min\{j-1,l-1\} = \begin{cases} i-l, & i \le l < j\,; \\ i-j, & i < j \le l\,; \\ 0, & l \le i < j\,. \end{cases}$$

Hence the proof. ☐

From representation (5.4), particularly in the form (5.12), we can derive conveniently various low parametric structures, presented below.

5.1.1 Low-parametric structures by Schoenmakers and Coffey

▶ By taking $\Delta_2, \ldots, \Delta_{M-1} =: \alpha \geq 0$, and $\Delta_M =: \beta \geq 0$, we obtain the following two-parameters form:

$$\rho_{i,j} = \exp\left[-|i-j|\left(\beta + \alpha\left(M - \frac{i+j+1}{2}\right)\right)\right], \quad i,j = 1,\ldots,M. \quad (5.13)$$

Proof. For $j < M$:

$$\begin{aligned}
\rho_{i,j} &= \exp\left[-\alpha \sum_{l=i+1}^{j}(l-i) - \alpha \sum_{l=j+1}^{M-1}(j-i) - \beta(j-i)\right] \\
&= \exp\left[\alpha i(M-1-i) - \alpha \sum_{l=i+1}^{j} l - \alpha j(M-1-j) - \beta(j-i)\right] \\
&= \exp\left[\alpha i(M-1-i) - \alpha\left(\frac{j(j+1)}{2} - \frac{i(i+1)}{2}\right) + \right. \\
&\qquad \left. -\alpha\, j(M-1-j) - \beta(j-i)\right] \\
&= \exp\left[\alpha i\left(M - \frac{1}{2} - \frac{i}{2}\right) - \alpha\frac{ij}{2} + \alpha\frac{ij}{2} + \right. \\
&\qquad \left. -\alpha\, j\left(M - \frac{1}{2} - \frac{j}{2}\right) - \beta(j-i)\right] \\
&= \exp\left[-(j-i)\left(\beta + \alpha\left(M - \frac{i+j+1}{2}\right)\right)\right].
\end{aligned}$$

For $j = M$:

$$\begin{aligned}
\rho_{i,M} &= \exp\left[-\alpha \sum_{l=i+1}^{M-1} (l-i) - \beta(M-i)\right] \\
&= \exp\left[\alpha i(M-1-i) - \alpha\left(\frac{(M-1)M}{2} - \frac{i(i+1)}{2}\right) - \beta(M-i)\right] \\
&= \exp\left[\alpha i\left(M - \frac{1}{2} - \frac{i}{2}\right) - \alpha M\frac{M-1}{2} - \beta(M-i)\right] \\
&= \exp\left[(i-M)\left(\beta + \alpha\frac{M-i-1}{2}\right)\right] .
\end{aligned}$$

□

Notice that for $\alpha = 0$ we get the simple correlation structure

$$\rho_{i,j} = e^{-\beta|i-j|}\,, \tag{5.14}$$

which is frequently used in practice in spite of the unrealistic fact that property III. is not satisfied, as the sub-diagonals are constant rather than increasing, i.e. the correlation between forward rates that are equally spaced is constant in time.

Now introduce two new parameters:

$$\rho_\infty := \rho_{1,M}\,, \qquad \text{and} \qquad \eta := \frac{\alpha}{2}(M-1)(M-2)\,,$$

where ρ_∞ is the correlation between the farthest forward rates in the family considered. Hence, by computing $\ln\rho_\infty$ in terms of α, β and inverting the definition of η, we obtain the old parameters in terms of the new ones as

$$\beta = -\frac{\alpha}{2}(M-2) - \frac{\ln\rho_\infty}{M-1}\,, \quad \alpha = \frac{2\eta}{(M-1)(M-2)}\,,$$

from which the form (5.13) transform into

$$\rho_{i,j} = \exp\left[-\frac{|i-j|}{M-1}\left(-\ln\rho_\infty + \eta\frac{M+1-i-j}{M-2}\right)\right] . \tag{5.15}$$

This re-parametrization improves the parameter stability: relatively small movements in the c-sequence associated with (5.15), and consequently in the correlations themselves, cause relatively small movements in the parameters ρ_∞ , η .

► Suppose $M > 2$ and take the Δ_i following a straight line for $i = 2, \ldots, M-1$ and choose one parameter for $i = M$. Precisely:

$$\Delta_2 = \alpha_1 \geq 0\,, \quad \Delta_{M-1} = \alpha_2 \geq 0\,, \quad \Delta_M = \beta \geq 0\,,$$

and for $i = 2, \ldots, M-1$ we have

$$\Delta_i = \alpha_1 \frac{M-i-1}{M-3} + \alpha_2 \frac{i-2}{M-3}\,.$$

Indeed, the equation representing the line, neither vertical nor horizontal, through the two distinct points $(2, \alpha_1)$ and $(M-1, \alpha_2)$ is

$$\Delta_i = \frac{\alpha_2 - \alpha_1}{M-1-2}(i-2) + \alpha_1\,,$$

where the coordinates of the points are (i, Δ_i) for each i.
Then we get the following three-parameter form:

$$\rho_{i,j} = \exp\left[-|i-j|\left(\beta - \tfrac{\alpha_2}{6M-18}(i^2+j^2+ij-6i-6j-3M^2+15M-7) + \right.\right.$$
$$\left.\left. + \tfrac{\alpha_1}{6M-18}(i^2+j^2+ij-3Mi-3Mj+3i+3j+3M^2-6M+2)\right)\right]\,. \tag{5.16}$$

Notice that (5.16) collapses to (5.13) if $\alpha_1 = \alpha_2 = \alpha$, in which case we would have an horizontal line for the first $M-2$ points.
In order to gain stability, as above, re-parameterize (5.16) by introducing $\rho_\infty := \rho_{1,M}$, which yields

$$\beta = -\frac{\ln \rho_\infty}{M-1} - \frac{\alpha_1}{6}(M-2) - \frac{\alpha_2}{3}(M-2)\,,$$

and by setting

$$\alpha_1 = \frac{6\eta_1 - 2\eta_2}{(M-1)(M-2)}\,, \quad \alpha_2 = \frac{4\eta_2}{(M-1)(M-2)}\,.$$

Then (5.16) becomes

$$\rho_{i,j} = \exp\left[-\tfrac{|i-j|}{M-1}\left(-\ln \rho_\infty + \eta_1 \tfrac{i^2+j^2+ij-3Mi-3Mj+3i+3j+2M^2-M-4}{(M-2)(M-3)} + \right.\right.$$
$$\left.\left. -\eta_2 \tfrac{i^2+j^2+ij-Mi-Mj-3i-3j+3M+2}{(M-2)(M-3)}\right)\right]\,. \tag{5.17}$$

Notice that for $\eta_1 = \eta_2 = \frac{\eta}{2}$ the structure (5.17) collapses to (5.15) again, as it is obvious.

▶ The calibration experiments of Schoenmakers and Coffey pointed out that correlation structure (5.17) suits very well in practice. However, calibrating a three-parameter matrix takes longer than a two-parameter one. Furthermore, the experiments reveal that $\eta_2 \approx 0$ in (5.17), i.e. the final point of the straight line in the Δ's is practically always close to 0. Thus we may adopt the following computationally improved correlation structure:

$$\begin{aligned} \rho_{i,j} &= \exp\Big[-\tfrac{|i-j|}{M-1}\Big(-\ln\rho_\infty + \\ &\qquad +\eta\tfrac{i^2+j^2+ij-3Mi-3Mj+3i+3j+2M^2-M-4}{(M-2)(M-3)}\Big)\Big], \end{aligned} \tag{5.18}$$

where the characteristics are the same of (5.17) apart from setting
$\eta_2 = 0\,,\ \eta := \eta_1\,.$

▶ Consider the sequence (c_i) defined by

$$c_i := e^{\beta(i-1)^\alpha}\,, \quad 1 \le i \le M\,,\ \beta > 0\,,\ 0 < \alpha < 1\,.$$

The associated ρ satisfies the assumptions of a correlation structure (5.1)-(5.2). Indeed:

$$\ln\left(\frac{c_i}{c_{i+1}}\right) = \ln c_i - \ln c_{i+1} = \beta\left((i-1)^\alpha - i^\alpha\right) < \beta\left(i^\alpha - (i-1)^\alpha\right)$$

because $i \mapsto i^\alpha$ is an increasing function but with decreasing slope.
Then, by the definitions (5.7)-(5.8) in the proof of Theorem 5.1.1, we get the coordinates

$$\begin{aligned} \Delta_i &= 2\beta(i-1)^\alpha - \beta(i-2)^\alpha - \beta i^\alpha\,, \quad 2 \le i \le 2M-1\,, \\ \Delta_M &= \beta - \sum_{l=2}^{M-1}\Delta_l = \beta - \beta\sum_{l=2}^{M-1}\left(2(l-1)^\alpha - (l-2)^\alpha - l^\alpha\right) \\ &= \beta\left(1 + (M-1)^\alpha - 2(M-2)^\alpha + (M-2)^\alpha + 1 - 2\right) \\ &= \beta\left((M-1)^\alpha - (M-2)^\alpha\right)\,, \end{aligned}$$

where $\Delta_i > 0$ for $2 \le i \le M$, as it must be.
Then, by introducing $\rho_\infty := \rho_{1,M} = \frac{1}{c_M}$, we get the correlations

$$\rho_{i,j} = \exp\left[\ln\rho_\infty\left|\left(\frac{i-1}{m-1}\right)^\alpha - \left(\frac{j-1}{m-1}\right)^\alpha\right|\right]\,. \tag{5.19}$$

The correlation structures (5.18) and (5.19) have similar properties, but calibration experiments of Schoenmakers and Coffey pointed out that (5.18) performs a little better.

5.1.2 Classical two-parameter structure

$$\rho_{i,j} = \rho_\infty + (1-\rho_\infty)\exp\left(-\beta|i-j|\right) \,, \quad \beta \geq 0 \,, \tag{5.20}$$

where ρ_∞ represents only asymptotically the correlation between the farthest forward rates in the family.
This formulation guarantees all the desirable properties apart from III., in the sense that sub-diagonals are flat. Notice that for $\rho_\infty = 0$ we get the simple structure (5.14), again.

5.1.3 Rebonato's three-parameter structure

Rebonato suggested the following perturbation of the classical structure:

$$\rho_{i,j} = \rho_\infty + (1-\rho_\infty)\exp\left[-|i-j|\left(\beta - \alpha\max\{i,j\}\right)\right] \,. \tag{5.21}$$

This structure recovers the property III., in fact

$$i \mapsto \rho_{i,j} = \rho_\infty + (1-\rho_\infty)\exp\left[-p\left(\beta - \alpha(i+p)\right)\right]$$

is increasing for $\alpha > 0$. Thus it may produce realistic market correlations for properly chosen $\rho_\infty, \beta > 0$ and small $\alpha > 0$. However, (5.21) does not fit in the general form (5.1)-(5.2) and its $(\alpha, \rho_\infty, \beta)$ domain of positivity is not explicitly specified, hence it is not guaranteed to be a correlation structure. To avoid this problem, at every step of a hypothetical calibration/optimization, it must be verified that the resulting matrix is positive semi-definite, for example through a constrained optimization. Furthermore for inappropriate values of ρ_∞, β and not small enough α, it may also happen that $\rho_{i,j} > 1$, thus the (5.21) is a weak characterization as correlation's parametrization.

5.2 Reduced rank parameterizations

Given ρ a positive definite symmetric matrix, it can be rewritten as

$$\rho = P\Delta P' \,, \tag{5.22}$$

where Δ is the diagonal matrix of (positive) eigenvalues of ρ in decreasing order and P the orthogonal matrix whose columns are the corresponding eigenvectors. Indeed:

$$\rho P = P\Delta \,, \quad \text{and} \quad P^{-1} = P' \,.$$

Let Λ be the diagonal matrix whose entries are the square roots of the corresponding ones of Δ, i.e. $\Delta = \Lambda\Lambda$.
Then define $C := P\Lambda$ in order to have

$$\rho = CC' \,, \quad C'C = \Delta \,.$$

The correlated M-dimensional Brownian motion is distributed as

$$dZ \sim \mathcal{N}(0, \rho dt) \,,$$

whereas the standard M-dimensional Brownian motion is

$$dW \sim \mathcal{N}(0, Idt) \,,$$

but we can replace dZ with CdW, where $\rho = CC'$.

If $rank\, \rho = r < M$, there exist a r-rank $M \times r$ matrix B such that

$$\rho = BB' \,, \quad \text{and} \quad dZ = BdW \,.$$

The advantage in this replacement is that now we have a r-dimensional random shock, given by the r-dimensional standard B.m. W.

Even when ρ is a full rank matrix, we may try and mimic the decomposition $\rho = CC'$ with a r-rank $M \times r$ correlation matrix B, by introducing a new noise correlation matrix $\rho^B = BB'$, with typically $r << M$.

5.2.1 Rebonato's angles method

Consider the $M \times r$ matrix

$$B = (b_{i,j})_{\substack{i=1,\dots,M \\ j=1,\dots,r}} ,$$

where for $i = 1, \dots, M$ the i-th row of B is

$$\begin{aligned} b_{i,1} &= \cos\theta_{i,1} \\ b_{i,k} &= \cos\theta_{i,k}\sin\theta_{i,1}\cdots\sin\theta_{i,k-1}\,, \quad 1 < k < r \\ b_{i,r} &= \sin\theta_{i,1}\cdots\sin\theta_{i,r-1} \end{aligned} \tag{5.23}$$

Notice that

$$\rho^B := BB'$$

results a positive semi-definite matrix and all its diagonal terms are equal to 1, thus it is a possible correlation matrix.

The number of parameters of this r-factor structure is $M(r-1)$, thus the angles parametrization is not necessarily reducing the dimension of the problem. In fact keeping for instance full-rank ρ, i.e. $r = M$, the number of parameters is then $M(M-1)$, that is twice the number of entries of a generic correlation matrix. To avoid a such inconvenient we have to ask not only $r < M$, but rather

$$r << \frac{M+1}{2} .$$

5.2.2 Reduced rank approximations of exogenous correlation matrices

When the correlation matrix is given exogenously as an historical estimation, instead of being a fitting parameter in the calibration to the swaption market, it has full rank M. The following two methods deal only with this kind of situation.

The target is to obtain an approximated r-rank correlation matrix $\rho^{(r)}$, with $r < M$:

$$M\text{-rank historical } \rho \longrightarrow r\text{-rank } \rho^{(r)} .$$

Two different approaches are presented below.

Eigenvalues zeroing with normalization

Consider the decomposition (5.22) and define $\Lambda^{(r)}$ as the $r \times r$ diagonal matrix obtained from Λ by taking away the $M-r$ smallest diagonal elements, together with the corresponding dimensions, and $P^{(r)}$ as the $M \times r$ matrix obtained from P by taking away the $M-r$ corresponding columns.
Thus define $B^{(r)} := P^{(r)}\Lambda^{(r)}$, getting the matrix

$$\overline{\rho}^{(r)} := B^{(r)}(B^{(r)})',$$

which is the best reduced rank approximation of ρ according to Frobenius norm.

Remark 11. $\overline{\rho}^{(r)}$ results positive semi-definite, but it does not features ones in the diagonal.

The solution is to interpret $\overline{\rho}^{(r)}$ as a covariance matrix and derive the associated correlation matrix:

$$\rho^{(r)}_{i,j} := \frac{\overline{\rho}^{(r)}_{i,j}}{\sqrt{\overline{\rho}^{(r)}_{i,i}\overline{\rho}^{(r)}_{j,j}}}.$$

Indeed, given a generic M-dimensional random variable X with covariance matrix $\Sigma = (\sigma_{i,j})$, which is a $M \times M$ positive semi-definite matrix, the associated correlation matrix is

$$\rho = \Lambda^{-1}\Sigma\Lambda^{-1},$$

where

$$\Lambda = \begin{pmatrix} \sqrt{\sigma_{1,1}} & & \\ & \ddots & \\ & & \sqrt{\sigma_{M,M}} \end{pmatrix},$$

and $\sigma_i^2 = \sigma_{i,i}$ is the variance of the i-th component X_i, $i = 1, \ldots, M$.

Optimization on angles parametrization (Rebonato and Jäkel, 1999)

Given the full rank correlation matrix ρ in input, we can find its r-rank better approximation $\rho(\theta) = B(\theta)B(\theta)'$ in terms of angles as in (5.23). Precisely, once defined the i-th row of $B(\theta)$ as

$$\begin{aligned} b_{i,1}(\theta) &= \cos\theta_{i,1} \\ b_{i,k}(\theta) &= \cos\theta_{i,k}\sin\theta_{i,1}\cdots\sin\theta_{i,k-1}\,, \quad 1 < k < r \\ b_{i,r}(\theta) &= \sin\theta_{i,1}\cdots\sin\theta_{i,k-1}\,, \end{aligned}$$

minimize with respect to (θ) the quantity

$$\sum_{i,j=1}^{M}(|\rho_{i,j} - \rho_{i,j}(\theta)|^2)\,.$$

This method gives the optimal solution through an unconstrained optimization.

Comparing the methods of *eigenvalues zeroing* and *angles optimization*, we may see a noticeable better fit with the last one, despite a few seconds of computation, with a low rank. This difference is going to attenuate by increasing the rank.
In most situations a rank-4 approximation is satisfactory with both methods, whereas in extreme cases we have to resort to a rank up the 7-th.

Chapter 6

Calibration of the LMM

6.1 Calibration of the LMM to Caplets

The LIBOR market model is calibrated to the most traded derivatives among the liquid ones, namely the caplets, in an almost automatic way. Indeed, the transition from caps to caplets is made by traders, then the calibration of the LMM parameters to caps and floors results trivial, thanks to the Proposition 2.1.1 and using market quoted volatilities.

Let assume that we are standing at time $t = 0$. Given an empirical term structure of implied spot volatilities, from Proposition 2.1.1 it follows that calibrating the model to caplets amounts to choose the deterministic LIBOR volatilities of forward rates $\sigma_1, \ldots, \sigma_M$ such that

$$v^2_{T_i-capl} = \frac{1}{T_{i-1}} \int_0^{T_{i-1}} \sigma_i(t)^2 dt\,, \quad i = 1, \ldots, M\,. \tag{6.1}$$

6.1.1 Parameterizations of Volatility of Forward Rates

The system (6.1) is highly undetermined, thus it needs structural assumptions about the shape of volatility functions. The most popular specifications fall into two main categories: piecewise-constant functions, in which case the σ's are constant in each expiry-maturity time interval in which the corresponding forward rates are alive, and functional parameterized forms, as shown below.

- General Piecewise-Constant volatilities (GPC):

$$\sigma_i(t) = \sigma_{i,\beta(t)}\,, \quad 0 < t \leq T_{i-1}\,, \tag{6.2}$$

where

$$\beta(t) = m \quad \text{if} \quad T_{m-2} < t \leq T_{m-1}\,, \quad m \geq 1\,. \tag{6.3}$$

The index $\beta(t)$ indicates the maturity of the first forward rate that has not expired yet by time t. These σ's can be organized in Table 6.1, where we have the time intervals in the columns and the relative rates in the rows.

	$t \in (0, T_0]$	$(T_0, T_1]$	$(T_1, T_2]$	$\cdots$	$(T_{M-2}, T_{M-1}]$
$F_1(t)$	$\sigma_{1,1}$	dead	dead	$\cdots$	dead
$F_2(t)$	$\sigma_{2,1}$	$\sigma_{2,2}$	dead	$\cdots$	dead
$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$F_M(t)$	$\sigma_{M,1}$	$\sigma_{M,2}$	$\sigma_{M,3}$	$\cdots$	$\sigma_{M,M}$

Table 6.1: General Piecewise-Constant volatilities.

To perform calibration, by inserting the expression (6.2) in the equation (6.1), we impose:

$$T_{i-1} v^2_{T_i - capl} = \int_0^{T_{i-1}} \sigma^2_{i,\beta(t)} dt = \frac{1}{T_{i-1}} \sum_{j=1}^{i} \tau_{j-2,j-1} \sigma^2_{i,j}\,, \tag{6.4}$$

for $i = 1, \ldots, M$.

There exists multiple configurations that can fit caplets perfectly. In fact:

$$\begin{aligned}
\sqrt{T_0} v_{T_1-capl} = \sqrt{(T_0 - 0)}\sigma_{1,1} \;\Leftrightarrow\; \sigma_{1,1} &= v_{T_1-capl} \\
\sqrt{T_0 \sigma^2_{2,1} + (T_1 - T_0)\sigma^2_{2,2}} &= \sqrt{T_1} v_{T_2-capl} \\
\sqrt{T_0 \sigma^2_{3,1} + (T_1 - T_0)\sigma^2_{3,2} + (T_2 - T_1)\sigma^2_{3,3}} &= \sqrt{T_2} v_{T_3-capl} \\
&\;\vdots
\end{aligned}$$

Then we can make some assumptions on this structure in order to reduce the number of parameters.

- Piecewise-constant volatilities dependent only on the time to maturity $T_i - T_{\beta(t)-1}$ of corresponding forward rates:

$$\sigma_i(t) = \sigma_{i,\beta(t)} = \eta_{i-(\beta(t)-1)}\,, \quad 0 < t \le T_{i-1}\,. \tag{6.5}$$

They are organized in Table 6.2.

	$t \in (0, T_0]$	$(T_0, T_1]$	$(T_1, T_2]$	$\cdots$	$(T_{M-2}, T_{M-1}]$
$F_1(t)$	η_1	dead	dead	$\cdots$	dead
$F_2(t)$	η_2	η_1	dead	$\cdots$	dead
$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$F_M(t)$	η_M	η_{M-1}	η_{M-2}	$\cdots$	η_1

Table 6.2: Time-to-maturity-dependent volatilities.

To perform calibration, by inserting the expression (6.5) in the equation (6.1), we impose:

$$T_{i-1} v^2_{T_i-capl} = \sum_{j=1}^{i} \tau_{j-2,j-1} \eta^2_{i-j+1}\,, \tag{6.6}$$

for $i = 1, \ldots, M$.
In this case we can find the parameters η's that exactly fit the market caplets volatilities, being each of those obtained in terms of the previous. In fact:

$$\sqrt{(T_0 - 0)}\eta_1 = \sqrt{T_0} v_{T_1-capl} \;\Leftrightarrow\; \eta_{1,1} = v_{T_1-capl}$$

$$\sqrt{T_0 \eta_2^2 + (T_1 - T_0)\eta_1^2} = \sqrt{T_1} v_{T_2-capl} \;\Leftrightarrow$$
$$\Leftrightarrow\; \eta_2 = \frac{1}{\sqrt{T_0}} \sqrt{T_1 v^2_{T_2-capl} - (T_1 - T_0)\eta_1^2}$$

$$\sqrt{T_0\eta_3^2 + (T_1 - T_0)\eta_2^2 + (T_2 - T_1)\eta_1^2} = \sqrt{T_2}v_{T_3-capl} \Leftrightarrow$$
$$\Leftrightarrow \eta_3 = \frac{1}{\sqrt{T_0}}\sqrt{T_2 v_{T_3-capl}^2 - (T_1 - T_0)\eta_2^2 - (T_2 - T_1)\eta_1^2}$$
$$\vdots$$

- Constant maturity-dependent volatilities:

$$\sigma_i(t) = \sigma_{i,\beta(t)} = s_i\,, \quad 0 < t \leq T_{i-1}\,, \tag{6.7}$$

as shown in Table 6.3.

	$t \in (0, T_0]$	$(T_0, T_1]$	$(T_1, T_2]$	$\cdots$	$(T_{M-2}, T_{M-1}]$
$F_1(t)$	s_1	dead	dead	$\cdots$	dead
$F_2(t)$	s_2	s_2	dead	$\cdots$	dead
$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$F_M(t)$	s_M	s_M	s_M	$\cdots$	s_M

Table 6.3: Constant maturity-dependent volatilities.

To perform calibration, by inserting the expression (6.7) in the equation (6.1), we impose:

$$T_{i-1}v_{T_i-capl}^2 = T_{i-1}s_i^2 \quad \Leftrightarrow \quad v_{T_i-capl}^2 = s_i^2\,, \tag{6.8}$$

for $i = 1, \ldots, M$.
Again, the parameters s's can exactly fit the market caplets volatilities.

- Separable piecewise-constant volatilities, factorized as follows:

$$\sigma_i(t) = \sigma_{i,\beta(t)} = \Phi_i \Psi_{\beta(t)}\,, \quad 0 < t \leq T_{i-1}\,, \tag{6.9}$$

leading to Table 6.4.

This structure includes the one in (6.7) as a special case when all Ψ's

	$t \in (0, T_0]$	$(T_0, T_1]$	$(T_1, T_2]$	$\cdots$	$(T_{M-2}, T_{M-1}]$
$F_1(t)$	$\Phi_1\Psi_1$	dead	dead	$\cdots$	dead
$F_2(t)$	$\Phi_2\Psi_1$	$\Phi_2\Psi_2$	dead	$\cdots$	dead
$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$F_M(t)$	$\Phi_M\Psi_1$	$\Phi_M\Psi_2$	$\Phi_M\Psi_3$	$\cdots$	$\Phi_M\Psi_M$

Table 6.4: $\Phi_i\Psi_{\beta(t)}$ structure

are equal to one. To perform calibration, by inserting (6.9) in (6.1), we impose:

$$T_{i-1}v^2_{T_i-capl} = \Phi_i^2 \sum_{j=1}^{i} \tau_{j-2,j-1}\Psi_j^2\,, \tag{6.10}$$

for $i = 1, \ldots, M$.

In this case, having read from the market the caplet volatilities $v_{T_i-capl} = v_i^{MKT}$, the parameters Φ's can be given in terms of the parameters Ψ's as

$$\Phi_i^2 = \frac{T_{i-1}(v_i^{MKT})^2}{\sum_{j=1}^{i} \tau_{j-2,j-1}\Psi_j^2}\,.$$

- Separable piecewise-constant volatilities, factorized as follows:

$$\sigma_i(t) = \sigma_{i,\beta(t)} = \Phi_i\Psi_{i-(\beta(t)-1)}\,, \quad 0 < t \leq T_{i-1}\,, \tag{6.11}$$

leading to Table 6.5.

This structure includes the one in (6.5) as a special case when all Φ's are equal to one. It concludes the piecewise-constant volatility models. To perform calibration, by inserting (6.11) in (6.1), we impose:

$$T_{i-1}v^2_{T_i-capl} = \Phi_i^2 \sum_{j=1}^{i} \tau_{j-2,j-1}\Psi_{i-j+1}^2\,, \tag{6.12}$$

for $i = 1, \ldots, M$.

Similarly to the previous case, the parameters Φ's can be given in terms

	$t \in (0, T_0]$	$(T_0, T_1]$	$(T_1, T_2]$	$\cdots$	$(T_{M-2}, T_{M-1}]$
$F_1(t)$	$\Phi_1\Psi_1$	dead	dead	$\cdots$	dead
$F_2(t)$	$\Phi_2\Psi_2$	$\Phi_2\Psi_1$	dead	$\cdots$	dead
$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$F_M(t)$	$\Phi_M\Psi_M$	$\Phi_M\Psi_{M-1}$	$\Phi_M\Psi_{M-2}$	$\cdots$	$\Phi_M\Psi_1$

Table 6.5: $\Phi_i\Psi_{i-(\beta(t)-1)}$ structure

of the parameters Ψ's as

$$\Phi_i^2 = \frac{T_{i-1}(v_i^{MKT})^2}{\sum_{j=1}^{i} \tau_{j-2,j-1}\Psi_{i-j+1}^2} .$$

Even the continuous parameterizations have always been very popular. The main examples are the following.

- Parametric linear exponential volatilities, analogue to form (6.5) :

$$\sigma_i(t) = \Psi(T_{i-1} - t; a, b, c, d) := (a(T_{i-1} - t) + d)e^{-b(T_{i-1}-t)} + c \,. \quad (6.13)$$

This time, in order to perform calibration, by inserting (6.13) in (6.1) we obtain a continuous expression:

$$T_{i-1}v_{T_i-capl}^2 = I^2(T_{i-1}; a, b, c, d) \,, \quad (6.14)$$

where

$$I^2(T_{i-1}; a, b, c, d) := \int_0^{T_{i-1}} \left([a(T_{i-1} - t) + d]e^{-b(T_{i-1}-t)} + c\right)^2 dt \,.$$

The parameters a, b, c, d can be used to fit the market the caplet volatilities through an optimization algorithm. This formulation can be perfected into the following richer parametric form.

- Parametric linear exponential volatilities, analogue to form (6.11) :

$$\sigma_i(t) = \Phi_i\Psi(T_{i-1} - t; a, b, c, d) = \Phi_i\left((a(T_{i-1} - t) + d)e^{-b(T_{i-1}-t)} + c\right) . \quad (6.15)$$

This form reduces to the previous by setting all the Φ's to one.
In order to perform calibration, by inserting (6.15) in (6.1) we obtain:

$$T_{i-1}v^2_{T_i-capl} = \Phi_i I^2(T_{i-1}; a, b, c, d)\,. \tag{6.16}$$

Now, having read from the market the caplet volatilities $v_{T_i-capl} = v_i^{MKT}$, the parameters Φ's can be given in terms of the parameters a, b, c, d as:

$$\Phi_i^2 = \frac{T_{i-1}(v_i^{MKT})^2}{I^2(T_{i-1}; a, b, c, d)}\,.$$

Notice that formulation (6.5) and (6.7) allow the complete determination of the parameters in Tables 6.2 and 6.3 respectively, whereas with formulations (6.2), (6.9) and (6.11) we cannot recover the whole tables of such parameters, respectively Tables 6.1, 6.4 and 6.5, since we have more unknown than equations. However, having parameters in excess can be helpful when we have to calibrate the model, either to the swaptions together with the caplets or only to the swaptions.

6.1.2 The Term Structure of Volatility

At each time T_j of the set of expiry-maturity dates of the LMM, the term structure of volatility is the graph of the average volatilities $V(T_j, T_{h-1})$ of the forward rates F_h with fixing times T_{h-1} and maturities T_h, in function of expiry times T_{h-1}. Namely, at time $t = T_j$, it is plotted the set of points

$$\{(T_{j+1}, V(T_j, T_{j+1})), (T_{j+2}, V(T_j, T_{j+2})), \ldots, (T_{M-1}, V(T_j, T_{M-1}))\}$$

where, for $h > j + 1$,

$$V^2(T_j, T_{h-1}) = \frac{1}{\tau_{j,h-1}} \int_{T_j}^{T_{h-1}} \frac{dF_h(t)dF_h(t)}{F_h(t)F_h(t)} = \frac{1}{\tau_{j,h-1}} \int_{T_j}^{T_{h-1}} \sigma_h^2(t)dt\,.$$

In particular, the term structure of volatility at time $t = 0$ is

$$\begin{aligned}&\{(T_0, V(0, T_0)), (T_1, V(0, T_1)), \ldots, (T_{M-1}, V(0, T_{M-1}))\} = \\ &= \{(T_0, v_{T_1-capl}), (T_1, v_{T_2-capl}), \ldots, (T_{M-1}, v_{T_M-capl})\}\,,\end{aligned}$$

that is the market caplet volatility curve. A typical example of this curve for annualized caplet volatilities from the Euro market, on May 4, 2011, is shown in Figure 6.1.2.

In about the 80 percent of cases, the caplet volatility structure occurred

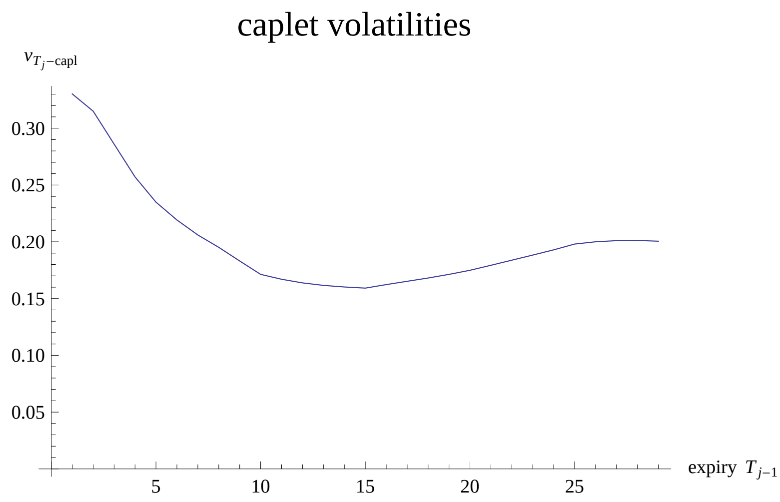

Figure 6.1: Term structure of volatility $T_j \mapsto v_{T_{j+1}-capl}$ from the Euro market, on May 4, 2011; the resettlement dates $T_0, \ldots, T_M$ are annualized and expressed in years.

with this initially humped shape. The short-term forward LIBOR rates (with maturity maximum in a year) depends largely on the choices of the monetary institutions and a bit on everything that happens on the market, e.g. even when you read a news on a newspaper this can influence the rates' trend, therefore they have a high volatility. The long-term forward LIBOR rates does not depend anymore neither on the expectations about the choices of institutions like the BCE nor on minor news, therefore they have a lower volatility. The little hump that is usually located between six months and two or three years is due to the fact that the institutions of central banks try to provide information in advance about their future actions, so that in a very short time the volatility is relatively low, while the peak remains in that period of time that is still influenced by actions and news but is out of the near future in which we have the forecasts.

Up to July 2007 the very-short-term volatilities was very low, but from the half of the same month, coinciding with the beginning of the crisis, these volatilities have abruptly risen. As a consequence, in August 2007, mainly in the USA and in GB, several banks registered big losses. Hence the central banks inverted the way of working, by flooding the market with liquidity. There are a few problems related to the choice of a parametrization for the instantaneous volatilities of forward rates: different assumptions about those imply different evolutions for the term structure of volatility. Moreover, different evolutions can change dramatically the price of some exotic products, nevertheless they fit the whole structure of caplets today.
Let's see the impact of different formulations by starting from the term structure of volatility today and getting its evolution in time.

- Formulation (6.5) gives

$$V^2(T_j, T_{h-1}) = \frac{1}{\tau_{j,h-1}} \sum_{k=j+1}^{h-1} \tau_{k-1,k} \eta_{h-k}^2 .$$

  Assuming for simplicity that the year fractions are all equal to τ, we have that $\tau_{j,h-1} = (h-1-j)\tau$ and another τ factorize out of the summation, thus

$$V^2(T_j, T_{h-1}) = \frac{1}{(h-1-j)} \sum_{k=j+1}^{h-1} \eta_{h-k}^2 ,$$

  which implies that

$$V(T_j, T_{h-1}) = V(T_{j+1}, T_h) .$$

  Therefore, the term structure of volatility simply shifts in time, i.e. when moving from time T_j to time T_{j+1} it moves from

$$\{(T_{j+1}, V(T_j, T_{j+1})), (T_{j+2}, V(T_j, T_{j+2})), \ldots, (T_{M-1}, V(T_j, T_{M-1}))\}$$

  to

$$\{(T_{j+2}, V(T_j, T_{j+1})), (T_{j+3}, V(T_j, T_{j+2})), \ldots, (T_{M-1}, V(T_j, T_{M-1}))\} .$$

The shape of volatility structure remains exactly the same, except that is shifted of one date in the future, consequently becomes shorter, in the sense that the tail of the graph is cut away. An example of this kind of evolution is shown in Figure 6.2, starting from the initial structure of Figure 6.1.2.

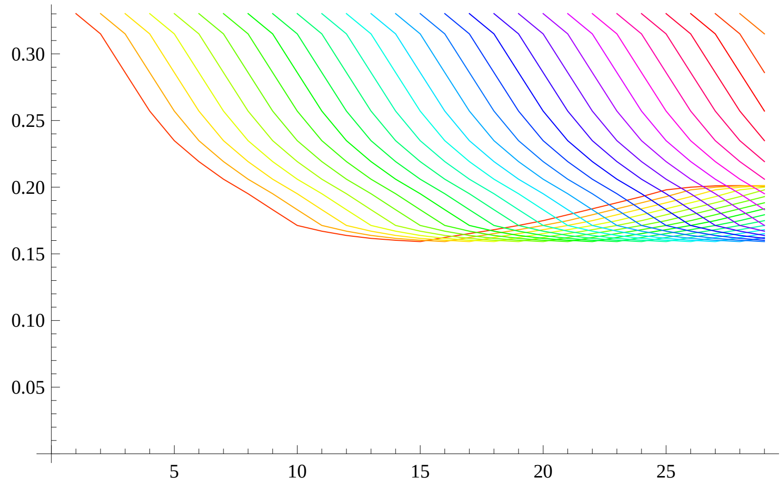

Figure 6.2: Evolution of the term structure of volatility of Figure 6.1.2 obtained by calibrating the parametrization (6.5).

Qualitatively, this is a desirable feature, since the actual shape of the market term structure does not change too much over time. Anyway, there remains the question of whether this formulation allows for a humped structure to be fitted at the initial time. We note that the map

$$T_{h-1} \mapsto \sqrt{T_{h-1}} V(0, T_{h-1}) = \sqrt{\tau \sum_{k=j+1}^{h-1} \eta_{h-k}^2}$$

is increasing. On the other hand, we know that the market volatility structure is decreasing following the hump, namely

$$V(0, T_{h-1}) \geq V(0, T_h)$$

typically for T_h larger that three or four years. Putting these two constraints together we obtain:

$$\begin{array}{c}\sqrt{T_{h-1}}V(0,T_{h-1}) \leq \sqrt{T_h}V(0,T_h) \\ V(0,T_{h-1}) \geq V(0,T_h)\end{array} \quad \Rightarrow \quad 1 \leq \frac{V(0,T_{h-1})}{V(0,T_h)} \leq \sqrt{\frac{T_h}{T_{h-1}}} = \sqrt{\frac{h}{h-1}}\,.$$

This means that for large h the obtained term structure gets almost flat at the end, i.e.

$$V(0,T_{h-1}) \approx V(0,T_h) \quad \text{for large } h\,.$$

Thus we can use the formulation (6.5) unless the market term structure is steeply decreasing also for large maturity.

- Formulation (6.7) gives

$$V^2(T_j,T_{h-1}) = s_h^2\,.$$

Since this equality holds for all $j = -1, 0, 1, \ldots, h$ and the right member is independent of j, the volatility term structure evolves simply by cutting off the head. In particular, we move from the term structure

$$\{(T_{j+1}, s_{j+2})), (T_{j+2}, s_{j+3}), \ldots, (T_{M-1}, s_M)\} \quad \text{at time } T_j$$

to

$$\{(T_{j+2}, s_{j+3})), (T_{j+3}, s_{j+4}), \ldots, (T_{M-1}, s_M)\} \quad \text{at time } T_{j+1}\,.$$

An example of this kind of evolution is shown in Figure 6.3, starting from the initial structure of Figure 6.1.2.
This behaviour is not desirable, because even if the term structure today features a hump around two years, this hump is disappearing in the structure in three years.

- Formulation (6.9) gives

$$V^2(T_j,T_{h-1}) = \frac{\Phi_h^2}{\tau_{j,h-1}} \sum_{k=j+2}^{h} \tau_{k-2,k-1}\Psi_k^2\,.$$

In this case it is hard to control the qualitative behaviour of the future term structure of volatilities, because it depends on the particular specification of both the Φ's and the Ψ's.

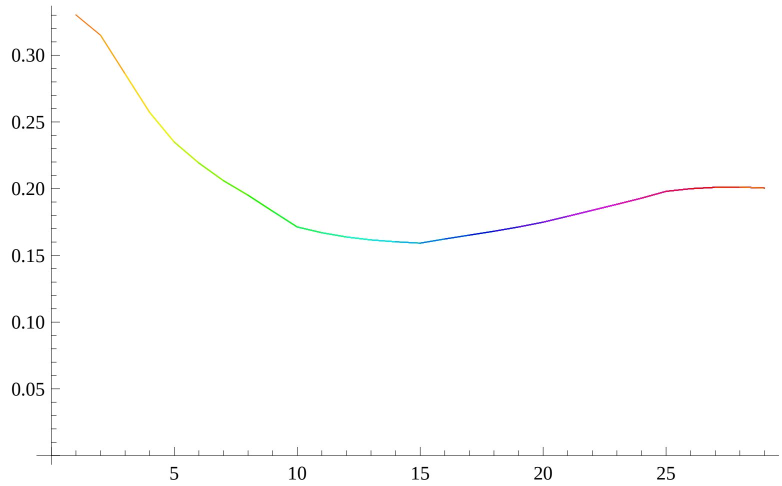

Figure 6.3: Evolution of the term structure of volatility of Figure 6.1.2 obtained with parametrization (6.7).

- Formulation (6.11) gives

$$V^2(T_j, T_{h-1}) = \frac{\Phi_h^2}{\tau_{j,h-1}} \sum_{k=j+1}^{h-1} \tau_{k-1,k} \Psi_{h-k}^2 \, .$$

By assuming again $\tau_{h-1,h} = \tau$ for all $h = 0, \dots, M-1$, it reduces to

$$V^2(T_j, T_{h-1}) = \frac{\Phi_h^2}{(h-1-j)} \sum_{k=j+1}^{h-1} \Psi_{h-k}^2 \, .$$

Then, if the Φ's are all equal, this formulation is analogous to the (6.5). Therefore, keeping the Φ's sufficiently close to each other, the qualitative behaviour of the future term structure of volatilities will not be affected and the hump remains unchanged. This formulation is considered the best among the piecewise-constant ones thanks to the abundance of parameters together with the controllability of the future evolution of the volatility structure.

- Formulation (6.13) gives

$$V^2(T_j, T_{h-1}) = \frac{1}{\tau_{j,h-1}} \int_{T_j}^{T_{h-1}} \Psi(T_{h-1} - t; a, b, c, d)^2 dt \, . \, .$$

Being the analogous continuous version of the parametrization (6.5), it maintains the shape of volatility structure as time passes, in particular its hump if initially present. Nevertheless, once again, it can't be used in the calibration if the initial structure is decreasing for large maturities.

- Formulation (6.13) gives

$$V^2(T_j, T_{h-1}) = \frac{\Phi_h}{\tau_{j,h-1}} \int_{T_j}^{T_{h-1}} \Psi(T_{h-1} - t; a, b, c, d)^2 dt\,..$$

Being the analogous continuous version of the parametrization (6.11), if all the Φ's are all close to each other the term structure maintains its humped as time passes, if initially humped.

Chapter 7

Calibration of the LMM to Swaptions

Since swaption prices are quoted in the market, calibrating the LIBOR market model to swaptions means reducing the distance between the market quotes and the prices obtained in the LMM, by working on the model parameters. But it is not all. Indeed, there are two points of interest: the computational cost and the financial plausibility. On the one hand, the step of calibration comes before anything else, hence it cannot be a slow procedure. On the other hand, in case we have to price other products that are very distant from those to which we have calibrated the model, we may find quite different values depending on which model's parametrization we have chosen.

In the LMM framework, the free parameters are those deriving from the instantaneous correlation and the volatility parameterizations.

Generally, traders translate swaption prices into implied Black's swaption volatilities and organize them in a table where the rows are indexed by the maturity time and the columns are indexed by the length of the underlying swap. An example of such a table of swaption volatilities is shown in Figure 7.1.
We must be careful to the shift of indexes of the resettlement dates of the

	1y	2y	3y	4y	5y	7y	10y
1y	0.327	0.291	0.273	0.26	0.251	0.236	0.221
2y	0.313	0.272	0.256	0.247	0.241	0.23	0.216
3y	0.284	0.253	0.24	0.232	0.227	0.219	0.207
4y	0.259	0.234	0.225	0.218	0.213	0.206	0.198
5y	0.238	0.22	0.212	0.206	0.201	0.195	0.19
7y	0.207	0.197	0.191	0.187	0.183	0.178	0.178
10y	0.175	0.171	0.167	0.165	0.164	0.164	0.168
15y	0.16	0.164	0.165	0.167	0.168	0.172	0.178
20y	0.18	0.186	0.188	0.191	0.193	0.198	0.201

Figure 7.1: Implied volatilities obtained by inverting the Black's formula for swaption, with swaption prices from the Euro market, May 4, 2011.

model $T_0, \dots, T_M$ with respect to the corresponding times in years. For instance, the first row in the above figure is related to $1\,\text{year} = T_0$, the second row to $2\,\text{years} = T_1$, and so on.

Furthermore, a thing to be extremely careful is the problem of "temporal misalignments" in the swaption matrix, in the sense that it is not necessarily uniformly updated. Indeed, although it is not stated, generally the most liquid swaptions are updated regularly, whereas other entries of the matrix refer to older market situations. This fact, together with the problem coming from the missing data, can cause troubles in the calibration.

A calibration of the LMM exclusively to swaptions aims to incorporate as much information as possible from the table of implied volatilities in the model parameters. In this case we can choose any of the methods we have seen in Chapter 4 to compute approximated swaption prices, then apply it to price all the swaptions present in the market table and finally find the parameters that minimize the distance between the corresponding prices. Instead, if we are going to calibrate the LMM both to caplets and to swaptions, we will proceed analogously but carrying out an optimization on less (or neither) parameters, because we can recover one of them in terms of the others from each implied caplet volatility, as we have seen in the different cases of Subsection 6.1.1. Anyway, we would get a good compromise between fitting (low errors) and gumption.

At this point, we have two possible ways of acting: to consider the instantaneous correlation either as input, estimated exogenously and introduced in the calibration leaving free only the volatility parameters, or output, also considered as fitting parameter, of the calibration.

7.1 Historical instantaneous correlation

The opportunity to introduce in the calibration a correlation matrix given exogenously allows us to incorporate the behaviour of the real market rates in our model and to unburden the optimization procedure. On the other hand, however, historical estimations reflect some problems deriving from the data sampling, e.g. outliers and non synchronous data. A way to tackle this undesirable feature is to use parameterized correlation matrices approximating the one estimated and preserving by construction the desired characteristics already described at the beginning of Chapter 5. In support of this, European swaption prices are relatively insensitive to correlation details and a more regular correlation structure can lead, through calibration, to more regular volatilities and to a more stable evolution of the volatility term structure.

7.1.1 Historical estimation

Initially, we have to recover market quotations for interest rates, generally forward LIBOR or forward swap rates, over time. They are characterized by a fixed time to maturity, contrarily to the forward rates modeled in the LMM, which have a fixed maturity date.
Therefore, we must compute first the zero coupon bond prices from the market rates, that will be in the form

$$p(t_1, t_1 + s_i)\,,\, p(t_2, t_2 + s_i)\,,\, \dots\,,\, p(t_l, t_l + s_i) \tag{7.1}$$

for every time to maturity s_i and time t_j of quotations, always expressed in years. In particular, we consider daily quotations and, since the first forward rate modeled expires in T_0, i.e. in one year from today, we go back only one year with the data. Then, we carry out a linear interpolation between the logarithms of the bond prices in (7.1) as functions of the maturity for each fixed evaluation time, in order to compute then the bond prices related to the fixed maturities $T_0, \ldots, T_M$, constituting the tenor structure of our LMM, and valued on the dates of the quotations, i.e. sequences

$$p(t_1, T_i)\,,\; p(t_2, T_i)\,, \ldots\,,\; p(t_l, T_i) \tag{7.2}$$

for every $i = 0, \ldots, M-1$. From these, we get the annual forward rates $F_0, \ldots, F_{M-1}$ valued daily from one year to date.
Notice that we loss the last forward rate F_M in the historical estimation.

At this point, based on the Gaussian approximation

$$\begin{pmatrix} \ln\left(\frac{F_1(t_{i+1})}{F_1(t_i)}\right) & \ldots & \ln\left(\frac{F_{M-1}(t_{i+1})}{F_{M-1}(t_i)}\right) \end{pmatrix} \sim \mathcal{N}(\mu, V)\,,$$

we use the following estimators for the mean and variance,

$$\begin{aligned} \hat{\mu}_i &:= \frac{1}{l-1}\sum_{k=1}^{l-1} \ln\left(\frac{F_i(t_{k+1})}{F_i(t_k)}\right)\,, \\ \hat{V}_{i,j} &:= \frac{1}{l-1}\sum_{k=1}^{l-1} \left(\ln\left(\frac{F_i(t_{k+1})}{F_i(t_k)}\right) - \hat{\mu}_i\right)\left(\ln\left(\frac{F_j(t_{k+1})}{F_j(t_k)}\right) - \hat{\mu}_j\right)\,, \end{aligned} \tag{7.3}$$

where l is the number of past evaluation times, so that our estimation of the historical correlation matrix ρ has elements

$$\hat{\rho}_{i,j} = \frac{\hat{V}_{i,j}}{\sqrt{\hat{V}_i}\sqrt{\hat{V}_j}}\,.$$

Example of Historical estimation

Below, are shown the results obtained by historical data spanning the year before March 29, 2001. In particular, we start with a table of rates quoted at times $t_1 = -1y, \ldots, t_l = 0$, $l = 260$ (daily quotations apart from holidays), specifically the EURIBOR rates for expires in one year and the forward swap rates for tenors of $2, \ldots, 20$ years from $t = 0$. Then, we compute the bond prices by performing a bootstrap and, through a log-linear interpolation, recover the same prices in the form of (7.1), with the annualized tenor structure $\{T_0 = 1y, \ldots, T_{19} = 20y\}$. From these, we recover the forward rates $F_1^*(t_i), \ldots, F_{18}^*(t_i)$ for $i = 1, \ldots, l$, and get the estimate, following (7.3)-(7.3), for the correlations between them, shown in Figures 7.2 and 7.3.

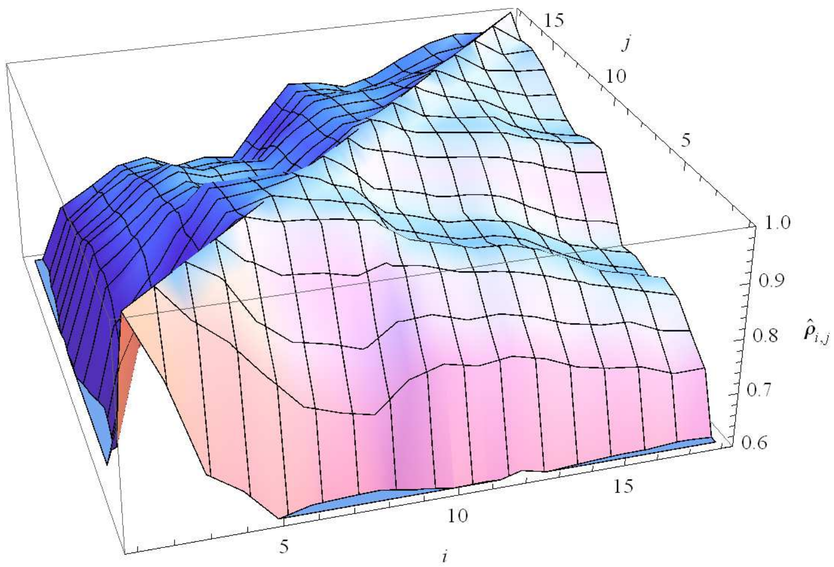

Figure 7.2: Three-dimensional plot of correlations $\hat{\rho}_{i,j}$ from the estimated matrix in Figure 7.3.

We must notice that the resulting $\hat{\rho}$ has the characteristics of a correlation matrix but it does not satisfy the financial expected properties I, II and III.

As mentioned above, once we get the historical estimate, we could consider a parametric correlation matrix that well approximates that estimate. To do this, we have two possibilities:

- choose any of the parameterizations in Chapter 5 and minimize some loss function of the distance between the parametric form and the estimate $\hat{\rho}$;

- directly estimate the correlations in the Schoenmakers and Coffey's semi-parametric structure (5.1)-(5.2).

	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	F_{14}	F_{15}	F_{16}	F_{17}	F_{18}
F_1	1.00	0.88	0.70	0.66	0.60	0.51	0.50	0.49	0.55	0.60	0.59	0.61	0.60	0.58	0.57	0.56	0.55	0.56
F_2	0.88	1.00	0.93	0.88	0.81	0.75	0.73	0.71	0.77	0.79	0.78	0.78	0.77	0.74	0.73	0.72	0.70	0.70
F_3	0.70	0.93	1.0	0.96	0.87	0.82	0.81	0.79	0.84	0.84	0.82	0.82	0.81	0.79	0.78	0.77	0.75	0.74
F_4	0.66	0.88	0.96	1.00	0.94	0.91	0.90	0.89	0.91	0.90	0.88	0.87	0.86	0.84	0.83	0.81	0.79	0.78
F_5	0.60	0.81	0.87	0.94	1.00	0.98	0.98	0.97	0.96	0.92	0.91	0.89	0.89	0.87	0.85	0.82	0.80	0.79
F_6	0.51	0.75	0.82	0.91	0.98	1.00	0.99	0.98	0.96	0.91	0.89	0.87	0.87	0.85	0.82	0.80	0.78	0.76
F_7	0.50	0.73	0.81	0.90	0.98	0.99	1.0	0.98	0.97	0.91	0.90	0.88	0.88	0.86	0.83	0.80	0.78	0.77
F_8	0.49	0.71	0.79	0.89	0.97	0.98	0.98	1.0	0.97	0.91	0.90	0.88	0.87	0.86	0.83	0.80	0.78	0.76
F_9	0.55	0.77	0.84	0.91	0.96	0.96	0.97	0.97	1.00	0.95	0.94	0.93	0.92	0.91	0.89	0.86	0.84	0.84
F_{10}	0.60	0.79	0.84	0.90	0.92	0.91	0.91	0.91	0.95	1.00	0.98	0.96	0.95	0.95	0.94	0.93	0.91	0.90
F_{11}	0.59	0.78	0.82	0.88	0.91	0.89	0.90	0.90	0.94	0.98	1.0	0.97	0.96	0.95	0.95	0.92	0.91	0.90
F_{12}	0.61	0.78	0.82	0.87	0.89	0.87	0.88	0.88	0.93	0.96	0.97	1.00	0.97	0.95	0.95	0.92	0.90	0.88
F_{13}	0.60	0.77	0.81	0.86	0.89	0.87	0.88	0.87	0.92	0.95	0.96	0.97	1.00	0.98	0.96	0.93	0.91	0.90
F_{14}	0.58	0.74	0.79	0.84	0.87	0.85	0.86	0.86	0.91	0.95	0.95	0.95	0.98	1.0	0.97	0.95	0.93	0.93
F_{15}	0.57	0.73	0.78	0.83	0.85	0.82	0.83	0.83	0.89	0.94	0.95	0.95	0.96	0.97	1.0	0.96	0.96	0.95
F_{16}	0.56	0.72	0.77	0.81	0.82	0.80	0.80	0.80	0.86	0.93	0.92	0.92	0.93	0.95	0.96	1.00	0.96	0.97
F_{17}	0.55	0.70	0.75	0.79	0.80	0.78	0.78	0.78	0.84	0.91	0.91	0.90	0.91	0.93	0.96	0.96	1.00	0.96
F_{18}	0.56	0.70	0.74	0.78	0.79	0.76	0.77	0.76	0.84	0.90	0.90	0.88	0.90	0.93	0.95	0.97	0.96	1.00

Figure 7.3: Estimated correlation matrix $\hat{\rho}$ obtained by historical data from the Euro market, March 29, 2011.

Example of Historically Optimized Rebonato's correlation

For example, by performing in Matlab an optimization on the fitting parameters of Rebonato's form in (5.21), i.e.

$$\rho_{i,j}^{\text{Reb}} = \rho_\infty + (1 - \rho_\infty) \exp\left[-|i - j| \left(\beta - \alpha \max\{i, j\}\right)\right] ,$$

wanting to approximate the historical estimate, we have obtained the following values:

$$\alpha = 2.249 * 10^{-14} , \ \beta = 0.0068 , \ \rho_\infty = 0.146 ,$$

which give the Rebonato's correlations in Figures 7.4 and 7.5, respectively plotted in the three-dimensional space and in the matrix form.

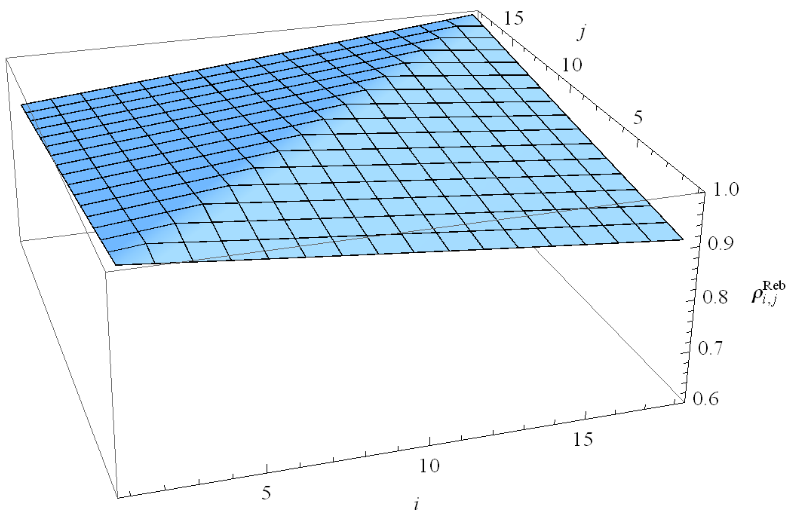

Figure 7.4: Three-dimensional plot of correlations $\rho_{i,j}^{\text{Reb}}(\alpha, \beta, \rho_{inf})$ with the values above for the parameters.

	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	F_{14}	F_{15}	F_{16}	F_{17}	F_{18}
F_1	1.00	0.99	0.99	0.98	0.98	0.97	0.97	0.96	0.95	0.95	0.94	0.94	0.93	0.93	0.92	0.92	0.91	0.91
F_2	0.99	1.00	0.99	0.99	0.98	0.98	0.97	0.97	0.96	0.95	0.95	0.94	0.94	0.93	0.93	0.92	0.92	0.91
F_3	0.99	0.99	1.00	0.99	0.99	0.98	0.98	0.97	0.97	0.96	0.95	0.95	0.94	0.94	0.93	0.93	0.92	0.92
F_4	0.98	0.99	0.99	1.00	0.99	0.99	0.98	0.98	0.97	0.97	0.96	0.95	0.95	0.94	0.94	0.93	0.93	0.92
F_5	0.98	0.98	0.99	0.99	1.00	0.99	0.99	0.98	0.98	0.97	0.97	0.96	0.95	0.95	0.94	0.94	0.93	0.93
F_6	0.97	0.98	0.98	0.99	0.99	1.00	0.99	0.99	0.98	0.98	0.97	0.97	0.96	0.95	0.95	0.94	0.94	0.93
F_7	0.97	0.97	0.98	0.98	0.99	0.99	1.00	0.99	0.99	0.98	0.98	0.97	0.97	0.96	0.95	0.95	0.94	0.94
F_8	0.96	0.97	0.97	0.98	0.98	0.99	0.99	1.00	0.99	0.99	0.98	0.98	0.97	0.97	0.96	0.95	0.95	0.94
F_9	0.95	0.96	0.97	0.97	0.98	0.98	0.99	0.99	1.00	0.99	0.99	0.98	0.98	0.97	0.97	0.96	0.95	0.95
F_{10}	0.95	0.95	0.96	0.97	0.97	0.98	0.98	0.99	0.99	1.00	0.99	0.99	0.98	0.98	0.97	0.97	0.96	0.95
F_{11}	0.94	0.95	0.95	0.96	0.97	0.97	0.98	0.98	0.99	0.99	1.00	0.99	0.99	0.98	0.98	0.97	0.97	0.96
F_{12}	0.94	0.94	0.95	0.95	0.96	0.97	0.97	0.98	0.98	0.99	0.99	1.00	0.99	0.99	0.98	0.98	0.97	0.97
F_{13}	0.93	0.94	0.94	0.95	0.95	0.96	0.97	0.97	0.98	0.98	0.99	0.99	1.00	0.99	0.99	0.98	0.98	0.97
F_{14}	0.93	0.93	0.94	0.94	0.95	0.95	0.96	0.97	0.97	0.98	0.98	0.99	0.99	1.00	0.99	0.99	0.98	0.98
F_{15}	0.92	0.93	0.93	0.94	0.94	0.95	0.95	0.96	0.97	0.97	0.98	0.98	0.99	0.99	1.00	0.99	0.99	0.98
F_{16}	0.92	0.92	0.93	0.93	0.94	0.94	0.95	0.95	0.96	0.97	0.97	0.98	0.98	0.99	0.99	1.00	0.99	0.99
F_{17}	0.91	0.92	0.92	0.93	0.93	0.94	0.94	0.95	0.95	0.96	0.97	0.97	0.98	0.98	0.99	0.99	1.00	0.99
F_{18}	0.91	0.91	0.92	0.92	0.93	0.93	0.94	0.94	0.95	0.95	0.96	0.97	0.97	0.98	0.98	0.99	0.99	1.00

Figure 7.5: Rebonato's three-parameter structure approximating the historically estimation $\hat{\rho}$ found in Figure 7.3.

Example of Historically Estimated S. & C.'s semi-parametric correlation

We start again from historical data spanning the past year before March 29, 2011, and we consider the table, obtained before, of historical forward rates $F_1^*(t_i), \dots, F_{18}^*(t_i)$ for $i = 1, \dots, l$.
Now, for each time $t_i,\ i = 1, \dots, l$, we consider the r.v. defined by

$$y_j^*(t_i) := \ln \frac{F_j^*(t_{i+1})}{F_j^*(t_i)}\,, \quad \text{for } j = 1, \dots, m\,, \quad m = 18\,.$$

Then, we assume the vector $y^*(t_i)$, having values in $\mathbb{R}^m$, following the model (5.3), that we rewrite in the following form:

$$y = A \cdot W = L \cdot \mathrm{diag}(\underline{a}) \cdot W \ \sim \mathcal{N}(0, A\,A')\,,$$

where

$$A = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ a_1 & a_2 & 0\cdots & 0 \\ & & & \vdots \\ \vdots & \vdots & \ddots & 0 \\ a_1 & a_2 & \cdots & a_m \end{pmatrix}, \quad L = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & 0\cdots & 0 \\ & & & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 1 & 1 & \cdots & 1 \end{pmatrix}, \quad \underline{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}.$$

By defining an auxiliary vector

$$w = \mathrm{diag}(\underline{a}) \cdot W \ \sim \mathcal{N}\left(0, \mathrm{diag}\left(\underline{a}^2\right)\right)\,,$$

where the square acts componentwise, we have

$$w = L^{-1} \cdot y\,.$$

Since we have an historical set of vectors $y^*(t_i)$, for $i = 1, \dots, l$, which we consider a sample of realizations of y, we obtain a consequent sample of realizations of w. Then, we compute the sample covariance matrix of the vector w, say $\hat{C}$, being an estimate of the actual one $\mathrm{diag}\,(\underline{a}^2)$, thus we estimate the vector $\underline{a}^2$ by extracting the diagonal from $\hat{C}$, say $\hat{\underline{a}}^2$. Finally, we get the vector

$$c = \sqrt{L \cdot \hat{\underline{a}}^2}\,,$$

whose elements $c_1, \dots c_m$ satisfy

$$c_i^2 = \sum_{k=1}^{i} \hat{a}_k^2 \,, \quad i = 1, \dots, m$$

and we obtain the correlation matrix $\rho^{\mathrm{SC}}(c)$ of kind (5.1)-(5.2), given by

$$\rho^{\mathrm{SC}}(c)_{ij} = \frac{\min\{c_i, c_j\}}{\max\{c_i, c_j\}}, \quad i, j = 1, \dots, m \,.$$

This historically estimated semi-parametric correlation matrix is shown in Figures 7.6 and 7.7, respectively plotted in the three-dimensional space and in the matrix form.

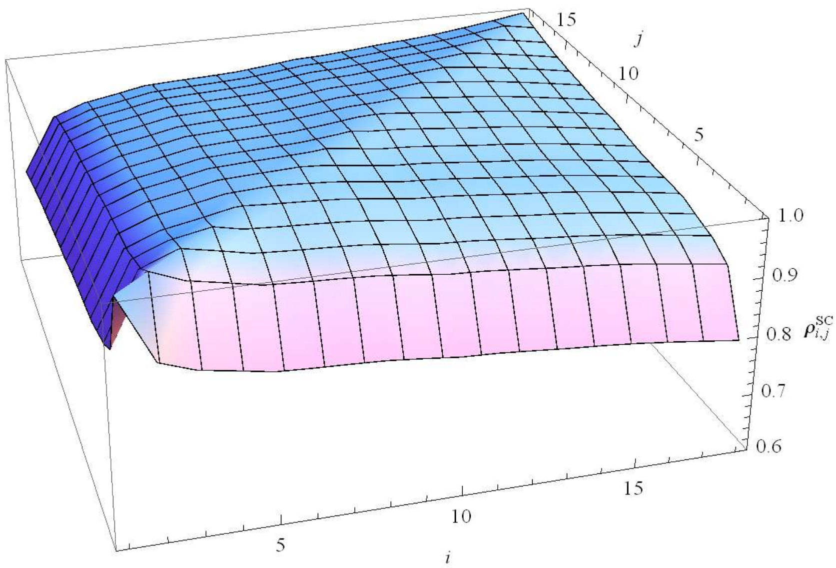

Figure 7.6: Three-dimensional plot of correlations $\rho_{i,j}^{\mathrm{SC}}$.

	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	F_{14}	F_{15}	F_{16}	F_{17}	F_{18}
F_1	1.00	0.89	0.87	0.86	0.85	0.84	0.84	0.84	0.83	0.83	0.82	0.82	0.82	0.81	0.81	0.80	0.79	0.79
F_2	0.89	1.00	0.98	0.97	0.95	0.95	0.95	0.95	0.94	0.93	0.93	0.92	0.92	0.92	0.91	0.90	0.90	0.89
F_3	0.87	0.98	1.00	0.99	0.98	0.97	0.97	0.97	0.96	0.95	0.95	0.95	0.94	0.94	0.93	0.93	0.92	0.91
F_4	0.86	0.97	0.99	1.00	0.99	0.98	0.98	0.98	0.97	0.97	0.96	0.96	0.95	0.95	0.94	0.94	0.93	0.92
F_5	0.85	0.95	0.98	0.99	1.00	1.0	0.99	0.99	0.98	0.98	0.97	0.97	0.96	0.96	0.95	0.95	0.94	0.93
F_6	0.84	0.95	0.97	0.98	1.0	1.00	1.0	0.99	0.99	0.98	0.98	0.97	0.97	0.96	0.96	0.95	0.94	0.93
F_7	0.84	0.95	0.97	0.98	0.99	1.0	1.00	1.0	0.99	0.98	0.98	0.98	0.97	0.97	0.96	0.95	0.94	0.94
F_8	0.84	0.95	0.97	0.98	0.99	0.99	1.0	1.00	0.99	0.99	0.98	0.98	0.97	0.97	0.96	0.96	0.95	0.94
F_9	0.83	0.94	0.96	0.97	0.98	0.99	0.99	0.99	1.00	0.99	0.99	0.98	0.98	0.97	0.97	0.96	0.95	0.94
F_{10}	0.83	0.93	0.95	0.97	0.98	0.98	0.98	0.99	0.99	1.00	1.0	0.99	0.99	0.98	0.98	0.97	0.96	0.95
F_{11}	0.82	0.93	0.95	0.96	0.97	0.98	0.98	0.98	0.99	1.0	1.00	0.99	0.99	0.98	0.98	0.97	0.96	0.95
F_{12}	0.82	0.92	0.95	0.96	0.97	0.97	0.98	0.98	0.98	0.99	0.99	1.00	0.99	0.99	0.98	0.98	0.97	0.96
F_{13}	0.82	0.92	0.94	0.95	0.96	0.97	0.97	0.97	0.98	0.99	0.99	0.99	1.00	1.0	0.99	0.98	0.97	0.96
F_{14}	0.81	0.92	0.94	0.95	0.96	0.96	0.97	0.97	0.97	0.98	0.98	0.99	1.0	1.0	0.99	0.99	0.98	0.97
F_{15}	0.81	0.91	0.93	0.94	0.95	0.96	0.96	0.96	0.97	0.98	0.98	0.98	0.99	0.99	1.00	0.99	0.98	0.97
F_{16}	0.80	0.90	0.93	0.94	0.95	0.95	0.95	0.96	0.96	0.97	0.97	0.98	0.98	0.99	0.99	1.00	0.99	0.98
F_{17}	0.79	0.90	0.92	0.93	0.94	0.94	0.94	0.95	0.95	0.96	0.96	0.97	0.97	0.98	0.98	0.99	1.0	0.99
F_{18}	0.79	0.89	0.91	0.92	0.93	0.93	0.94	0.94	0.94	0.95	0.95	0.96	0.96	0.97	0.97	0.98	0.99	1.00

Figure 7.7: Schoenmakers and Coffey's semi-parametric correlation structure estimated by historical data from the Euro market, March 29, 2011.

Example of Historically Optimized S. & C.'s three-parameter correlation

By performing in Matlab an optimization on the fitting parameters of Schoenmakers and Coffey's three-parameter form in (5.17), i.e.

$$\rho_{i,j}^{\text{SCpar}} = \exp\left[-\frac{|i-j|}{m-1}\left(-\ln\rho_\infty + \eta_1\frac{i^2+j^2+ij-3mi-3mj+3i+3j+2m^2-m-4}{(m-2)(m-3)} + \right.\right.$$
$$\left.\left. -\eta_2\frac{i^2+j^2+ij-mi-mj-3i-3j+3m+2}{(m-2)(m-3)}\right)\right],$$

wanting to approximate the historical estimate $\hat{\rho}$, we have obtained the following values:

$$\eta_1 = 0.4856\,,\ \eta_2 = 0.00\,,\ \ln\rho_\infty = -0.5395\,,$$

which give the correlations in Figures 7.4 and 7.5, respectively plotted in the three-dimensional space and in the matrix form.

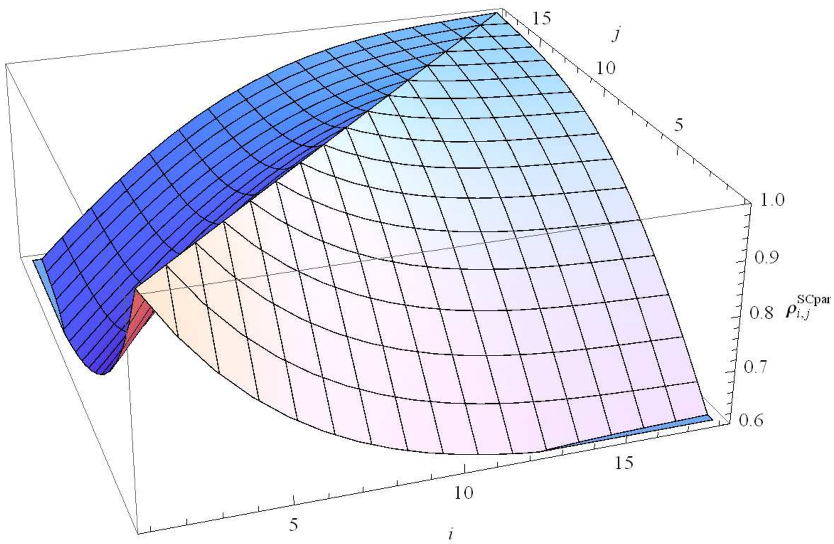

Figure 7.8: Three-dimensional plot of correlations $\rho_{i,j}^{\text{SCpar}}$.

	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	F_{14}	F_{15}	F_{16}	F_{17}	F_{18}
F_1	1.00	0.91	0.85	0.79	0.75	0.71	0.68	0.66	0.64	0.62	0.61	0.60	0.60	0.59	0.59	0.59	0.58	0.58
F_2	0.91	1.00	0.92	0.86	0.81	0.77	0.74	0.72	0.70	0.68	0.67	0.66	0.65	0.65	0.64	0.64	0.64	0.64
F_3	0.85	0.92	1.00	0.93	0.88	0.84	0.80	0.77	0.75	0.74	0.72	0.71	0.71	0.70	0.70	0.69	0.69	0.69
F_4	0.79	0.86	0.93	1.00	0.94	0.90	0.86	0.83	0.81	0.79	0.77	0.76	0.75	0.75	0.75	0.74	0.74	0.74
F_5	0.75	0.81	0.88	0.94	1.00	0.95	0.91	0.88	0.85	0.84	0.82	0.81	0.80	0.79	0.79	0.79	0.78	0.78
F_6	0.71	0.77	0.84	0.90	0.95	1.00	0.96	0.93	0.90	0.88	0.86	0.85	0.84	0.84	0.83	0.83	0.83	0.82
F_7	0.68	0.74	0.80	0.86	0.91	0.96	1.00	0.97	0.94	0.92	0.90	0.89	0.88	0.87	0.87	0.86	0.86	0.86
F_8	0.66	0.72	0.77	0.83	0.88	0.93	0.97	1.00	0.97	0.95	0.93	0.92	0.91	0.90	0.90	0.89	0.89	0.89
F_9	0.64	0.70	0.75	0.81	0.85	0.90	0.94	0.97	1.00	0.98	0.96	0.95	0.94	0.93	0.92	0.92	0.92	0.92
F_{10}	0.62	0.68	0.74	0.79	0.84	0.88	0.92	0.95	0.98	1.00	0.98	0.97	0.96	0.95	0.95	0.94	0.94	0.94
F_{11}	0.61	0.67	0.72	0.77	0.82	0.86	0.90	0.93	0.96	0.98	1.00	0.99	0.98	0.97	0.96	0.96	0.96	0.95
F_{12}	0.60	0.66	0.71	0.76	0.81	0.85	0.89	0.92	0.95	0.97	0.99	1.00	0.99	0.98	0.98	0.97	0.97	0.97
F_{13}	0.60	0.65	0.71	0.75	0.80	0.84	0.88	0.91	0.94	0.96	0.98	0.99	1.00	0.99	0.99	0.98	0.98	0.98
F_{14}	0.59	0.65	0.70	0.75	0.79	0.84	0.87	0.90	0.93	0.95	0.97	0.98	0.99	1.00	0.99	0.99	0.99	0.98
F_{15}	0.59	0.64	0.70	0.75	0.79	0.83	0.87	0.90	0.92	0.95	0.96	0.98	0.99	0.99	1.00	1.0	0.99	0.99
F_{16}	0.59	0.64	0.69	0.74	0.79	0.83	0.86	0.89	0.92	0.94	0.96	0.97	0.98	0.99	1.0	1.00	1.0	0.99
F_{17}	0.58	0.64	0.69	0.74	0.78	0.83	0.86	0.89	0.92	0.94	0.96	0.97	0.98	0.99	0.99	1.0	1.00	1.0
F_{18}	0.58	0.64	0.69	0.74	0.78	0.82	0.86	0.89	0.92	0.94	0.95	0.97	0.98	0.98	0.99	0.99	1.0	1.00

Figure 7.9: Schoenmakers and Coffey's three-parameter structure approximating the historically estimation $\hat{\rho}$ found in Figure 7.3.

We denote the average squared relative error by MSE% and the average simple relative error by ME%. In Table 7.1 we compare the errors characterizing the three correlation get above, the Rebonato's and Schoenmakers and Coffey's optimized three-parameter and the estimated Schoenmakers and Coffey's semi-parametric structure, in terms of the historical estimation.

	ME%	MSE%
Reb. three-par.	0.159833	0.246294
S. & C. semi-par.	0.136001	0.189073
S. & C. three-par.	0.0622644	0.0887776

Table 7.1: Average relative errors, both simple and squared, between historical estimation in Figures 7.3-7.2 and, respectively, the ones in Figures 7.5-7.4, 7.7-7.6 and 7.9-7.8.

7.2 Cascade Calibration

In Subsection 6.1.1 we presented some possible parameterizations of the forward volatilities, that largely refer to the most general one, the GPC formulation (6.2) shown in Table 6.1. This structure has a high number of parameters, thus we then made some assumptions on it in order to reduce the number of free parameters to be involved in an optimization algorithm of the calibration. However, there exists an alternative method such that, given exogenously the correlations, the calibration of the LMM to the swaptions can be carried out through closed-form formulae. Substantially, we would like to have a calibration that is:

- univocal, i.e. without all the indeterminacies seen in the calibration to the caplets;

- exact, i.e. that avoids the problem of significant errors with respect to the table of the market swaption volatilities;

- computationally efficient.

This method meets in a large part our expectations.

We assume the GPC volatility structure and we start from an historical correlation matrix ρ and with the table of the swaption volatilities from the market, denoting by $V_{\alpha,\beta}$ the Black volatility for the swaption with underlying swap rate $S_{\alpha,\beta}$.

By recalling the Rebonato's formula for swaptions, the approximated Black volatility in the LMM is

$$(v_{\alpha,\beta}^{\mathrm{LMM}})^2 := \sum_{i,j=\alpha+1}^{\beta} \frac{w_i(0)w_j(0)F_i(0)F_j(0)\rho_{i,j}}{T_\alpha\, S_{\alpha,\beta}(0)^2} \int_0^{T_\alpha} \sigma_i(t)\sigma_j(t)dt\,.$$

We apply this to the GPC formulation and equate it to the market swaption volatility:

$$(V_{\alpha,\beta})^2 \approx \sum_{i,j=\alpha+1}^{\beta} \frac{w_i(0)w_j(0)F_i(0)F_j(0)\rho_{i,j}}{T_\alpha\, S_{\alpha,\beta}(0)^2} \sum_{h=0}^{\alpha} \tau_h \sigma_{i,h+1}\sigma_{,j,h+1}\,, \tag{7.4}$$

where we must remember that the w's depend on the specific α, β considered.

The cascade calibration moves along the swaption table from left to right and from top to down and we apply, at each step, the approximating formula (7.4) to compute a new unknown volatility.

First, we analyze the simplest case of the calibration to the upper-left triangular part of a swaption matrix providing all the market data involved, which leads to sensible results, then the calibration to a rectangular submatrix, again providing all the market volatilities involved. Finally, we consider the extended triangular case in which we calibrate again to the upper-left triangular part of a swaption matrix, where, however there are missing data.

7.2.1 Triangular Cascade Calibration Algorithm

In this case we calibrate the GPC volatility formulation to the upper-left triangular part of a swaption matrix where there are no missing data from the market. To give an overview of this procedure, in the Table 7.2 we show the outcomes obtained by applying it to an example of just six swaptions.

Let's see in detail how to proceed in this example. We start from the entry

Length / Maturity	1year	2years	3years
$T_0 = 1$year	$\mathbf{V_{0,1}}$ $\sigma_{1,1}$	$\mathbf{V_{0,2}}$ $\sigma_{1,1}$ $\sigma_{2,1}$	$\mathbf{V_{0,3}}$ $\sigma_{1,1}$ $\sigma_{2,1}$ $\sigma_{3,1}$
$T_1 = 2$years	$\mathbf{V_{1,2}}$ $\sigma_{2,1}, \sigma_{2,2}$	$\mathbf{V_{1,3}}$ $\sigma_{2,1}, \sigma_{2,2}$ $\sigma_{3,1}, \sigma_{3,2}$	–
$T_2 = 3$years	$\mathbf{V_{2,3}}$ $\sigma_{3,1}, \sigma_{3,2}, \sigma_{3,3}$	–	–

Table 7.2: Table summarizing the swaption volatilities to which we calibrate the LMM through a triangular cascade calibration and the dependence of the GPC forward volatilities on them, where the blue ones are the new parameters determined at each step.

$V_{0,1}$ in position $(1,1)$, i.e. the swaption maturing in T_0 e living up to T_1, that is an option on a single forward LIBOR rate collapsing to a spot LIBOR rate, and the equation to solve has as unique unknown the volatility of F_1 from now to $T_0 = 1$year. Indeed:

$$S_{0,1}(0) = w_1(0)F_1(0) \quad \Rightarrow \quad (V_{0,1})^2 \approx \sigma_{1,1}^2 \, .$$

Hence the parameter $\sigma_{1,1}$ is calibrated exactly.

Then we move to the entry $(1,2)$, $V_{0,2}$, which involves the rates F_1, F_2 over the time from now to a year:

$$\begin{aligned} S_{0,2}(0)^2(V_{0,2})^2 \approx\ & w_1(0)^2F_1(0)^2\sigma_{1,1}^2 + w_2(0)^2F_2(0)^2\sigma_{2,1}^2 + \\ & +2\rho_{1,2}w_1(0)F_1(0)w_2(0)F_2(0)\sigma_{1,1}\sigma_{2,1} \, , \end{aligned}$$

where the only unknown is $\sigma_{2,1}$, which solves an algebraic second-order equation assuming existence and uniqueness of a positive solution.

Moving to the entry $(1,3)$, $V_{0,3}$, the rate F_3 is added among others, still over $[0,T_0]$, and we have

$$\begin{aligned}S_{0,3}(0)^2(V_{0,3})^2 \approx\ & w_1(0)^2F_1(0)^2\sigma_{1,1}^2 + w_2(0)^2F_2(0)^2\sigma_{2,1}^2 + w3(0)^2F_3(0)^2\sigma_{3,1}^2 + \\ & +2\rho_{1,2}w_1(0)F_1(0)w_2(0)F_2(0)\sigma_{1,1}\sigma_{2,1} + \\ & +2\rho_{1,3}w_1(0)F_1(0)w_3(0)F_3(0)\sigma_{1,1}\sigma_{3,1} + \\ & +2\rho_{2,3}w_2(0)F_2(0)w_3(0)F_3(0)\sigma_{2,1}\sigma_{3,1}\,,\end{aligned}$$

where the only unknown is $\sigma_{3,1}$, which solves again a second-order equation, by the same assumption of existence and uniqueness.

Now we move on the second row to the entry $(2,1)$, $V_{1,2}$, where only F_2 is at stake but over the two time subintervals $[0,T_0]$, $[T_0,T_1]$. The formula (7.4) gives

$$T_1(V_{1,2})^2 \approx \tau_0\sigma_{2,1}^2 + \tau_1\sigma_{2,2}^2\,,$$

with the only unknown $\sigma_{2,2}$.

Moving on the right to the entry $(2,2)$, $V_{1,3}$, the formula (7.4) gives

$$\begin{aligned}T_1S_{1,3}(0)^2(V_{1,3})^2 \approx\ & w_2(0)^2F_2(0)^2(\tau_0\sigma_{2,1}^2 + \tau_1\sigma_{2,2}^2) + \\ & +w_3(0)^2F_3(0)^2(\tau_0\sigma_{3,1}^2 + \tau_1\sigma_{3,2}^2) + \\ & +2\rho_{2,3}w_2(0)F_2(0)w_3(0)F_3(0)(\tau_0\sigma_{2,1}\sigma_{3,1} + \tau_1\sigma_{2,2}\sigma_{3,2})\,,\end{aligned}$$

where the only unknown is $\sigma_{3,2}$.

Finally, we move to the only entry $(3,1)$ on the third row, $V_{2,3}$, where our formula gives

$$T_2(V_{2,3})^2 \approx \tau_0\sigma_{3,1}^2 + \tau_1\sigma_{3,2}^2 + \tau_2\sigma_{3,3}^2\,,$$

with the only unknown $\sigma_{3,3}$.

Notice that each time we compute an only new parameter as a function of a market swaption volatility and of the parameters previously found.

The procedure illustrated in the example above can be generalized in the following scheme.

Cascade Calibration Algorithm (CCA)

Brigo and Mercurio (2001,2002)

1. *Select the dimension* s *of the swaption matrix that is of interest for the calibration.*

2. *Set* $\alpha = 0$.

3. *Set* $\beta = \alpha + 1$.

4. *Solve the following equation in* $\sigma_{\beta,\alpha+1}$ *:*

$$A_{\alpha,\beta}\sigma^2_{\beta,\alpha+1} + B_{\alpha,\beta}\sigma_{\beta,\alpha+1} + C_{\alpha,\beta} = 0\,, \tag{7.5}$$

where

$$\begin{aligned}
A_{\alpha,\beta} &= w_\beta(0)^2 F_\beta(0)^2 \tau_\alpha\,,\\
B_{\alpha,\beta} &= 2\sum_{j=\alpha+1}^{\beta-1} w_\beta(0)F_\beta(0)w_j(0)F_j(0)\,\rho_{\beta,j}\,\tau_\alpha\,\sigma_{j,\alpha+1}\,,\\
C_{\alpha,\beta} &= \sum_{i,j=\alpha+1}^{\beta-1} w_i(0)F_i(0)w_j(0)F_j(0)\,\rho_{i,j}\sum_{h=0}^{\alpha}\tau_h\,\sigma_{i,h+1}\,\sigma_{j,h+1} +\\
&\quad +2\sum_{j=\alpha+1}^{\beta-1} w_\beta(0)F_\beta(0)w_j(0)F_j(0)\,\rho_{\beta,j}\sum_{h=0}^{\alpha-1}\tau_h\,\sigma_{\beta,h+1}\,\sigma_{j,h+1} +\\
&\quad +w_\beta(0)^2F_\beta(0)^2\sum_{h=0}^{\alpha-1}\tau_h\,\sigma^2_{\beta,h+1} - T_\alpha\, S_{\alpha,\beta}(0)^2(V_{\alpha,\beta})^2\,.
\end{aligned}$$

Since $A_{\alpha,\beta}, B_{\alpha,\beta} > 0$*, the equation (7.5) admits a unique positive solution, namely*

$$\sigma_{\beta,\alpha+1} = \frac{-B_{\alpha,\beta} + \sqrt{B^2_{\alpha,\beta} - 4A_{\alpha,\beta}C_{\alpha,\beta}}}{2A_{\alpha,\beta}}\,,$$

if and only if $C_{\alpha,\beta} < 0$.

5. *Set* $\beta = \alpha + 1$; *if* $\beta \le s$, *go back to point* 4, *else set* $\alpha = \alpha + 1$.

6. *If* $\alpha < s$, *go back to point* 3, *else stop.*

Practical experiences confirm that for non-pathological swaption data the condition $C_{\alpha,\beta} < 0$ is generally verified. Instead, in problematic situations, we shall make some adjustments to this method, for example as in the extended case. Notice that, in a triangular CCA of dimension s, the entries of the swaption matrix involved are

$$(i, j) \quad \text{s.t.} \quad (i + j) \leq (s + 1)\,,$$

hence we calibrate the model to Black swaption volatilities

$$V_{\alpha,\beta} \quad \text{s.t.} \quad 0 \leq \alpha \leq s - 1, \quad \alpha + 1 \leq \beta \leq s\,.$$

Indeed:

$$\begin{cases} \alpha = i - 1 \\ \beta = \alpha + j = i + j - 1 \end{cases} \iff \begin{cases} i = \alpha + 1 \\ j = \beta - \alpha\,. \end{cases}$$

This kind of calibration does not need further assumptions and it is independent of the dimension s, provided all the market data involved, in the sense that the output of the calibration to a sub-matrix of a swaption table V will be a subset of the output of the calibration to V.

Example of a 5-dimensional Cascade Calibration

We applied the CCA to the 5×5 triangular sub-matrix of the implied Black swaption volatilities shown in Figure 7.10, where we highlight the fact that there are maturities with missing data by showing the corresponding rows and columns empty. We chose $s = 5$, as it is the larger maturity with no missing data before.

Again, we have to pay attention to the shift of the indexes of the resettlement dates of the model with respect to the corresponding times in years.

The resulting calibrated forward volatilities are shown in Figure 7.11.

7.2.2 Rectangular Cascade Calibration Algorithm

If we want to calibrate the LMM to a whole rectangular swaption matrix with all the entries provided by the market, we need to make some

	1 y	2 y	3 y	4 y	5 y	6 y	7 y	8 y	9 y	10 y
1 y	0.34	0.30	0.28	0.27	0.26		0.24			0.22
2 y	0.32	0.28	0.26	0.25	0.24		0.23			0.22
3 y	0.29	0.26	0.24	0.23	0.23		0.22			0.21
4 y	0.26	0.24	0.23	0.22	0.21		0.21			0.20
5 y	0.24	0.22	0.21	0.21	0.20		0.20			0.20
6 y										
7 y	0.21	0.20	0.19	0.19	0.18		0.18			0.18
8 y										
9 y										
10 y	0.18	0.17	0.17	0.17	0.17		0.17			0.17
11 y										
12 y										
13 y										
14 y										
15 y	0.16	0.16	0.17	0.17	0.17		0.17			0.18
16 y										
17 y										
18 y										
19 y										
20 y	0.18	0.18	0.19	0.19	0.19		0.20			0.21

Figure 7.10: Implied Black swaption volatilities, from the Euro market, April 26, 2011.

$$\sigma = \begin{pmatrix} 0.342 & 0 & 0 & 0 & 0 \\ 0.274301 & 0.36172 & 0 & 0 & 0 \\ 0.264304 & 0.245348 & 0.344644 & 0 & 0 \\ 0.252787 & 0.24233 & 0.220029 & 0.311907 & 0 \\ 0.247875 & 0.231487 & 0.20753 & 0.191782 & 0.293202 \end{pmatrix}$$

Figure 7.11: Forward volatilities calibrated to the 5×5 sub-matrix of the swaption table in Figure 7.10.

adjustments to the CCA algorithm, because, in general, there are no market swaption tables large enough to contain our rectangular one in its upper-triangular part.

By recalling the initial example of dimension $s = 3$, in the Table 7.3 we show the outcomes of the rectangular cascade calibration, involving nine swaptions.

We can see how, in each entry of the last column except from the first, we

Length / Maturity	1year	2years	3years
$T_0 = 1$year	$\mathbf{V_{0,1}}$ $\sigma_{1,1}$	$\mathbf{V_{0,2}}$ $\sigma_{1,1}$ $\sigma_{2,1}$	$\mathbf{V_{0,3}}$ $\sigma_{1,1}$ $\sigma_{2,1}$ $\sigma_{3,1}$
$T_1 = 2$years	$\mathbf{V_{1,2}}$ $\sigma_{2,1}, \sigma_{2,2}$	$\mathbf{V_{1,3}}$ $\sigma_{2,1}, \sigma_{2,2}$ $\sigma_{3,1}, \sigma_{3,2}$	$\mathbf{V_{1,4}}$ $\sigma_{2,1}, \sigma_{2,2}$ $\sigma_{3,1}, \sigma_{3,2}$ $\sigma_{4,1}, \sigma_{4,2}$
$T_2 = 3$years	$\mathbf{V_{2,3}}$ $\sigma_{3,1}, \sigma_{3,2}, \sigma_{3,3}$	$\mathbf{V_{2,4}}$ $\sigma_{3,1}, \sigma_{3,2}, \sigma_{3,3}$ $\sigma_{4,1}, \sigma_{4,2}, \sigma_{4,3}$	$\mathbf{V_{2,5}}$ $\sigma_{3,1}, \sigma_{3,2}, \sigma_{3,3}$ $\sigma_{4,1}, \sigma_{4,2}, \sigma_{4,3}$ $\sigma_{5,1}, \sigma_{5,2}, \sigma_{5,3}$

Table 7.3: Table summarizing the swaption volatilities to which we calibrate the LMM through a rectangular cascade calibration and the dependence of the GPC forward volatilities on them, where the blue ones are the new parameters determined at each step.

have multiple unknown forward volatilities. The easiest way to interrelate them is to assume they are all equal and this assumption makes sense, because the multiple unknowns are always volatilities of a single forward LIBOR rate over a few adjacent intervals of time.

This method is described in detail in the following scheme.

Rectangular Cascade Calibration Algorithm (RCCA)

Brigo and Morini (2002), Morini (2002) *The RCCA algorithm recovers the first three points of the CCA algorithm, unchanged. At point* 5 *the new condition for* β *becomes*

$$\beta - \alpha \leq s\,.$$

Moreover, in the case

$$\beta = s + \alpha\,,$$

at point 4 *we have to assume all the unknowns to be equal, i.e.*

$$\sigma_{\beta,1} = \sigma_{\beta,2} = \ldots \sigma_{\beta,\alpha+1}\,.$$

Hence, instead of (7.5), the new equation to solve is

$$A^*_{\alpha,\beta}\sigma^2_{\beta,\alpha+1} + B^*_{\alpha,\beta}\sigma_{\beta,\alpha+1} + C^*_{\alpha,\beta} = 0\,, \tag{7.6}$$

where

$$\begin{aligned}
A^*_{\alpha,\beta} &= w_\beta(0)^2 F_\beta(0)^2 \sum_{h=0}^{\alpha} \tau_h\,,\\
B^*_{\alpha,\beta} &= 2\sum_{j=\alpha+1}^{\beta-1} w_\beta(0)F_\beta(0)w_j(0)F_j(0)\,\rho_{\beta,j}\,\tau_\alpha\,\sigma_{j,\alpha+1} +\\
&\quad +2\sum_{j=\alpha+1}^{\beta-1} w_\beta(0)F_\beta(0)w_j(0)F_j(0)\,\rho_{\beta,j}\sum_{h=0}^{\alpha-1}\tau_h\,\sigma_{j,h+1}\,,\\
C^*_{\alpha,\beta} &= \sum_{i,j=\alpha+1}^{\beta-1} w_i(0)F_i(0)w_j(0)F_j(0)\,\rho_{i,j}\sum_{h=0}^{\alpha}\tau_h\,\sigma_{i,h+1}\,\sigma_{j,h+1} +\\
&\quad -T_\alpha\,S_{\alpha,\beta}(0)^2(V_{\alpha,\beta})^2\,.
\end{aligned}$$

Example of a 5-dimensional Rectangular Cascade Calibration

We applied the RCCA to the whole 5×5 sub-matrix of the implied Black swaption volatilities shown in Figure 7.10 and the resulting calibrated forward volatilities are shown in Figure 7.12.

7.2.3 Extended Triangular Cascade Calibration Algorithm

In case we want to calibrate the model to a larger dimension of the swaption table, we have to tackle the problem of missing data. We refer to the

$$\sigma = \begin{pmatrix} 0.342 & 0 & 0 & 0 & 0 \\ 0.274301 & 0.36172 & 0 & 0 & 0 \\ 0.264304 & 0.245348 & 0.344644 & 0 & 0 \\ 0.252787 & 0.24233 & 0.220029 & 0.311907 & 0 \\ 0.247875 & 0.231487 & 0.20753 & 0.191782 & 0.293202 \\ 0.233045 & 0.233045 & 0.199793 & 0.204828 & 0.177152 \\ 0.220006 & 0.220006 & 0.220006 & 0.177201 & 0.178626 \\ 0.206154 & 0.206154 & 0.206154 & 0.206154 & 0.17995 \\ 0.190429 & 0.190429 & 0.190429 & 0.190429 & 0.190429 \end{pmatrix}$$

Figure 7.12: Forward volatilities calibrated to the 5×5 sub-matrix of the swaption table in Figure 7.10.

algorithm we are describing as to the "ExtCCA Algorithm".

It is essential to observe that the swaption volatility $V_{\alpha,\beta}$, which is located in the $(\alpha - 1, \beta - \alpha)$ entry of the table, involves the forward volatilities

$$\{\sigma_{i,j}\}_{\substack{i=\alpha+1,\dots,\beta \\ j=1,\dots,\alpha+1}} .$$

Among these, as we are in the triangular case, the new unknown is always $\sigma_{\beta,\alpha+1}$.

Whenever the data $V_{\alpha,\beta}$ is not quoted by the market, we compute the unknown forward volatility by means of the following devices:

$$\begin{cases} \sigma_{\beta,\alpha+1} = \sigma_{\beta-1,\alpha+1}, & \text{if } \alpha = 0; \\ \sigma_{\beta,\alpha+1} = \frac{\sigma_{\beta,\alpha}+\sigma_{\beta-1,\alpha}}{2}, & \text{if } \alpha > 0. \end{cases}$$

Then, we create a fictitious data $\tilde{V}_{\alpha,\beta}$, by the approximating formula (7.4), using the σ's already calibrated and the one defined above, in order to fill in the incomplete market table.

Example of a 10 and a 15-dimensional ExtCCA

Firstly, we applied the ExtCCA to the 10×10 triangular sub-matrix of the implied Black swaption volatilities shown in Figure 7.10 and the resulting calibrated forward volatilities are shown in Figure 7.13.

We exploit the calibrated σ's to reconstruct the missing swaption volatilities falling inside our triangular part of interest. Thus, the new obtained swaption table is shown in Figure 7.14.

Then, we applied the ExtCCA to the largest 15×15 triangular matrix deriving from the implied Black swaption volatilities shown in Figure 7.10 and the resulting calibrated forward volatilities are shown in Figure 7.15.

As above, we exploit the calibrated σ's to reconstruct the missing swaption volatilities falling inside our triangular part of interest, thus obtaining the swaption table shown in Figure 7.16.

Conclusions about the Cascade Calibration

We may conclude by pointing out the main features of the cascade calibration. Its positive aspects are:

- it makes use of the correlation matrix in input from a historical estimation;
- it is a fast method, thanks to the analytical closed-form formulae;
- it is exact, i.e. if Rebonato's approximation for the Black implied volatility of swaptions is used, the swaption market prices are fitted exactly;
- given correlation, it has a unique solution, under some homogeneity assumptions;
- it induces a one-to-one relation between the model σ's and the market swaption volatilities.

However, on the other hand, we have encountered a few numerical problems, for example the condition $C_{\alpha,\beta} < 0$ is not always satisfied and some negative or imaginary σ's may come out. In particular, we came up against this facts only in the last example of calibration, by carrying out an ExtCCA with dimension $s = 15$, whereas is a too large because of the considerable amount of missing data. Moreover, the temporal misalignments in the market data provided by a single broker and the correlation coming from a different calibration need more attention.

$$\sigma = \begin{pmatrix}
0.342 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.271535 & 0.363801 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.260861 & 0.244808 & 0.347639 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.249284 & 0.240942 & 0.221234 & 0.314935 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.244662 & 0.230049 & 0.208129 & 0.193539 & 0.295445 & 0 & 0 & 0 & 0 & 0 \\
0.244662 & 0.217784 & 0.199226 & 0.205704 & 0.178541 & 0.236993 & 0 & 0 & 0 & 0 \\
0.196503 & 0.220583 & 0.241022 & 0.180193 & 0.179805 & 0.179173 & 0.240164 & 0 & 0 & 0 \\
0.196503 & 0.273879 & 0.247231 & 0.098049 & 0.168667 & 0.174236 & 0.139744 & 0.189954 & 0 & 0 \\
0.196503 & 0.196503 & 0.207105 & 0.227168 & 0.133219 & 0.150943 & 0.201287 & 0.170515 & 0.180234 & 0 \\
0.23603 & 0.216267 & 0.206385 & 0.163969 & 0.195569 & 0.164394 & 0.0756796 & 0.138483 & 0.154499 & 0.142269
\end{pmatrix}$$

Figure 7.13: Forward volatilities calibrated with a 10-dimensional ExtCCA to the swaption table in Figure 7.10.

	1 y	2 y	3 y	4 y	5 y	6 y	7 y	8 y	9 y	10 y
1 y	0.34	0.30	0.28	0.27	0.26	0.25	0.24	0.23	0.22	0.22
2 y	0.32	0.28	0.26	0.25	0.24	0.23	0.23	0.23	0.22	0.22
3 y	0.29	0.26	0.24	0.23	0.23	0.23	0.22	0.22		0.21
4 y	0.26	0.24	0.23	0.22	0.21	0.21	0.21			0.20
5 y	0.24	0.22	0.21	0.21	0.20	0.20	0.20			0.20
6 y	0.21	0.21	0.20	0.20	0.19					
7 y	0.21	0.20	0.19	0.19	0.18		0.18			0.18
8 y	0.19	0.19	0.18							
9 y	0.19	0.18								
10 y	0.18	0.17	0.17	0.17	0.17		0.17			0.17
11 y										
12 y										
13 y										
14 y										
15 y	0.16	0.16	0.17	0.17	0.17		0.17			0.18
16 y										
17 y										
18 y										
19 y										
20 y	0.18	0.18	0.19	0.19	0.19		0.20			0.21

Figure 7.14: Black swaption volatilities partially reconstructed after a 10-dimensional ExtCCA to the market swaption table in Figure 7.10.

$$\sigma = \begin{pmatrix}
0.342 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.271535 & 0.363801 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.260861 & 0.244808 & 0.347639 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.249284 & 0.240942 & 0.221234 & 0.314935 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.244662 & 0.230049 & 0.208129 & 0.193539 & 0.295445 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.244662 & 0.217784 & 0.199226 & 0.205704 & 0.178541 & 0.236993 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.196503 & 0.220583 & 0.241022 & 0.180193 & 0.179805 & 0.179173 & 0.240164 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.196503 & 0.273879 & 0.247231 & 0.098049 & 0.168667 & 0.174236 & 0.139744 & 0.189954 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.196503 & 0.196503 & 0.207105 & 0.227168 & 0.133219 & 0.150943 & 0.201287 & 0.170515 & 0.180234 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.23603 & 0.216267 & 0.206385 & 0.163969 & 0.195569 & 0.164394 & 0.0756796 & 0.138483 & 0.154499 & 0.142269 & 0 & 0 & 0 & 0 & 0 \\
0.23603 & 0.166229 & 0.191248 & 0.198816 & 0.175855 & 0.185712 & 0.0453731 & 0.0605264 & 0.0995048 & 0.236383 & 0.189326 & 0 & 0 & 0 & 0 \\
0.23603 & 0.23603 & 0.0946263 & 0.142937 & 0.170877 & 0.173366 & 0.179539 & 0.112456 & 0.0864912 & 0.191997 & 0.21419 & 0.201758 & 0 & 0 & 0 \\
0.23603 & 0.23603 & 0.23603 & 0.190344 & 0.166641 & 0.168759 & -0.0114842 & 0.0840273 & 0.0982416 & 0.0947651 & 0.143381 & 0.178785 & 0.190272 & 0 & 0 \\
0.23603 & 0.23603 & 0.23603 & 0.23603 & 0.0610866 & 0.113864 & 0.141311 & 0.0649134 & 0.0744704 & 0.210479 & 0.152622 & 0.148002 & 0.163394 & 0.176833 & 0 \\
0.23603 & 0.23603 & 0.23603 & 0.23603 & 0.23603 & 0.148558 & 0.131211 & 0.136261 & 0.100587 & 0.0875288 & 0.149004 & 0.150813 & 0.149407 & 0.1564 & 0. + 0.248824\, i
\end{pmatrix}$$

Figure 7.15: Forward volatilities calibrated with a 15-dimensional ExtCCA to the swaption table in Figure 7.10.

	1 y	2 y	3 y	4 y	5 y	6 y	7 y	8 y	9 y	10 y	11 y	12 y	13 y	14 y	15 y
1 y	0.34	0.30	0.28	0.27	0.26	0.25	0.24	0.23	0.22	0.22	0.22	0.22	0.22	0.22	0.22
2 y	0.32	0.28	0.26	0.25	0.24	0.23	0.23	0.23	0.22	0.22	0.22	0.22	0.22	0.22	
3 y	0.29	0.26	0.24	0.23	0.23	0.23	0.22	0.22	0.21	0.21	0.21	0.21	0.21		
4 y	0.26	0.24	0.23	0.22	0.21	0.21	0.21	0.21	0.20	0.20	0.20	0.20			
5 y	0.24	0.22	0.21	0.21	0.20	0.20	0.20	0.19	0.20	0.20	0.20				
6 y	0.21	0.21	0.20	0.20	0.19	0.19	0.19	0.19	0.19	0.19					
7 y	0.21	0.20	0.19	0.19	0.18	0.18	0.18	0.18	0.18	0.18					
8 y	0.19	0.19	0.18	0.18	0.17	0.17	0.17	0.18							
9 y	0.19	0.18	0.17	0.17	0.17	0.17	0.17								
10 y	0.18	0.17	0.17	0.17	0.17	0.17	0.17			0.17					
11 y	0.17	0.17	0.17	0.17	0.17										
12 y	0.18	0.17	0.17	0.17											
13 y	0.17	0.17	0.17												
14 y	0.17	0.17													
15 y	0.16	0.16	0.17	0.17	0.17		0.17			0.18					
16 y															
17 y															
18 y															
19 y															
20 y	0.18	0.18	0.19	0.19	0.19		0.20			0.21					

Figure 7.16: Black swaption volatilities partially reconstructed after a 15-dimensional ExtCCA to the market swaption table in Figure 7.10.

Appendix A

Preliminary theory

Definition A.1. We define an N-dimensional *Itô process* any stochastic process $X = (X_t)_{t\in[0,T]}$ whose dynamics is given by

$$dX_t = \mu_t dt + \sigma_t dW_t$$

where W is a d-dimensional standard Brownian motion on $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$, $\mu, \sigma \in \mathbb{L}^2_{loc}(\Omega \times [0,T])$, μ having values in $\mathbb{R}^N$ and σ having values in $\mathbb{R}^{N\times d}$. Equivalently, in the integrated form, for $i = 1, \ldots, N$ the i-th component of X follows:

$$X^i_t = X^i_0 + \int_0^t \mu^i_s ds + \sum_{j=1}^{d} \int_0^t \sigma^{i,j}_s dW^j_s \,.$$

Definition A.2. Given an N-dimensional Itô process X, the associated *co-variation process* is the s.p. $(< X, X >_t)_{t\in[0,T]}$ with values in $\mathbb{R}^{N\times d}$ defined by

$$< X^i, X^j >_t := \lim_{|\Sigma|\to 0} \sum_{k=1}^{n} (X^i_{t_k} - X^i_{t_k-1})(X^j_{t_k} - X^j_{t_k-1}) \,,$$

where $\Sigma = \{(0 = t_0, \ldots, t_n = T) \,|\, t_0 < \ldots < t_n \,,\, n \in \mathbb{N}\}$.

Lemma A.0.1. *With the notations above, let $C_t = \sigma_t \sigma^*_t$, we have*

$$< X^i, X^j >_t = \int_0^t C^{i,j}_s ds \,, \ \text{or equivalently} \quad d < X^i, X^j >_t = C^{i,j}_t dt \,.$$

We use also the notation $\quad d < X^i, X^j >_t \equiv (dX^i)\,(dX^j)$.

Lemma A.0.2 (Itô's formula). *Given an N-dimensional Itô process with dynamics*

$$dX_t = \mu_t dt + \sigma_t dW_t$$

and a function $F = F(t,x) \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}^N)$, then the s.p. Y defined by $Y_t = F(t, X_t)$ is an Itô process with dynamics

$$dY_t = \partial_t F\, dt + \nabla F \cdot dX_t + \frac{1}{2} \sum_{i,j=1}^{N} \partial_{x_i x_j} F\, d < X^i, X^j >_t\,, \quad F \equiv F(t, X_t)\,.$$

Definition A.3. Given a d-dimensional B.m. W on $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$, $Z \in \mathbb{R}^N$, $\mu : [0,T] \times \mathbb{R}^N \to \mathbb{R}^N$, $\sigma : [0,T] \times \mathbb{R}^N \to \mathbb{R}^{N \times d}$, the s.p. $(X_t)_{t \in [0,T]}$ solves the SDE with coefficients Z, μ, σ with respect to W if:

i) $\mu(t, X_t), \sigma(t, X_t) \in \mathbb{L}^2_{loc}(\Omega \times [0,T])$;

ii) $X_t = Z + \int_0^t \mu(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s$.

A statistical tool

Here, we give a basic and simple result that will be very useful as commonplace in pricing.

Lemma A.0.3. *Let S be a random variable Log-Normally distributed with mean m and variance Σ^2, i.e. $S = C\, e^X$ with $X \sim \mathcal{N}(m, \Sigma^2)$ and $C \in \mathbb{R}$, and K be a real positive constant, $K \in \mathbb{R}^+$. Then:*

$$E\left[(S-K)^+\right] = C\, e^{\frac{\Sigma^2}{2}+m}\, \Phi\left(\frac{\ln \frac{C}{K} + m + \Sigma^2}{\Sigma}\right) - K\, \Phi\left(\frac{\ln \frac{C}{K} + m}{\Sigma}\right), \quad (A.1)$$

where Φ is the cumulative distribution function for the $\mathcal{N}(0,1)$ distribution.

Proof.

$$
\begin{aligned}
E\left[(S-K)^+\right] &= \int_{\mathbb{R}} \max\{C\,e^x - K, 0\}\,\frac{1}{\sqrt{2\pi}\Sigma}\,e^{-\frac{(x-m)^2}{2\Sigma^2}}\,dx \\
&= \frac{1}{\sqrt{2\pi}\Sigma}\int_{\{C\,e^x > K\}} (C\,e^x - K)\,e^{-\frac{(x-m)^2}{2\Sigma^2}}\,dx \\
&= \frac{1}{\sqrt{2\pi}\Sigma}\left(\int_{\{x>\ln\frac{K}{C}\}} C\,e^x\,e^{-\frac{(x-m)^2}{2\Sigma^2}}\,dx - K\int_{\{x>\ln\frac{K}{C}\}} e^{-\frac{(x-m)^2}{2\Sigma^2}}\,dx\right) \\
&= \frac{1}{\sqrt{2\pi}\Sigma}\left(\int_{\ln\frac{K}{C}}^{+\infty} C\,e^{-\frac{1}{2\Sigma^2}(x^2+m^2-2xm-2x\Sigma^2}\,dx + \right. \\
&\qquad \left. -K\int_{\ln\frac{K}{C}}^{+\infty} e^{-\frac{(x-m)^2}{2\Sigma^2}}\,dx\right) \\
&= \frac{1}{\sqrt{2\pi}\Sigma}\left(\int_{\frac{\ln\frac{K}{C}-m-\Sigma^2}{\Sigma}}^{+\infty} C\,e^{-\frac{z^2}{2}}\,e^{\frac{\Sigma^2}{2}+m}\,\Sigma\,dz + \right. \qquad \text{(A.2)} \\
&\qquad \left. -K\int_{\frac{\ln\frac{K}{C}-m}{\Sigma}}^{+\infty} e^{-\frac{u^2}{2}}\,\Sigma\,dx\right) \qquad \text{(A.3)} \\
&= \frac{1}{\sqrt{2\pi}}\left(C\,e^{\frac{\Sigma^2}{2}+m}\int_{\frac{\ln\frac{K}{C}-m-\Sigma^2}{\Sigma}}^{+\infty} e^{-\frac{z^2}{2}}\,dz - K\int_{\frac{\ln\frac{K}{C}-m}{\Sigma}}^{+\infty} e^{-\frac{u^2}{2}}\,dx\right) \\
&= C\,e^{\frac{\Sigma^2}{2}+m}\,\Phi\left(\frac{\ln\frac{C}{K}+m+\Sigma^2}{\Sigma}\right) - K\,\Phi\left(\frac{\ln\frac{C}{K}+m}{\Sigma}\right), \qquad \text{(A.4)}
\end{aligned}
$$

where the two summands in (A.2)-(A.3) are obtained respectively with the two following changes of variable:

$$
z = \frac{x-m-\Sigma^2}{\Sigma}, \quad u = \frac{x-m}{\Sigma}.
$$

□

Corollary A.0.4. *If S is a random variable Log-Normally distributed with mean $-\frac{1}{2}\Sigma^2$ and variance Σ^2, i.e. $S = C\,e^X$ with $X \sim \mathcal{N}(-\frac{1}{2}\Sigma^2, \Sigma^2)$ and $C \in \mathbb{R}$, and K is a real positive constant, $K \in \mathbb{R}^+$, then:*

$$
E\left[(S-K)^+\right] = C\Phi(d_1(K, C, \Sigma^2)) - K\Phi(d_2(K, C, \Sigma^2)), \qquad \text{(A.5)}
$$

where

$$d_1(K,S,\Sigma^2) := \frac{\ln\left(\frac{S}{K}\right)+\frac{\Sigma^2}{2}}{\Sigma}, \tag{A.6}$$

$$d_2(K,S,\Sigma^2) := d_1(K,C,\Sigma^2)-\Sigma\,. \tag{A.7}$$

Proof. It follows immediately from (A.1) by substituting m with $-\frac{1}{2}\Sigma^2$. □

Appendix B

Change of measure

The change of probability measure will be essential to deal with the concept of martingale measure, which plays a central role in discrete as well as continuous market modeling.

For mathematical reasons, financial prices are expressed in terms of expectation values, but if we use the probability measure of the real world we'll be in wrong, because the market would not be free of arbitrage opportunities. Here comes the martingale measure, which leads to the risk-neutral price for derivatives in an arbitrage-free market.

In the following, we are assuming to have a probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$ and, on this, a d-dimensional standard Brownian motion $(W_t)_{t\in[0,T]}$.

Definition B.1. Let $\lambda \in \mathbb{L}^2_{loc}(\Omega \times [0,T])$ be a d-dimensional process, the *exponential martingale* associated to λ is the s.p. Z^λ defined by

$$Z_t^\lambda = \exp\left(-\int_0^t \lambda_s \cdot dW_s - \frac{1}{2}\int_0^t |\lambda_s|^2 ds\right), \quad t \in [0,T].$$

Its dynamics is

$$dZ_t^\lambda = -Z_t^\lambda \lambda_t \cdot dW_t .$$

Indeed, by Ito's formula:

$$dZ_t^\lambda = Z_t^\lambda\left(-\lambda_t \cdot dW_t - \frac{1}{2}|\lambda_t|^2 dt\right) + \frac{1}{2}Z_t^\lambda |\lambda_t|^2 dt = -Z_t^\lambda \lambda_t \cdot dW_t .$$

Since $\lambda \in \mathbb{L}^2_{loc}$ and Z^λ is a continuous adapted process, than $Z^\lambda \lambda \in \mathbb{L}^2_{loc}$, so that Z^λ is a continuous local martingale. Moreover, being positive, it is also a super-martingale, i.e.

$$E[Z^\lambda_t] \leq [Z^\lambda_0] = 1\,, \quad t \in [0,T]\,.$$

There are some situations in which Z^λ becomes a strict martingale.

Lemma B.0.5. *If there exists a constant C such that*

$$\int_0^T |\lambda_t|^2 dt \leq C \quad a.s.\,,$$

then the exponential martingale associated to λ, Z^λ, is a strict martingale such that

$$E\left[\sup_{0\leq t\leq T} (Z^\lambda_t)^p\right] < \infty\,, \quad p \geq 1\,.$$

In particular $Z^\lambda \in \mathbb{L}^p(\Omega, P)$ for every $p \geq 1$.

Theorem B.0.6 (Novikov condition). *If $\lambda \in \mathbb{L}^2_{loc}(\Omega \times [0,T])$ is such that*

$$E\left[\exp\left(\frac{1}{2}\int_0^T |\lambda_s|^2 ds\right)\right] < \infty\,, \tag{B.1}$$

then the the exponential martingale associated to λ, Z^λ, is a strict martingale.

Let us endow the space $(\Omega, \mathcal{F}, P)$ with the Brownian filtration $\mathcal{F}^W = (\mathcal{F}^W_t)_{t\in[0,T]}$.

Theorem B.0.7 (Martingal representation). *Let $M = (M_t)_{t\in[0,T]}$ be a $\mathcal{F}^W$-local martingale, then there exists a unique (just up to $(m \otimes P)$-equivalence) process $u \in \mathbb{L}^2_{loc}(\mathcal{F}^W)$ such that*

$$M_t = M_0 + \int_0^t u_s \cdot dW_s\,, \quad t \in [0,T]\,.$$

The following basic theorem is used in the main topic of this chapter.

Theorem B.0.8 (Bayes' Formula). *Let P, Q be probability measures on $(\Omega, \mathcal{F})$, with $Q \ll_{\mathcal{F}} P$, $X \in \mathrm{L}^1(\Omega, Q)$, and $\mathcal{G}$ a sub-σ-algebra of $\mathcal{F}$. Let L be the Radon-Nikodym derivative of Q with respect to P, i.e.*

$$Q(F) = \int_F L dP\,, \quad F \in \mathcal{F}\,,$$

denoted by $L = \frac{dQ}{dP} = \frac{dQ}{dP}|_{\mathcal{F}}$. Then we have

$$E^Q[X \mid \mathcal{G}] = \frac{E^P[XL \mid \mathcal{G}]}{E^P[L \mid \mathcal{G}]}\,.$$

Proof. First denote $B := E^P[L|\mathcal{G}]$ and prove that $Q(B > 0) = 1$:

$$\{B = 0\} \in \mathcal{G} \quad \Rightarrow \quad Q(\{B = 0\}) = \int_{\{B=0\}} L dP = \int_{\{B=0\}} B dP = 0\,.$$

Then denote $A := E^Q[X|\mathcal{G}]$ and prove that $AB = E^P[XL|\mathcal{G}]$: for all $G \in \mathcal{G}$

$$\begin{aligned} \int_G AB\, dP &= \int_G E^P[AL|\mathcal{G}]\, dP = \int_G AB\, dP = \int_G A\, dQ = \\ &= \int_G E^Q[X|\mathcal{G}]\, dQ = \int_G X\, dQ = \int_G XL\, dP\,. \end{aligned}$$

□

Thus, if Q is a probability measure on $(\Omega, \mathcal{F})$ defined by

$$Z_T^\lambda = \frac{dQ}{dP}\,, \tag{B.2}$$

then, for every $X \in \mathrm{L}^1(\Omega, Q)$ we have

$$E^Q[X \mid \mathcal{F}_t] = \frac{E^P[XZ_T^\lambda \mid \mathcal{F}_t]}{E^P[Z_T^\lambda \mid \mathcal{F}_t]}\,, \quad t \in [0, T]\,.$$

Consequently we have the following lemma.

Lemma B.0.9. *Let Z^λ be a P-martingale and Q the probability measure on $(\Omega, \mathcal{F})$ defined by (B.2). Then a process $(M_t)_{t\in[0,T]}$ is a Q-martingale if and only if $(M_t Z_t^\lambda)_{t\in[0,T]}$ is a P-martingale.*

Theorem B.0.10 (**Girsanov's Theorem**). *Let Z^λ be a P-martingale and Q the probability measure on $(\Omega, \mathcal{F})$ defined by (B.2). Then the process W^λ defined by*

$$W_t^\lambda = W_t + \int_0^t \lambda_s \cdot ds\,, \quad t \in [0, T]\,,$$

is a Brownian motion on $(\Omega, \mathcal{F}, Q, (\mathcal{F}_t))$.

In financial applications, the processes playing the role of λ are often bounded, so that the martingale property of Z^λ follows from Lemma B.0.5.

Theorem B.0.11 (**Change of drift**). *Let Q be a probability measure on $(\Omega, \mathcal{F})$ equivalent to P, i.e. $Q \sim P$. Then, the Radon-Nikodym derivative of Q with respect to P is an exponential martingale:*

$$\frac{dQ}{dP}|_{\mathcal{F}_t^W} = Z_t^\lambda\,, \quad dZ_t^\lambda = -Z_t^\lambda \lambda_t \cdot dW_t\,,$$

with associated process $\lambda \in \mathbb{L}^2_{loc}(\Omega \times [0, T])$, *and the process W^λ defined by*

$$dW_t = dW_t^\lambda - \lambda_t dt \tag{B.3}$$

is a Brownian motion on $(\Omega, \mathcal{F}, Q, (\mathcal{F}_t))$.

Proof. By denoting $Z_t := \frac{dQ}{dP}|_{\mathcal{F}_t^W}$, we have

$$Z_t = E^P\left[\frac{dQ}{dP}\,|\,\mathcal{F}_t^W\right]\,, \quad t \in [0, T]\,.$$

Indeed, for all $A \in \mathcal{F}_t^W$,

$$\int_A Z_t dP = \int_A dQ = \int_A \frac{dQ}{dP} dP\,.$$

Thus Z is a positive P-martingale, in fact $Z_t \geq 0\ \forall t \in [0, T]$ and, $\forall s < t$,

$$E^P\left[Z_t\,|\,\mathcal{F}_s^W\right] = E^P\left[E^P\left[\frac{dQ}{dP}\,|\,\mathcal{F}_t^W\right]\,|\,\mathcal{F}_s^W\right] = E^P\left[\frac{dQ}{dP}\,|\,\mathcal{F}_s^W\right] = Z_s\,.$$

Then, by the martingale representation theorem B.0.7, there exists a unique (P-a.s.) d-dimensional process $u \in \mathbb{L}^2_{loc}(\mathcal{F}^W)$ such that

$$dZ_t = u_t \cdot dW_t\,, \quad \text{or equivalently} \quad dZ_t = -Z_t \lambda_t \cdot dW_t\,,$$

having defined λ as the process

$$\lambda_t = \frac{u_t}{Z_t}\,, \quad t \in [0,T]\,.$$

Since $u \in \mathbb{L}^2_{loc}$ and Z is a continuous adapted process, than $\lambda \in \mathbb{L}^2_{loc}$.
Hence Z is the exponential martingale associated with λ.
Finally, the Girsanov theorem states that W^λ in (B.3) is a B.m. on $(\Omega, \mathcal{F}, Q, (\mathcal{F}_t))$.

□

Remark 12. Let X be an N-dimensional Itô process of the form

$$dX_t = \mu_t dt + \sigma_t dW_t\,,$$

μ having values in $\mathbb{R}^N$ and σ having values in $\mathbb{R}^{N\times d}$.
Given $Q \sim P$, then the Q-dynamics of X is

$$dX_t = (\mu_t - \sigma_t \lambda_t)dt + \sigma_t dW_t^\lambda\,.$$

Appendix C

Change of Measure with Correlation in Arbitrage Theory

In a continuous-time market model, where considering a probability space $(\Omega, \mathcal{F}, P)$, the sources of risk are usually represented by a d-dimensional *correlated Brownian motion* $W = (W^1, \dots, W^d)$ on $(\Omega, \mathcal{F}, P)$, endowed with the Brownian filtration $(\mathcal{F}_t^W)_{t\in[0,T]}$. We consider the correlation constant in time, thus define

$$W_t = A \cdot \bar{W}_t \,,$$

where $\bar{W}$ is a standard d-dimensional Brownian motion and $A = (A^{ij})_{i,j=1,\dots,d}$ is a non-singular $d \times d$ constant matrix.
We denote $\rho := AA^*$ and assume that, for $i = 1, \dots, d$,

$$\rho^{ii} = \sum_{j=1}^{d} (A^{ij})^2 = 1 \quad \text{a.s.}$$

Remark 13. For $i = 1, \dots, d$, W^i is a standard 1-dimensional Brownian motion.

Proof. All the properties of a real standard B.m. are verified:

1.
$$W_0^i = \sum_{j=1}^{d} A^{ij} \bar{W}_0^j = 0\,;$$

2. W is a continuous adapted process on $(\Omega, \mathcal{F}, P)$;

3.
$$W_{t+h}^i - W_t^j = \sum_{j=1}^{d} A^{ij} \left(\bar{W}_{t+h}^j - \bar{W}_t^j\right) \sim \mathcal{N}\left(0, h\sum_{j=1}^{d} (A^{ij})^2\right) = \mathcal{N}(0,h)$$

and $\left(W_{t+h}^i - W_t^j\right)$ does not depend on $\mathcal{F}_t$.

□

The co-variation process of W has components given by

$$d < W^i, W^j >_t = \rho^{ij} dt\,, \quad i,j = 1,\ldots,d\,.$$

Notice that for all t it coincides with the covariance matrix of W. Indeed:

$$\bar{W}_t \sim \mathcal{N}(0, t\, Id_d) \quad \Rightarrow \quad W_t \sim \mathcal{N}(0, t\, AA')$$

$$\Rightarrow \quad \mathrm{Cov}(W_t^i, W_t^j) \;=\; \rho^{ij}\, t \;= < W^i, W^j >_t \;.$$

Remark 14. ρ is the correlation matrix of $W(t)$ for each fixed time t.

Indeed,

$$\mathrm{Corr}(W_t^i, W_t^j) = \frac{\mathrm{Cov}(W_t^i, W_t^j)}{\mathrm{Std}(W_t^i)\mathrm{Std}(W_t^j)} = \frac{\rho^{ij}\, t}{\sqrt{t}\sqrt{t}} = \rho^{ij}\,,$$

where we denote by $\mathrm{Std}(W_t^i)$ the standard deviation of W_t^i, i.e.

$$\mathrm{Std}(W_t^i) := \sqrt{\mathrm{Cov}(W_t^i, W_t^i)}\,.$$

When modeling a financial market, we assume the following hypothesis.

- There are N risky assets whose price process is $S = (S^1, \dots, S^N)$ and one locally no-risky asset B, that satisfy respectively the following dynamics:

$$dS_t^i = \mu_t^i S_t^i dt + \sigma_t^i S_t^i dW_t^i\,, \quad \mu^i = b^i + \frac{(\sigma^i)^2}{2}\,, \quad i = 1, \dots, N \tag{C.1}$$

and

$$dB_t = r_t B_t dt\,, \quad B_0 = 1\,, \tag{C.2}$$

where $b, r \in \mathbb{L}^1_{loc}(\Omega \times [0,T])$, σ^i positive $\in \mathbb{L}^2_{loc}(\Omega \times [0,T])\ \forall i$.
The dynamics in (C.1) derives from supposing

$$S_t^i = e^{X_t^i}\,, \quad dX_t^i = b_t^i dt + \sigma_t^i dW_t^i\,, \quad i = 1, \dots, N\,.$$

In fact, by Itô's formula,

$$dS_t^i = \partial_x S_t^i dX_t^i + \frac{1}{2}\partial_{xx} S_t^i d < X^i, X^i >_t = S_t^i(b_t^i dt + \sigma_t^i dW_t^i) + \frac{1}{2} S_t^i (\sigma^i)^2 dt\,.$$

The integrated form of the solution of the SDE (C.1) is:

$$S_t^i = S_0^i \exp\left(\int_0^t \sigma_s^i dW_s^i + \int_0^t \left(\mu_s^i - \frac{(\sigma^i)^2}{2} \right) ds \right)\,,$$

obtained by searching the two processes A, B such that
$S_t^i = S_0^i \exp\left(\int_0^t A_s dW_s^i + \int_0^t B_s ds \right)$, then by applying the Itô's formula and equating it to (C.1).
Notice that B, although representing a locally no-risky asset, is a stochastic process, because r is a $\mathcal{F}^W$-progressively measurable process. Anyway it has a smaller degree of randomness with respect to the other assets, because it has bounded variation and consequently null co-variation process. Indeed:

$$\begin{aligned} r \in \mathbb{L}^1_{loc} \quad &\Rightarrow \quad r(w) \in \mathrm{L}^1_{[0,t]}\ a.s. \text{ for } w \in \Omega \\ \Rightarrow \quad \textstyle\int_0^t r_s ds \in \mathrm{BV} \quad &\Rightarrow \quad \exp\left(\textstyle\int_0^t r_s ds \right) \in \mathrm{BV}\,. \end{aligned}$$

- r and σ satisfy Lemma B.0.5.

- W is a d-dimensional Brownian motion with constant correlation matrix ρ and $d \geq N$.

Theorem C.0.12 (Change of drift with correlation). *Let Q be a probability measure on $(\Omega, \mathcal{F})$ equivalent to P, i.e. $Q \sim P$. Then, the Radon-Nikodym derivative of Q with respect to P is an exponential martingale:*

$$\frac{dQ}{dP}|_{\mathcal{F}_t^W} = Z_t^\lambda\,, \quad dZ_t^\lambda = -Z_t^\lambda \lambda_t \cdot dW_t\,,$$

with associated process $\lambda \in \mathbb{L}^2_{loc}(\Omega \times [0,T])$, and the s.p. W^λ defined by

$$dW_t = dW_t^\lambda - \rho\lambda_t dt \tag{C.3}$$

is a Brownian motion on $(\Omega, \mathcal{F}, Q, (\mathcal{F}_t^W))$ with correlation matrix ρ.

Proof. Let $Z_t = \frac{dQ}{dP}|_{\mathcal{F}_t^W}$, by the martingale representation theorem for the standard B.m., there exists a unique (P-a.s.) d-dimensional process $\overline{\lambda} \in \mathbb{L}^2_{loc}(\mathcal{F}^W)$ such that

$$dZ_t = -Z_t\overline{\lambda}_t \cdot d\overline{W}_t = -Z_t\overline{\lambda}_t \cdot (A^{-1}dW_t) = -Z_t\lambda_t \cdot dW_t\,,$$

where $\lambda_t := (A^{-1})'\overline{\lambda}_t$. Notice that, by Itô's formula,

$$\begin{aligned} Z_t &= \exp\left(-\int_0^t \overline{\lambda}_s \cdot d\overline{W}_s - \frac{1}{2}\int_0^t |\overline{\lambda}_s|^2 ds\right) \\ &= \exp\left(-\int_0^t < A^*\lambda_s, A^{-1}dW_s > -\frac{1}{2}\int_0^t < A^*\lambda_s, A^*\lambda_s > ds\right) \\ &= \exp\left(-\int_0^t \lambda_s \cdot dW_s - \frac{1}{2}\int_0^t < \rho\lambda_s, \lambda_s > ds\right)\,. \end{aligned}$$

Then, by the Girsanov theorem, we have a standard Q-B.m. $\overline{W}^{\overline{\lambda}}$, $d\overline{W}_t^{\overline{\lambda}} = d\overline{W}_t + \overline{\lambda}_t dt$, $t \in [0,T]$. Multiplying by A this equation,

$$dW_t^\lambda := A\, d\overline{W}_t^{\overline{\lambda}} = dW_t + \rho\lambda_t dt$$

is a correlated Q-B.m. with correlation matrix ρ. □

Remark 15. Analogously to the standard case, let X be an N-dimensional Itô process of the form

$$dX_t = \mu_t dt + \sigma_t dW_t \,,$$

μ having values in $\mathbb{R}^N$ and σ having values in $\mathbb{R}^{N \times d}$.
Given $Q \sim P$, then the Q-dynamics of X is

$$dX_t = (\mu_t - \sigma_t \rho \lambda_t) dt + \sigma_t dW_t^{\lambda} \,.$$

Appendix D

Change of numeraire

Now we characterize the arbitrage-free financial markets by introducing the previously mentioned equivalent martingale measure.

Definition D.1. An *equivalent martingale measure* (*EMM*) Q with numeraire B is a probability measure on $(\Omega, \mathcal{F})$ such that:

i) Q is equivalent to P;

ii) the process of the discounted prices $\tilde{S} = (\tilde{S}_t)_{t\in[0,T]}$ defined by

$$\tilde{S}_t = \frac{S_t}{B_t} = e^{-\int_0^t r_s ds} S_t \,,\ t \in [0,T]\,,$$

is a strict Q-martingale. In particular, the *risk-neutral pricing formula*

$$S_t = E^Q \left[e^{-\int_t^T r_s ds} S_T \mid \mathcal{F}_t^W \right] , \quad t \in [0,T]\,,$$

holds.

We consider a market model (S, B) of the form (C.1)-(C.2) and assume that the class $\mathcal{Q}$ of the EMMs is not empty.

Definition D.2. Let $Q \in \mathcal{Q}$ be an EMM with numeraire B. A s.p. U is called a *Q-price process* if:

i) U is strictly positive;

ii) the process of its discounted price $\tilde{U}$, $\tilde{U}_t = \frac{U_t}{B_t}$, $t \in [0,T]$, is a strict Q-martingale.

Practically, a Q-price process has all the features of a true price. The martingale property leads to the risk-neutral pricing formula under Q:

$$U_t = E^Q \left[D(t,T) U_T \,|\, \mathcal{F}_t^W \right], \quad t \in [0,T],$$

where $D(t,T) = \frac{B_t}{B_T} = e^{-\int_t^T r_s ds}$ is the standard discount factor.

Notice that any risky asset S^i is a Q-price process.

The so called *numeraire* is a s.p. that represents a basic standard by which the prices of all other assets are measured.

Definition D.3. Let U be a Q-price process, a probability measure Q^U on $(\Omega, \mathcal{F})$ is called an *EMM with numeraire U* if:

i) Q^U is equivalent to P;

ii) the processes of U-discounted prices $\frac{S_t}{U_t}$, $\frac{B_t}{U_t}$ are strict Q^U-martingales. In particular, the *risk-neutral pricing formulae*

$$S_t = E^{Q^U} \left[D^U(t,T) S_T \,|\, \mathcal{F}_t^W \right], \tag{D.1}$$

$$B_t = E^{Q^U} \left[D^U(t,T) B_T \,|\, \mathcal{F}_t^W \right], \quad t \in [0,T] \tag{D.2}$$

hold, where $D^U(t,T) = \frac{U_t}{U_T}$ is the U-discount factor.

Theorem D.0.13. *Let $Q \in \mathcal{Q}$ be an EMM with numeraire B and let U be a Q-price process. Consider the probability measure Q^U on $(\Omega, \mathcal{F})$ defined by*

$$\frac{dQ^U}{dQ} = \frac{D(0,T)}{D^U(0,T)} = \frac{U_T B_0}{B_T U_0}.$$

Then, for any $X \in \mathrm{L}^1(\Omega, Q)$, we have

$$E^Q \left[D(t,T) X \,|\, \mathcal{F}_t^W \right] = E^{Q^U} \left[D^U(t,T) X \,|\, \mathcal{F}_t^W \right], \quad t \in [0,T]. \tag{D.3}$$

In particular Q^U is an EMM with numeraire U and the risk-neutral price of a European derivative X is equal to

$$E^{Q^U} \left[D^U(t,T) X \,|\, \mathcal{F}_t^W \right], \quad t \in [0,T].$$

Proof. Denote
$$Z_t := \frac{D(0,t)}{D^U(0,t)} = \frac{U_t B_0}{B_t U_0}, \quad t \in [0,T].$$
Since U is a Q- price process,

$$Z_t = \frac{B_0}{U_0} E^Q\left[\frac{U_T}{B_T} \,|\, \mathcal{F}_t^W\right] = E^Q\left[\frac{D(0,T)}{D^U(0,T)} \,|\, \mathcal{F}_t^W\right] = E^Q\left[Z_T \,|\, \mathcal{F}_t^W\right],$$

hence Z is a strictly positive Q-martingale. Then, by the Bayes' formula,

$$\begin{aligned} E^{Q^U}\left[X \,|\, \mathcal{F}_t^W\right] &= \frac{E^Q\left[X Z_T \,|\, \mathcal{F}_t^W\right]}{E^Q\left[Z_T \,|\, \mathcal{F}_t^W\right]} = E^Q\left[X \frac{Z_T}{Z_t} \,|\, \mathcal{F}_t^W\right] \\ &= E^Q\left[X \frac{D(0,T)}{D^U(0,T)} \,|\, \mathcal{F}_t^W\right], \end{aligned} \tag{D.4}$$

because
$$\frac{Z_T}{Z_t} = \frac{U_T B_0}{B_T U_0}\frac{B_t U_0}{U_t B_0} = \frac{U_T B_0}{B_T U_0} = \frac{D(0,T)}{D^U(0,T)}.$$
Now, taking $(D^U(t,T)X)$ in place of simply X in (D.4), we obtain

$$\begin{aligned} E^Q\left[D(t,T)X \,|\, \mathcal{F}_t^W\right] &= E^Q\left[\frac{D(t,T)}{D^U(t,T)}(D^U(t,T)X) \,|\, \mathcal{F}_t^W\right] \\ &= E^{Q^U}\left[D^U(t,T)X \,|\, \mathcal{F}_t^W\right]. \end{aligned}$$

Moreover $Q^U \sim Q$, indeed $\exists \frac{dQ^U}{dQ}, \frac{dQ}{dQ^U} > 0$, $\frac{dQ^U}{dQ}, \frac{dQ}{dQ^U} \in \mathbb{L}^1$, therefore $Q^U \sim P$. Finally, (D.3) prove (D.1) and (D.2). □

Corollary D.0.14. *Let U, V be Q-price processes with corresponding EMMs Q^U, Q^V, respectively. Then, we have*

$$\frac{dQ^V}{dQ^U}|\mathcal{F}_t^W = \frac{V_t U_0}{U_t V_0}. \tag{D.5}$$

Proof.

$$\begin{aligned} \frac{dQ^V}{dQ^U}|\mathcal{F}_t^W &= E^{Q^U}\left[\frac{dQ^V}{dQ}\frac{dQ}{dQ^U} \,|\, \mathcal{F}_t^W\right] = E^{Q^U}\left[\frac{V_T U_0}{U_T V_0} \,|\, \mathcal{F}_t^W\right] \\ &= \frac{U_0}{V_0} E^{Q^U}\left[\frac{V_T}{U_T} \,|\, \mathcal{F}_t^W\right] = \frac{U_0}{V_0}\frac{V_t}{U_t}. \end{aligned}$$

The last equality follows from Theorem D.0.13, since V is a Q-price process.

□

Forward measure

Let $p(t,T)$ be the price at time t, under a fixed EMM Q with numeraire B, of the zero coupon bond with maturity T:

$$p(t,T) = E^Q\left[e^{-\int_t^T r_s ds} \,|\, \mathcal{F}_t^W\right], \quad t \le T. \tag{D.6}$$

Clearly $p(t,T)$ is a Q-price, because it is just defined through the risk neutral pricing formula under the selected EMM Q. Thus there exist an EMM associated with it, that is Q^T, called *T-forward measure.*
By Theorem D.0.13, the risk neutral price H of a European derivative X, at time t, is equal to

$$H_t = E^{Q^T}\left[\frac{p(t,T)}{p(T,T)} X \,|\, \mathcal{F}_t^W\right] = p(t,T)\, E^{Q^T}\left[X \,|\, \mathcal{F}_t^W\right]. \tag{D.7}$$

This pricing formula in terms of a Q^T-expectation does not involve the stochastic discount factor $e^{-\int_t^T r_s ds}$, as instead the Q-expectation does. On the other hand, it needs to know the distribution of X under Q^T, which can be deduced by a change of drift (Theorem B.3) in terms of the Radon-Nikodym derivative $\frac{dQ^T}{dQ} = \frac{B_0}{p(0,T)B_T}$ (from (D.0.14)).

Now we move towards the cardinal theorem of the change of measure induced by numeraires that are Itô processes.

Lemma D.0.15. *Let U, V be two positive Itô processes of the form*

$$dU_t = (\cdots)dt + \sigma_t^U \cdot dW_t\,,$$
$$dV_t = (\cdots)dt + \sigma_t^V \cdot dW_t\,,$$

where W is a correlated d-dimensional Brownian motion and $\sigma^U, \sigma^V \in \mathbb{L}^2_{loc}(\Omega \times [0,T]; \mathbb{R}^d)$ are the diffusion coefficients. Then, $\frac{V_t}{U_t}$ is an Itô process of the form

$$d\frac{V_t}{U_t} = (\cdots)dt + \frac{V_t}{U_t}\left(\frac{\sigma_t^V}{V_t} - \frac{\sigma_t^U}{U_t}\right) \cdot dW_t\,. \tag{D.8}$$

Proof. Apply the Itô formula to the function F of the 2-dimensional Itô process (U,V), $F(U,V) = \frac{V}{U}$:

$$\begin{aligned} d\frac{V_t}{U_t} &= -\frac{V_t}{(U_t)^2}dU_t + \frac{dV_t}{U_t} + \frac{V_t}{(U_t)^3}d < U,U >_t - \frac{1}{(U_t)^2}d < U,V >_t \\ &= (\cdots)dt - \frac{V_t}{(U_t)^2}\sigma_t^U \cdot dW_t + \frac{1}{U_t}\sigma_t^V \cdot dW_t\,. \end{aligned}$$

□

Theorem D.0.16 (Change of numeraire). *Let U,V be two Q-price processes of the form*

$$dU_t = (\cdots)dt + \sigma_t^U \cdot dW_t\,,$$
$$dV_t = (\cdots)dt + \sigma_t^V \cdot dW_t\,,$$

where W is a correlated d-dimensional Brownian motion with correlation matrix ρ and $\sigma^U, \sigma^V \in \mathbb{L}^2_{loc}(\Omega \times [0,T];\mathbb{R}^d)$. Let Q^U, Q^V be the EMMs related to U,V respectively and W^U, W^V be the corresponding Brownian motions. Then:

$$dW_t^U = dW_t^V + \rho\left(\frac{\sigma_t^V}{V_t} - \frac{\sigma_t^U}{U_t}\right) \cdot dt\,. \tag{D.9}$$

Proof. Apply the formula (D.5) from Corollary D.0.14:

$$\frac{dQ^V}{dQ^U}|\mathcal{F}_t^W = \frac{V_t U_0}{U_t V_0} =: Z_t\,.$$

From Lemma D.0.15 the dynamics of Z under Q is

$$\begin{aligned} dZ_t &= \frac{U_0}{V_0}\left[(\cdots)dt + \frac{V_t}{U_t}\left(\frac{\sigma_t^V}{V_t} - \frac{\sigma_t^U}{U_t}\right) \cdot dW_t\right] \\ &= (\cdots)dt + Z_t\left(\frac{\sigma_t^V}{V_t} - \frac{\sigma_t^U}{U_t}\right) \cdot dW_t\,, \end{aligned}$$

but we know from the theorem of Change of drift with correlation that Z is an exponential martingale under Q^U, with dynamics given by

$$dZ_t = -Z_t\lambda_t \cdot dW_t^U\,.$$

As we see in the formula (C.3), the diffusion coefficient is never involved in the change of measure, so that the percentage diffusion process under Q^U

equals the one under Q, i.e.

$$\lambda_t = -\frac{\sigma_t^V}{V_t} + \frac{\sigma_t^U}{U_t} \,.$$

Therefore, by applying (C.3), we have

$$dW_t^U = dW_t^V - \rho\lambda_t dt = dW_t^V + \rho\left(\frac{\sigma_t^V}{V_t} - \frac{\sigma_t^U}{U_t}\right) dt \,.$$

□

Bibliography

[1] Andrea Pascucci. *Calcolo stocastico per la finanza*. Springer-Verlag Italia, Via Decembrio 28, 20137, Milano, 2008.

[2] Andrea Pascucci. *PDE and Martingale methods in option pricing*, chapter 10. Springer-Verlag Italia, Via Decembrio 28, 20137, Milano, 2011.

[3] Tomas Björk. *Arbitrage Theory in Continuous Time*. Oxford University Press Inc., New York, United States, third edition, 2009.

[4] Damiano Brigo and Fabio Mercurio. *Interest Rate Models, Theory and Practice*. Springer-Verlag Berlin Heidelberg, Germany, second edition, 2006.

[5] Paolo Foschi. *Tools for interest rate derivatives*. February, 2011.

[6] Massimo Morini. *Interest rate and the credit crunch*, March, 2011.

[7] Schoenmakers John and Coffey Brian. *Stable implied calibration of a multi-factor LIBOR model via a semi-parametric correlation structure*. Weierstraß-Institute für Analysis und Stochastik, October, 2000.

[8] Schoenmakers John and Coffey Brian. *LIBOR rate models, related derivatives and model calibration*.Weierstraß-Institute für Analysis und Stochastik, April, 1999

[9] Riccardo Rebonato. *On the simultaneous calibration of multi-factor log-normal interest-rate models to Black volatilities and to the correlation matrix*. 1999.

[10] Damir Filipović. *Term-Structure Models, A Graduate Course*, chapter 7,11, pages 105–107,197–211. Springer-Verlag Berlin Heidelberg, 2009.