Westfälische
Wilhelms-Universität
Münster

Calibration of a Libor Market Model with Stochastic Volatility

Master's Thesis

by

Hendrik Hülsbusch

Submitted in Partial Fulfillment for the

Degree of Master of Science
in
Mathematics

Supervisor:

PD. Dr. Volkert Paulsen

Münster, August 27, 2014

Contents

List of Figures

List of Tables

Eidesstattliche Erklärung

Hiermit versichere ich, Hendrik Hülsbusch, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Gedanklich, inhaltlich oder wörtlich übernommenes habe ich durch Angabe von Herkunft und Text oder Anmerkung belegt bzw. kenntlich gemacht. Dies gilt in gleicher Weise für Bilder, Tabellen, Zeichnungen und Skizzen, die nicht von mir selbst erstellt wurden.

Hendrik Hülsbusch, Münster, August 27, 2014

1. Introduction

The market for interest rate derivatives is one of the biggest financial markets in the world and is easily bigger than the stock market. A huge part of the market for options on interest rates is the over-the-counter (OTC) market. Different from stock exchanges where prices of products are publicly quoted on consent, in the OTC market the involved parties negotiate over prices only between themselves behind closed doors without making them public. In the global OTC derivatives market positions of almost 600 trillion USD were outstanding in 2013 [39] and all of those derivatives can be found as positions in the balance sheets of financial institutions. To evaluate those positions and to set prices in OTC trades sophisticated models are needed.

The main underlyings for products in the world of interest rate derivatives are forward rates. Roughly speaking, those rates give the interest rate as of today for some future time period. The most famous forward rates are the Libor/Euribor forward rates. Those rates dictate the conditions to which big and liquid financial institutions lend money to each other. A majority of the derivatives that have to be priced depend on more than one forward rate.

In 1997 the first Libor Market Models (LMMs) to describe a set of forward rates consistent with each other were published in [2] and [11]. This was a real breakthrough, because, firstly, the forward rates are modeled directly and, secondly, it enabled market participants to evaluate whole books of options depending on a range of forward rates in arbitrage free manners. One drawback in the early simple LMM is the incapably to incorporate the observable smile effect due to the deterministic volatility structure. So the model is only capable to evaluate European options on those strikes which are used for the model calibration and, even worse, it can only be used to evaluate a European options for *exactly* one strike. In most cases this strike is at the money. Obviously a reasonable model should be able to price options on any strike.

A simple model for forwards rates that is capable to incorporate the smile effect is the SABR model and was introduced by *Hagan* [18]. It is popular since, because it easy to understand and to calibrate at the same time. The SABR model is a one asset

model with stochastic volatility and thus gives a way to incorporate the market smile. Therefore, it is possible to evaluate a book of European options on more than one strike – at least for books depending on only one underlying. The model is a step in the right direction, but for the sake of evaluating whole balance sheets consisting of options on more than one underlying it is not enough.
Rebonato proposed in [36] the SABR-LMM, which is a hybrid of the SABR model and the LMM. The SABR-LMM is a market model which can do both, it incorporates the market smile and it describes the dynamics of a set of forward rates. Simultaneously, it tries to preserve the simple SABR dynamics for the single assets as close as possible.
A main issue is to calibrate the SABR-LMM to the market to reflect the dynamics of the real world. The goal of this work is to tackle this problem by giving the right frameworks for an implicit calibration to current market prices. We will focus on the calibration to cap and swaption prices and on the calibration to cap and constant maturity swap (CMS) spread option prices. The first was introduced in [36] and is revised thoroughly in this work. For the second calibration approach we extend the work of *Kienitz* & *Wittke* [22] to our SABR-LMM environment.
A subproblem in the calibration task is to find an appropriate parametrization of the model coefficients and structures. A *stylized* parametrization is required to guarantee a stable implementation. To describe the correlation-structure of the SABR-LMM through a proper parametrization we will research the two approaches coming from *Lutz* [29] and *Schoenmakers* & *Coffey* [25].
In addition, we will test the calibration methods and the different parameterizations on two different data sets consisting of real market prices from two different dates.

The work is organized as follows. In chapter 2 we introduce basic products, derivatives and bootstrapping techniques. Then, in chapter 3 we explain the simple SABR model and the SABR-LMM. In addition we calculate the involved asset dynamics under common measures. In chapter 4 we approximate the induced swap rate dynamics for the SABR-LMM in a simple SABR framework. The parametrization of the SABR-LMM is covered in chapter 5 were we explain how we stylize the model volatilities and correlations. The implicit model calibration to cap and swaption prices is explained in chapter 6. Afterwards, we introduce in chapter 7 the concept CMS spread options and describe their dynamics in the SABR-LMM model. Further, we show how to calibrate the correlation structure to CMS spread option prices. In chapter 8 we outline the out-carried implementations of all calibration procedures explained in the previous chapters and test the calibration methods by reprising the involved products using

Monte Carlo simulations. Here we use market prices from two different dates. Last, chapter 9 concludes the results of this works.

Acknowledgments

I would like to thank Christoph Moll for providing the opportunity to write my thesis in cooperation with zeb/information.technology GmbH & Co.KG.

Further, my gratitude goes to Niels Linnemann for being a great motivator (or mentor) and for his power to have the right questions for every situation. I thank Josef Üre for his full support throughout my study.

Last but not least, I thank my love Jana for her amazing patience and encouragement during the months of writing.

2. Preliminaries

This work discusses the calibration of a market model for forward rates. To understand the model we first need to understand the concept of forwards and to comprehend the calibration processes we first have to grasp the concept behind derivatives on forward rates. For this we give a proper mathematical environment and explain all the needed concepts. Amongst others we give definitions for basic financial instruments and show how to bootstrap the needed market data. This data will be essential in the calibration part. Further we give a method to evaluate options market consistent.

2.1. Forward Rates and Swaps

In this section we first introduce two basic instruments that will be the basis of this work. We show how they are related and give important formulas that will accompany us throughout this thesis. After that, we establish derivatives on these products and give evaluations formulas for pricing. This section follows [7] and [8].

2.1.1. Basic Definitions

To talk about financial requires a concept of time and time steps. A *tenor structure* $(T_i)_{i\in\{0,\dots,N\}}$ is a set of real numbers consists of all time points of interest to the model. Usually the structure starts at $T_0 = 0$ which can be interpreted as the valuation date we are looking at. At this point in time all market prices are known. Formally a tenor structure can be defined as

Definition 1 (Tenor Structure).
A tenor structure $(T_i)_{i\in\{0,\dots,N\}}$ *is a finite, strictly monotonously increasing sequence of non-negative real numbers, hence*

$$0 \le T_0 < T_1 < \cdots < T_N < \infty.$$

Further we define

$$\delta_i := T_{i+1} - T_i, \quad i \in \{0, \dots, N-1\}$$

as the i-th time step.

The T_i don't have to be equally spaced in general, so we do not force δ_i to be constant. But in fact, when it comes to implementation it seems possible (and in fact turns out to be possible) to set $\delta_i = \delta$ constant without getting to inflexible.
The most basic product in market is the zero coupon bond and can be defined as follows

Definition 2 (Zero Coupon Bond).
A zero coupon bond, pays at some tenor point T *its notional* N *and has no other payments in between. We write* $B(t,T)$ *for its price at time* $t \le T$*, where* $B(T,T) = N$ *and call* T *maturity and* $T-t$ *time to maturity of the zero coupon bond.*

In most cases we just write *bond* instead of zero coupon bond. In this work it holds $N = 1$ throughout and the bond price is assumed to be positive all the time, which means $B(t,T) > 0$ for all t. In addition we assume the price process $B(t,T)$ is decreasing in T. Hence, it holds $B(t,S) < B(t,T)$, if and only if $S > T$. That means the longer the time to the payment the less worth is the bond. This agrees with our intuition about pricing.
Further, we define bond prices $B(t,T)$ as the discount factors for the periods $[0,T]$. This makes sense, since a bond pays in T exactly one and therefore gives today's value of one unit of money in T.
One of the most important concepts for this work is the forward rate. A forward rate F_t^i for the time interval $[T_i, T_{i+1}]$ at time t gives the current interest rate for that time interval, which is consistent with the bond prices $(B(t,T_i))_{i\in\{T_i \mid T_i \ge t\}}$. Here F_t^i is normalized linear to one year, that means over the period $[T_i, T_{i+1}]$ the occurring interest rate is $(1 + \delta_i F_t^i)$. This strange definition comes from market convention. Mathematically the rate is defined as

Definition 3 (Forward Rate)**.**
The forward rate at time t over time $[T_i, T_{i+1}]$ is defined as

$$F_t^i := \frac{1}{\delta_i} \frac{B(t,T_i) - B(t,T_{i+1})}{B(t,T_{i+1})}. \tag{2.1}$$

Remark. In particular, our assumptions about the bond prices imply $F_t^i \geq 0$ for all $t < T_i$ and $i \in 0, \ldots, N-1$. For clarification we emphasis: The forward rate F^i over the period $[T_i, T_{i+1}]$ has the *expiry date* T_i, which means that the rate is fixed to a certain value after this time, and that any payments are done a time step later, at the *settlement date* T_{i+1}.

To check the interpretation of a forward rate as an interest rate for a certain period define the value

$$r_{t,T} := \Big(\frac{1}{B(t,T)}\Big)^{1/(T-t)} - 1 \geq 0$$

This is the risk-less interest rate the bond pays until maturity. If we consider a bond with maturity T_i, $i > 1$, the risk-less rate r_{0,T_i} is an interest rate over more than one time period.
Following the mentioned intution about forwards it seems equally plausible to follow one of the two following strategies of which one is equivalent to the other. The first is to buy a bond with maturity T_{i+1} and get for $(T_{i+1} - t)$ periods the interest rate $r_{t,T_{i+1}}$. The second is to buy a bond with maturity T_i and a *forward rate agreement* for the period $[T_i, T_{i+1}]$. A forward rate agreement (FRA) is a contract that guarantees at time t an interest rate of exactly F_t^i over the period $[T_i, T_{i+1}]$. Both strategies should have the same payoff in the end. To verify this claim we calculate both portfolio payoffs in T_{i+1} and obtain

$$\begin{aligned}(1 + r_{t,T_i})^{(T_i - t)}(1 + \delta_i F_t^i) &= \frac{1}{B(t,T_i)}\Big(1 + \frac{B(t,T_i) - B(t,T_{i+1})}{B(t,T_{i+1})}\Big) \\ &= \frac{1}{B(t,T_{i+1})} \\ &= (1 + r_{t,T_{i+1}})^{(T_{i+1} - t)}.\end{aligned}$$

Another important financial product is the swap. A swap over a time horizon $[T_m, T_n]$ is a contract between two parties – the long and short party – which exchanges the forward rates F^i, $i \in \{m, \ldots, n-1\}$, in each period against a fixed rate K. In a *payer*

swap the long party pays the fixed rate K and has to receives the floating rates F^i. In a *receiver* swap the long party receives the fixed rate K and pays the floating rates F^i.
A swap over the period $[T_m, T_n]$ *expires* in T_m. On that date all rates are fixed to the value $F^i_{T_m}$ and don't change over the exchange time from T_m to T_n. The difference $T_n - T_m$ is the *tenor* of the swap and describes the length of the exchange period. For the long party a payer swap at time $t \leq T_m$ has the value

$$\begin{aligned} \mathrm{Swap}_t^{m,n} &:= \sum_{i=n}^{m-1} \delta_i B(t, T_{i+1})(F_t^i - K) \\ &= \sum_{i=n}^{m-1} B(t, T_i) - B(t, T_{i+1}) - \delta_i B(t, T_{i+1}) K \\ &= B(t, T_m) - B(t, T_m) - \sum_{i=n}^{m-1} \delta_i B(t, T_{i+1}) K, \end{aligned} \tag{2.2}$$

since the $B(t, T_{i+1})$ are the discount factors and the forward rate F^i is paid in T_{i+1}. In the market the value of K is chosen such that the expression in (2.2) is equal to zero. In this case K is called *swap rate*. Since it holds

$$\begin{aligned} 0 &= B(t, T_m) - B(t, T_m) - \sum_{i=n}^{m-1} \delta_i B(t, T_{i+1}) K \\ \Leftrightarrow K &= \frac{B(t, T_n) - B(t, T_m)}{\sum_{i=n}^{m-1} \delta_i B(t, T_{i+1})}, \end{aligned} \tag{2.3}$$

we define the following:

Definition 4 (Swap Rate)**.**
The swap rate at time $t \leq T_m$ over the period $[T_m, T_n]$ is given as

$$S_t^{m,n} := \frac{B(t, T_m) - B(t, T_n)}{A_t^{m,n}}, \tag{2.4}$$

where we define the swap numéraire as

$$A_t^{m,n} := \sum_{i=m}^{n-1} \delta_i B(t, T_{i+1}).$$

Remark. A swap rate over a time interval can be interpreted, due to the relation in (2.3), as the average interest rate over this period.

An important feature of a swap rate is that it can be written as a weighted sum of the involved forward rates. To realize this we write

$$\begin{aligned} S_t^{m,n} &= \frac{B(t,T_m) - B(t,T_n)}{A_t^{m,n}} \\ &= \frac{\sum_{i=m}^{n-1} B(t,T_i) - B(t,T_{i+1})}{A_t^{m,n}} \\ &= \sum_{i=m}^{n-1} \frac{\delta_i B(t,T_{i+1})}{A_t^{m,n}} F_t^i \\ &=: \sum_{i=m}^{n-1} \omega_i^{m,n}(t) F_t^i, \end{aligned} \tag{2.5}$$

where the weights are defined as

$$\omega_i^{m,n}(t) := \frac{\delta_i B(t,T_{i+1})}{A_t^{m,n}}. \tag{2.6}$$

The last equations (2.5) and (2.6) are extremely important since they enable us to see the direct link between forward rates, which we are planing to describe in a market model, and swaps, one of the most liquid products in the market. Further the sum structure shows that swaps depend on the interplay of the forwards. This will be relevant in the calibration part of this thesis.

2.1.2. Basic Derivatives

Later we want to calibrate our model to market prices. It is practice to use call or put-like derivatives for this purpose, since these simple products are the most liquid ones. High liquidity favors the reliability of the observed prices since the associated products are more likely traded on a census price. We will start with derivatives one forwards and then come to options on swaps.

The most simple derivative on a forward rate is a caplet, which is a simple call option. It enables to hedge against rising interest rates for a period of length δ_i.

Definition 5 (Payoff of a Caplet)**.**
A caplet on a forward rate F^i with strike K pays in T_{i+1} the following

$$\delta_i(F^i_{T_i} - K)^+. \tag{2.7}$$

So a caplet payment at the settlement date T_{i+1} is fixed one period earlier at the expiry date T_i.
However, in the market almost no caplets are quoted directly. They are quoted in whole portfolios of caplets which are called *caps*. A cap over the period $[T_m, T_n]$ is a sum of caplets with expiry dates T_i, $i \in \{m, \ldots, n-1\}$, where each caplet has the same strike K. This implies the following proposition:

Proposition 1 (Cap Price)**.**
The cap price of a cap ranging from T_m to T_n and strike K is given as

$$C^{m,n}(K) := \sum_{i=m}^{n-1} C^i(K), \tag{2.8}$$

where $C^i(K)$ is defined as the value of the caplet on F_i with strike K.

Remark. As for swaps the difference $T_n - T_m$ is the *tenor* and T_m the *expiry date* of the cap.

Apart from caplets, floorlets which form the counterparts of the caplets and are puts on forward rates exists. Therefore we can define the payoff of a floorlet as:

Definition 6 (Payoff of a Floorlet)**.**
A floorlet on a forward rate F^i with strike K pays in T_{i+1} the following

$$\delta_i(K - F^i_{T_i})^+. \tag{2.9}$$

Again those derivatives are not quoted directly in the market. There are only floors – a sum of floorlets – quoted. Floors can be seen as the counterpart of caps. The price of a floorlet is given as:

Proposition 2 (Floor Price)**.**
The floor price of a floor ranging from T_m to T_n and strike K is given as

$$P^{m,n}(K) := \sum_{i=m}^{n-1} P^i(K), \tag{2.10}$$

where $P^i(K)$ is defined as the value of the floorlet on F_i with strike K.

Remark. Similar to caps the difference $T_n - T_m$ is the *tenor* and T_m the expiry date of the floor.

Up to now we have discussed derivatives on forward rates. Now we want introduce options on swap rates. Those options are often referred to as *swaptions*. A swaption with strike K gives the right to enter a payer swap or receiver swap, respectively, with strike K. In our case, the swaption and swap have the same expiry dates all the time. Therefore, because of (2.2), the payoff of a payer swaption in T_n is given as

$$\Big(\sum_{i=m}^{n-1} \delta_i B(T_m, T_{i+1})(F^i_{T_m} - K)\Big)^+.$$

With the result in (2.5) we are able to rewrite this payoff in the following proposition:

Proposition 3 (Payoff of a Swaption)**.**
The payoff of a swaption on a payer swap $S^{m,n}$ at time T_n is given as

$$A^{m,n}_{T_m}(S^{m,n}_{T_m} - K)^+, \tag{2.11}$$

where $A^{m,n}$ is the swap numéraire from definition 4.

We want to emphasize that, unlike as in the case for caps it is not possible to decompose the payment (2.11) nor the value of a swaption in more elementary payoffs or prices. This is a huge distinguish feature of caps/floors and swaptions.

2.2. Bootstrapping Market Data

In the calibration part we will rely on some fundamental data which we will assume as given. This includes the prices of caplets and floorlets in any given tenor as well as prices for swaptions with any expiry date and any tenor. Further, we will need the current forward rates at the valuation date. Unfortunately, those cannot be obtained directly and have to be stripped as well.

2.2.1. The Bootstrapping of Forward Rates

To calculate the current forward rate which are consistent with the corresponding swap prices we are going to rely on the definition of forward rates as a quotient of bond prices (2.1) and on the definition of swap rates as in (2.4). Our plan is to

calculate the forward rates based on a set of swap rates starting at the valuation date and having growing tenors up to the maximal tenor $T_N - T_0$. Those swaps are quoted for a very long tenors up to 50 years. Therefore they provide the right environment to calculate all the needed forward rates.
We want to calculate the forward prices basing on bond prices. To achieve this we now calculate the needed bond prices iteratively.
It is clear that the first forward rate F_0^0 starting at the valuation date and settling in T_1 corresponds to the swap rate $S_0^{0,1}$. From this we get

$$B(0,T_1) = \frac{1}{1+\delta_0 F_0^0}. \tag{2.12}$$

This is our initial value. Next, let us consider definition 4, namely

$$S_0^{0,n} = \frac{B(0,T_0) - B(0,T_n)}{A_0^{0,n}} = \frac{1 - B(0,T_n)}{A_0^{0,n}},$$

which is equivalent to

$$\begin{aligned}
S_0^{0,n} A_0^{0,n} &= 1 - B(0,T_n) \\
\Leftrightarrow S_0^{0,n}\left(A_0^{0,n-1} + \delta_{n-1} B(0,T_n)\right) &= 1 - B(0,T_n) \\
\Leftrightarrow B(0,T_n) &= \frac{1 - S_0^{0,n} A_0^{0,n-1}}{1 + \delta_{n-1} S_0^{0,n}} \\
\Leftrightarrow B(0,T_n) &= \frac{1 - S_0^{0,n} \sum_{i=0}^{n-2} \delta_i B(0,T_{i+1})}{1 + \delta_{n-1} S_0^{0,n}}. \qquad (2.13)
\end{aligned}$$

On the left hand side of (2.13) we find the n-th bond price and on the right hand side we find a function depending on the n-th swap rate and the first $n-1$ bond prices. Therefore, the formula gives us a way to calculate the bond prices one by one by just knowing the swap rates $S_0^{0,n}$ for each tenor point T_n. If we calculated all bond prices, we can calculate the forward rates through the formula in (2.1)

$$F_0^i = \frac{1}{\delta_i} \frac{B(0,T_i) - B(0,T_{i-1})}{B(0,T_{i-1})}.$$

However, not all needed swap rates can be found in the market and have to be interpolated. We decided to interpolate linear. This method doesn't guarantee positive forwards, but in our case we did not get any.

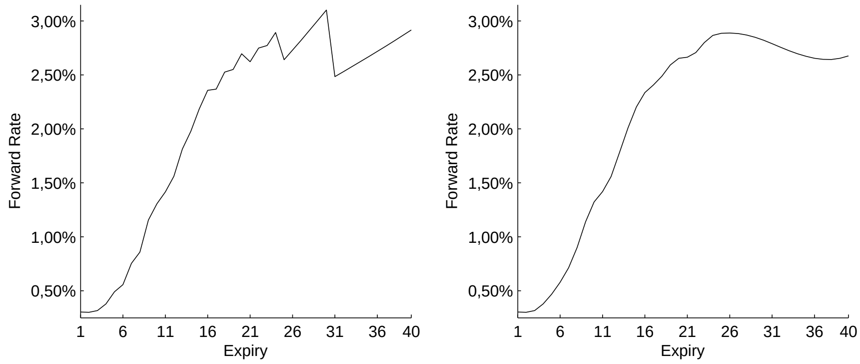

Figure 2.1.: The stripped forward rates following the approach above. On the left we used linear interpolation and get shape that is in [14] reffered to as a *saw tooth* shape. On the right we used spline interpolation and got a smoother shape, but a smaller maximal forward rate. The data is from the 21.07.2014 and was obtained from Bloomberg. The stripping was implemented in F# and the plot was done in Matlab.

Still, there are other possibilities. In [14] the (C^1/C^2) spline interpolation is suggested and explained, but this method doesn't guarantee positive forwards either. Another possibility is the *Forward Monoton Convex Spline* introduced in [33]. This method incorporates the idea of occurred interest, meaning that a forward rate F^i is paid over the interval $[T_i, T_{i+1}]$ and not only at T_{i+1}.
All the above methods work only in an environment of greater certainty about the input data. If the validity of the data is questionable one could build a swap curve by using a Nelson-Siegel or Svensson curve as described in [1] and [16]. Those curves have a parametrization that forces them in a range of idealized swap curves. Since those interpolation methods are behind the scope of this work we stick to linear interpolation.

2.2.2. The Stripping of Caplet Volatilities

Caplet volatilities will be one of the corner stones of our calibration procedure later on. As described in section 2.1.2 caplets are not directly quoted in the market, but indirectly as caps. In this section we will describe a stable approach to calculate caplet volatilities from cap volatilities. This procedure is called *caplet stripping*. For the general framework we rely on [24].

In the market we find for each set of strikes $(K_i)_i$ a set of cap prices, for caps $\big(C^{1,j}(K_i)\big)_{i,\{1<j\leq N\}}$ with expiry date T_1 growing tenors up to $T_N - T_1$, given in Black volatilities $\sigma^{\text{cap}}(j, K_i)$. So in the market caps are quoted indirectly. The cap price can be obtained via

$$\begin{aligned} C^{1,j}(K_i) &= \sum_{k=1}^{j-1} C^k(K_i) \\ &= \sum_{k=1}^{j-1} C^k(F_0^k, K_i, \sigma^{\text{cap}}(j, K_i), T_k), \end{aligned} \tag{2.14}$$

where $C^k(F_0^k, K_i, \sigma(j, K_i), T_k)$ is the price of the k-th caplet assuming that F_t^k follows Black's model [10]. Therefore, it holds due to (2.7)

$$C^k(F_0^k, K_i, \sigma(j, K_i), T_k) = \delta_k B(0, T_{k+1})(F_0^k \mathcal{N}(d_1) - K_i \mathcal{N}(d_2)), \tag{2.15}$$

where

$$d_{1/2} := \frac{\ln\big(\frac{F_0^k}{K_i}\big) \pm \frac{1}{2}(\sigma(j, K_i))^2 T_k}{\sigma(j, K_i)\sqrt{T_k}}.$$

So a cap is priced by using an *all in* volatility $\sigma^{cap}(j, K_i)$ for all caplets. Knowing this we want to calculate the caplet prices for all caplets $\big(C^j(K_i)\big)_{j,i}$ in Black volatilities $\big(\sigma^{\text{cpl}}(j, K_i)\big)_{j,i}$ at any tenor point T_j and strike K_i.
To achieve this, we first fix a strike K_i and therefore only consider the set of caps $\big(C^{1,j}(K_i)\big)_{\{1<j\leq N\}}$. As in the case for forward rates the stripping of caplet volatilities is done iteratively as follows: It is clear from (2.8) that the cap price $C^{1,2}(K_i)$ agrees with the caplet price $C^1(K_i)$. This is our initial value. Then we solve iteratively the following equations for $1 < k < N$

$$C^k(F_0^k, K_i, \sigma^{\text{cpl}}(j, K_i), T_k) = C^{1,k+1}(K_i) - \sum_{j=1}^{k-1} C^j(F_0^j, K_i, \sigma^{\text{cpl}}(j, K_i), T_j) \tag{2.16}$$

to obtain all caplet volatilities $\sigma^{\text{cpl}}(j, K_i)$. We do this for all strikes and get the whole caplet volatility surface.
Similar as in the bootstrapping of forward rates not all cap volatilities for all tenors we are interested in may exist in the market. We gain the missing tenors by spline interpolation, since we want a smooth volatility surface.

Remark. It is clear from the definition of the payoff of a floorlet (2.9) and from the definition of the price of a floor (2.10): It is possible to strip floorlet volatilities $\left(\sigma^{\text{flt}}(j, K_i)\right)_{j,i}$ in the same fashion as stripping caplet volatilities. The floorlet volatilities $\sigma^{\text{flt}}(j, K_i)$ then agree with the caplet volatilities $\sigma^{\text{cpl}}(j, K_i)$ for the same strikes, underlying prices and expiries, because of the call-put parity.

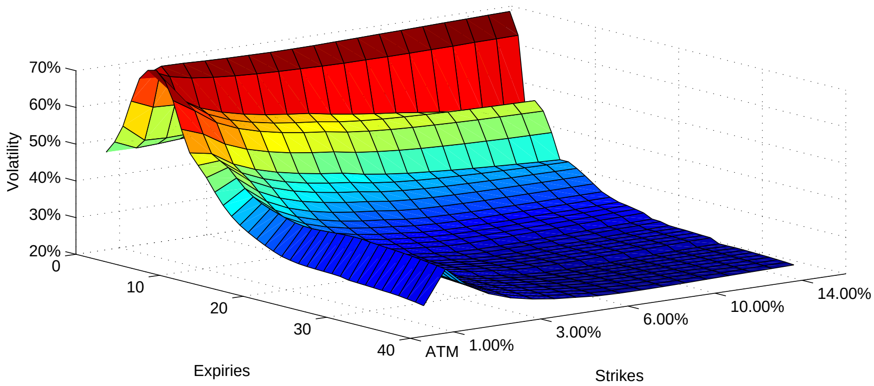

Figure 2.2.: The volatility surface stripped from Euro caps as of the 21.07.2014 with a half year tenor ($\delta_i \equiv 0.5$) and a time horizon of over 20 years. We obtained the data from Bloomberg. The implemention was carried out in F# and the plot was done in Matlab.

3. The SABR and SABR-LMM model

The main goal of this work is to provide an environment to price options depending on a range of forward rates consistently. To achieve that we set up a model that can do both. It will incorporate the observable smile effect and provide a dependency structure for the modeled assets. We will develop the model in two steps. In the first one we give a simple model that can only handle one asset, but is capable of incorporating the smile effect. Further, the model provides a analytic formula that translates the coefficients which are describing the model into an implied model smile. This feature will come in handy later on, since it will enable us to calibrate efficiently to given prices. Then we extend the model to a full market model. By doing so we try to preserve the dynamics from the simpler model as well as possible.

3.1. The SABR Model

The SABR (σ, α, β, ρ) model is a model of stochastic volatility and can describe exactly one asset F. It was first published by *Hagan* [18] and it has been popular since, because it easy to understand and to calibrate. The stochastic volatility gives a way to incorporate the market smile. Further there exists a formula that gives depending of the model parameters the smile generated by the model, whereas a change of model parameters can be directly interpreted in changes of the model induced smile in a logical way. In addition there exist very effective simulation schemes which reduce the pricing procedure through Monte Carlo simulations drastically [6]. For this reason it is a perfect tool to manage a book of options on a single asset. Theoretically F can be any asset but in our context F will be only a forward rate or a swap rate.

Definition 7. *(The SABR Model)*
In the SABR model the dynamics of an asset F maturing at time T is given by

$$\begin{aligned} dF_t &= \sigma_t F_t^\beta dW_t, \quad F_0 = F(0), t \leq T \\ d\sigma_t &= \alpha \sigma_t dZ_t, \quad \sigma_0 = \sigma(0) \\ d\langle W_\bullet, Z_\bullet \rangle_t &= \rho dt, \end{aligned} \tag{3.1}$$

where $\sigma,\ \alpha \in \mathbb{R}^+$ und $\beta \in [0,1], \rho \in [-1,1]$. Further W and Z are one dimensional Wiener processes. To be consistent we define $dF_t \equiv 0$ for $t > T$.

The process σ_t is the *volatility* of the model and σ_0 as the *level* of the volatility. The coefficient α is named the *vol-vol*, that's short for the volatility of volatility and ρ is the so called correlation or *skew*.
The coefficient β is the CEV parameter. In the special case $\beta = 0$ we get a normal model, since in this case F_t is approximately normally distributed. With this choice for β the process $(F_t)_t$ can get negative and in Asian markets practitioners often choose this parametrization to model forwards, since they tend to be negative in this markets from time to time [18]. If we set $\beta = 1$, we obtain a log-normal model. If $\beta \in (0,1)$, we get a CEV model. The choice of beta will be important in the calibration part later in this work.
We do not favor the SABR model only because of the simple structure above other ones like *Heston* [26] or *Bates* [37], which also could be extended to a full Libor-Market-Model [12]. The main advantage of the SABR model over the other ones is the analytic function for implied Black volatility depending on strike and underlying price. If this volatility is put in to Blacks pricing formula, it yields the model price for a call. To describe the implied Black volatility closer, we first consider Black's Model [10] in which an asset F follows the SDE

$$dF_t = \sigma^{\text{imp}} F_t dW_t, \quad F_0 = F(0),$$

where $\sigma^{\text{imp}} > 0$ is a real number and W a Wiener process. It is well known [10] that the call price for a call with strike K, expiry date T_{ex} and settlement date T_{set} can be calculated as

$$C(F_0, K) = B(0, T_{\text{set}}) \Big[F_0 \mathcal{N}(d_1) - K \mathcal{N}(d_2) \Big] \tag{3.2}$$

and, respectively the put price for a put with strike K, expiry date T_{ex} and settlement date T_{set} can be calculated as

$$P(F_0, K) = B(0, T_{\text{set}})\Big[K\mathcal{N}(-d_2) - F_0\mathcal{N}(-d_1)\Big],$$

where

$$d_{1/2} = \frac{\ln\left(\frac{F_0}{K}\right) \pm (\sigma^{\text{imp}})^2 T_{ex}}{\sigma^{\text{imp}}\sqrt{T_{\text{ex}}}} \tag{3.3}$$

and $B(0, T_{\text{set}})$ is the discount factor for the time interval $[0, T_{set}]$. The volatility σ^{imp} which yields in (3.2) and (3.3) the same price for a call and put the SABR model would produce is called *implied (Black) volatility.* It is well known that in real market situations the implied volatility depends on the underlying price and strike. Thus, a volatility surface can be observed which reduces to a volatility smile, if we fix the underlying price F_0 to some value. In the SABR model exists a formula to calculate the implied volatility surface based on the model parameters σ_0, α, β and ρ. The formula is given in [18] and was improved in [23]. The improved formula goes as follows

$$\begin{aligned}\sigma_{\text{I}}(F, K, \beta, \alpha, \nu, \rho, T_{\text{exp}}) :=& I_{\text{H}}^0(F, K, \beta, \alpha, \nu, \rho) \\ & \times (1 + I_{\text{H}}^1(F, K, \beta, \alpha, \nu, \rho)T_{\text{exp}}),\end{aligned} \tag{3.4}$$

where

$$I_{\text{H}}^1(F, K, \beta, \alpha, \nu, \rho) := \frac{(\beta-1)^2\alpha^2}{24(F_0K)^{1-\beta}} + \frac{\rho\nu\alpha\beta}{4(F_0K)^{(1-\beta)/2}} + \frac{2-3\rho^2}{24}\nu^2$$

and

$$I_{\text{H}}^0(F, K, \beta, \alpha, \nu, \rho) := \begin{cases} \alpha K^{\beta-1} & \text{, if } \frac{F_0}{K} = 1 \\ \frac{\ln(F_0/K)\alpha(1-\beta)}{F_0^{1-\beta} - K^{1-\beta}} & \text{, if } \nu = 0 \\ \nu\ln(\frac{F_0}{K}) / \ln\left(\frac{\sqrt{1-2\rho z_1+z_1^2}+z_1-\rho}{1-\rho}\right) & \text{, if } \beta = 1 \\ \nu\ln(\frac{F_0}{K}) / \ln\left(\frac{\sqrt{1-2\rho z_2+z_2^2}+z_2-\rho}{1-\rho}\right) & \text{, if } \beta < 1 \end{cases},$$

where

$$z_1 := \frac{\nu\ln(F_0/K)}{\alpha}$$

and

$$z_2 := \frac{\nu}{\alpha} \frac{F_0^{1-\beta} - K^{1-\beta}}{1-\beta}.$$

All the above expressions are purely analytic and no numerical integrations or some

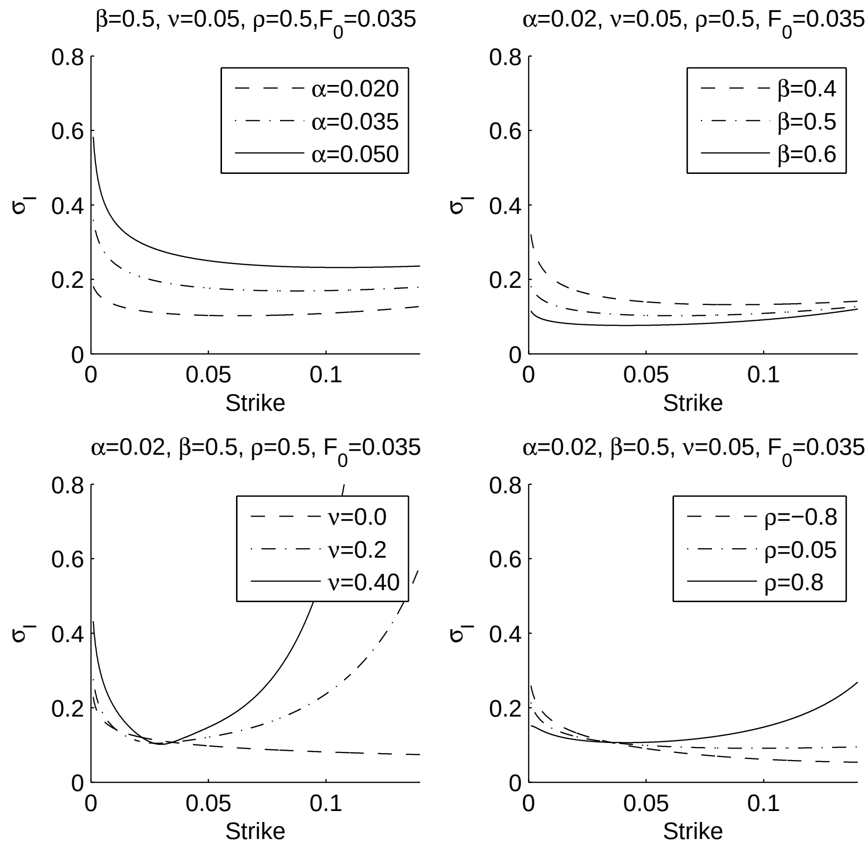

Figure 3.1.: The graphic shows how the implied volatility σ_{I} changes with changes of the SABR model parameters. α has an impact on the level the smile, whereas a higher ν produces a more pronounced shape. A change in β effects the left end of the smile, that is in the area with small strikes. The parameter ρ has a general impact on the *skew* of the smile. The plots give the impression that changes in ν and ρ can substitute changes in β very well. The plot was done in Matlab.

similar cumbersome procedures are needed. This makes this formula highly tractable and efficient.
However, we want to emphasis that there exist other approximation for the implied volatility. For example, other formulations are given in [38], [30] and [27], whereas the formulation in the last source is the most exact one according to market opinion. But, the implementation of those significant more complicate formulas is behind the scope of this work. Note however that it was shown in [18] and [23] that the above version in (3.4) works quite well.
Clearly, the formula enables us to calculate prices, which our model produces, for puts and calls with different strikes and underlying prices without doing cumbersome Monte Carlo simulations. Further, the formula enables us to calculate prices on portfolios of put and calls, like straddles, butterfly spreads, covered calls, protective puts, etc.
But, we can go the other way around as well. It is market practice to quote prices of calls and put in Black volatilities indirectly. Hence, if we observe implied volatilities of European options we can easily calibrate the SABR model to market prices by minimizing the difference of quoted volatilities and implied model volatilities depending on the model parameters α, β, ν and ρ.
However, in this work we will fix β to 0.5 or 1.0, depending on the assets we are looking at. So we only have to estimate the three parameters α, ν and ρ. We have two reasons to do so. First, the impact of β and ρ, in combination with ν, on the shape of the curve is very similar as can be seen in figure 3.1. By fixing β we obtain a more unique solution. Second, we want to model forward rates and to set $\beta = 0.5$ seems to be market conform, as argued in [36]. Discussion with traders showed that most of them indeed choose $\beta = 0.5$ in their CEV models and β fixed at this value leads to a lower variation of the other parameters over time. So the model calibration is longer approximately valid and a longer validity speaks in favor of a fixed β.
To estimate the SABR model parameters we simply minimize the square of the sum over the squared errors between market prices and model prices. That implies our estimated parameters $\widehat{\alpha}$, $\widehat{\rho}$ and $\widehat{\nu}$ are obtained by

$$(\widehat{\alpha}, \widehat{\rho}, \widehat{\nu}) = \underset{\alpha,\rho,\nu}{\arg\min} \sqrt{\sum_i \left[\sigma_{\mathrm{M}}(F_0, K_i) - \sigma_{\mathrm{I}}(F_0, K_i, \beta, \alpha, \nu, \rho, T_i)\right]^2}, \tag{3.5}$$

where $\sigma_{\mathrm{M}}(F_0, K_i)$ is the in the market quoted Black volatility for a call or put with strike K_i, underlying price F_0 and expiry date T_i. The minimization problem in (3.5) can be tailored to ones needs by multiplying a weights. This technique can be used to weight uncertain data lower than certain one or to emphasis on a range of strikes. In this case the minimization problem becomes

$$(\widehat{\alpha}, \widehat{\rho}, \widehat{\nu}) = \underset{\alpha,\rho,\nu}{\arg\min} \sqrt{\sum_i \omega_i \big[\sigma_{\mathrm{M}}(F_0, K_i) - \sigma_{\mathrm{I}}(F_0, K_i, \beta, \alpha, \nu, \rho, T_i)\big]^2}. \tag{3.6}$$

For example, by choosing $(\omega_i)_i = (\sigma_{\mathrm{M}}(F_0, K_i)^{-1})_i$ the relative differences will be minimized. If not written differently we will use (3.5).

3.2. The SABR-LMM Model

In this chapter we will combine the simple SABR model with the classic Libor market model (LMM) under deterministic volatility as developed in [11] and [2]. In a LMM a number of forward rates with a dependency structure are modeled. The dependency structure is given through the correlation which describes the comovement of the assets in the model. The problematic part in a simple LMM is the lack of possibility to model smile effect which is observable in the real market. This means we are only able to evaluate caplets or swaptions on those strikes which are used for the model calibration and, even worse, it is only possible to evaluate caplets or swaptions for *exactly* one strike. In most cases those strikes are at the money. Obviously a reasonable model should be able to price options on any strike. The SABR model can reproduce the smile, but since it is a one-asset model no dependency of two or more processes can be considered. It is definitely no good solution to model a number of assets simply by taking a number of uncorrelated SABR processes. For example, this shows the valuation of swaptions based on forward rates.

So the LMM and the SABR together have the needed features plus the SABR model gives us the useful formula for the implied volatility. In the following we will combine both models and develop two stable calibration methods. The first method will be a calibration on caplets and swaptions and the second will be a calibration on caplets and CMS spread options. In both cases we will heavily depend on the formula for implied volatility to hit quoted market prices. The overall goal in both approaches will be to keep the SABR dynamics for the forward rates as close as possible, since that model has so many good characteristics. The SABR-LMM is defined as follows:

Definition 8. *(The SABR-LMM Model for Forward Rates)*
In a N-dimensional SABR-LMM model the N forward rates $(F^i)_{i\in\{1,\dots,N-1\}}$ have under their individual forward measure $\mathbb{P}^i$ the following dynamics:

$$dF_t^i = \sigma_t^i (F_t^i)^\beta dW_t^i, \quad t < T_i, \quad F_0 = F(0) \tag{3.7}$$
$$\sigma_t^i = g_t^i k_t^i \tag{3.8}$$
$$dk_t^i = h_t^i k_t^i dZ_t^i, \quad t < T_i, \quad k_0 = k(0) \tag{3.9}$$
$$d\langle W^i, W^j\rangle_t = \rho_{ij} dt, \ i,j \in \{1,\dots,N-1\} \tag{3.10}$$
$$d\langle Z^i, Z^j\rangle_t = r_{ij} dt, \ i,j \in \{1,\dots,N-1\} \tag{3.11}$$
$$d\langle W^i, Z^j\rangle_t = R_{ij} dt, \ i,j \in \{1,\dots,N-1\}, \tag{3.12}$$

where $\beta \in [0,1]$, $\rho_{i,j}, r_{i,j}, R_{i,j} \in [-1,1]$ for all $i,j \in \{1,\dots,N-1\}$ and the deterministic functions $g,h : \mathbb{R}_+ \to \mathbb{R}$ fulfill

$$\int_0^T g_i^2(s)ds < \infty \text{ and } \int_0^T h_i^2(s)ds < \infty \text{ for all } i \in \{1,\dots,N-1\} \text{ and } 0 < T \le T_i.$$

Further, we set for completeness

$$F_t^i = F_{T_i}^i \text{ for all } t > T_i.$$

We define the super correlation matrix of the model as

$$P := \begin{pmatrix} \rho & R \\ R^T & r \end{pmatrix}. \tag{3.13}$$

The Matrix $(\rho_{ij})_{i,j}$ consists of all forward/forward correlations, the entries of $(r_{ij})_{i,j}$ are the volatility/volatility correlations and $(R_{ij})_{i,j}$ carries all the forward/volatility correlations. Notice, only P, $(\rho_{ij})_{i,j}$ and $(r_{ij})_{i,j}$ are symmetric. The matrix $(R_{ij})_{i,j}$ is asymmetric in general.

Remark. From time to time we will use the forward rate F^0 which is not contained in the SABR-LMM above. This forward rate is the interest rate for the period $[T_0, T_1]$. Since we assume that all prices in T_0 are known F^0 is not stochastic. Obviously, its dynamics doesn't have to be modeled.

The SABR-LMM incorporates not only the SABR into the LMM model it has even

time dependent parameters given through g^i and h^i. The stochastic volatilities σ^i of the forward rates F^i can be separated into a deterministic part g^i and in a stochastic part k^i. Therefore g^i is often called the deterministic volatility of F^i and k^i the stochastic volatility, respectively. Further, the function h^i describes the deterministic volatility of volatility.
We would like to highlight the following feature: If k^i is constant for all i, e.g. $h^i \equiv 0$ for all i, we obtain an ordinary LMM. This is because the stochastic volatility vanishes.

3.3. The SABR-LMM Dynamics under any Forward Measure $\mathbb{P}^l$

To simulate option prices in a Monte Carlo setup or to examine the model it is necessary to express all asset dynamics in the model under one common measure. A possible choice for such a measure is the forward measure $\mathbb{P}^l$ for some l. In the special case $l = N-1$ we call $\mathbb{P}^{N-1}$ the *terminal* measure and in most cases we use this measure for our simulation. In the following, we always assume that our modeled assets are forward rates with dynamics as given in definition 8. The idea is to calculate the change of measures by change of numéraire techniques.

Theorem 1 (SABR-LMM Dynamics under different $\mathbb{P}^l$).
In the SABR-LMM model, as in definition 8, the dynamics under a certain forward measure $\mathbb{P}^l$ of the forward rates F^j and the stochastic volatilities k^j are given as

$$dF_t^j = \sigma_t^j (F_t^j)^\beta \times \begin{cases} -\sum_{j+1 \le i \le l} \frac{\rho_{i,j}\delta_i \sigma_t^i (F_t^i)^\beta}{1+\delta_i F_t^i} dt + dW_t^j & , \text{ if } j < l \\ dW_t^j & , \text{ if } j = l \\ \sum_{l+1 \le i \le j} \frac{\rho_{i,j}\delta_i \sigma_t^i (F_t^i)^\beta}{1+\delta_i F_t^i} dt + dW_t^j & , \text{ if } j > l \end{cases} \tag{3.14}$$

and

$$dk_t^j = h_t^j k_t^j \times \begin{cases} -\sum_{j+1 \le i \le l} \frac{R_{j,i}\delta_i \sigma_t^i (F_t^i)^\beta}{1+\delta_i F_t^i} dt + dZ_t^j & , \text{ if } j < l \\ dZ_t^j & , \text{ if } j = l \\ \sum_{l+1 \le i \le j} \frac{R_{j,i}\delta_i \sigma_t^i (F_t^i)^\beta}{1+\delta_i F_t^i} dt + dZ_t^j & , \text{ if } j > l \end{cases} \tag{3.15}$$

where $\sigma_t^j = h_t^j k_t^j$ stays the same.

Proof. We will carry out the proof with by means of induction. First, we concentrate on the dynamics of the F^i. It holds per definition, since F^i is a forward rate:

$$F_t^i = \frac{1}{\delta_i}\Big(\frac{B(t,T_i) - B(t,T_{i+1})}{B(t,T_{i+1})}\Big) \text{ for all } t \leq T_i,$$

where the $B(t,T_i)$ are strictly positive bond prices at time t for bonds which pay at maturity T_i exactly one unit of money. Further, the F^i are local martingales under $\mathbb{P}^i$. Therefore the probability measure $\mathbb{P}^i$ can be seen as the measure under which every tradeable asset relative to the numéraire $B(t,T_{i+1})$ is a local martingale. In the first step we calculate the dynamics of F_t^i under $\mathbb{P}^{i-1}$ and therefore relatively to $B(t,T_i)$. For this we need the Bayes formula. We give the formula without proof.

Proposition 4. *(Bayes' Formula)*
Let $(\Omega, \mathcal{A})$ be a measurable space with probability measures $\mathbb{P}$ and $\mathbb{Q}$. Further let $\mathcal{B} \subseteq \mathcal{A}$ be a sub-σ-algebra. Then it holds for an integrable and $\mathcal{A}$ measurable random variable Y

$$\mathbb{E}^{\mathbb{Q}}\big[Y \mid \mathcal{B}\big] = \mathbb{E}^{\mathbb{P}}\big[\frac{d\mathbb{Q}}{d\mathbb{P}} Y \mid \mathcal{B}\big] \ \Big(\mathbb{E}^{\mathbb{P}}\big[\frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathcal{B}\big]\Big)^{-1} \quad \mathbb{P}\text{-}a.s. \ . \tag{3.16}$$

It follows that

$$\frac{d\mathbb{P}^{i-1}}{d\mathbb{P}^i}\Big|_{\mathcal{F}_t} = \frac{B(t,T_i)}{B(t,T_{i+1})} \frac{B(0,T_{i+1})}{B(0,T_i)} \quad \mathbb{P}^i\text{-a.s. } , \tag{3.17}$$

since the expression is a probability measure, because $(1 + \delta_i F_t^i)\frac{B(0,T_i)}{B(0,T_{i+1})} = \frac{d\mathbb{P}^{i-1}}{d\mathbb{P}^i}\big|_{\mathcal{F}_t}$ is positive $\mathbb{P}^{i+1}$-martingale with an expected value of 1. Further, let $\Big(\frac{X_t}{B(t,T_i)}\Big)_{0 \leq t \leq T_i}$ be a $\mathbb{P}^{i-1}$ martingale. Then it holds with the Base formula (3.16) for $t \leq T_i$

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}^i}\Big[\Big(\frac{d\mathbb{P}^{i-1}}{d\mathbb{P}^i}\Big)\frac{X_{T_i}}{B(T_i,T_i)} \mid \mathcal{F}_t\Big] &= \mathbb{E}^{\mathbb{P}^i}\Big[\Big(\frac{B(T_i,T_i)}{B(T_i,T_{i+1})}\frac{B(0,T_{i+1})}{B(0,T_i)}\Big)\frac{X_{T_i}}{B(T_i,T_i)} \mid \mathcal{F}_t\Big] \\
&= \mathbb{E}^{\mathbb{P}^{i-1}}\Big[\frac{X_{T_i}}{B(T_i,T_i)} \mid \mathcal{F}_t\Big] \\
&\times \mathbb{E}^{\mathbb{P}^i}\Big[\frac{B(T_i,T_i)}{B(T_i,T_{i+1})}\frac{B(0,T_{i+1})}{B(0,T_i)} \mid \mathcal{F}_t\Big] \\
&= \frac{X_t}{B(t,T_i)}
\end{aligned}$$

This shows, that the measure implied by density in (3.17) agrees with the probability measure $\mathbb{P}^{i-1}$. Therefore the notation $\frac{d\mathbb{P}^{i-1}}{d\mathbb{P}^i}\big|_{\mathcal{F}_t}$ is justified. Now it follows with the

Ito-Formulas [9] and by considering the SABR-LMM model dynamics from definition 8:

$$\begin{aligned} d\Big[\ln\Big(\frac{d\mathbb{P}^{i-1}}{d\mathbb{P}^{i}}\big|_{\mathcal{F}_t}\Big)\Big] &= d\ln((1+\delta_i F_t^i)\frac{B(0,T_i)}{B(0,T_{i+1})}) \\ &= d\ln(1+\delta_i F_t^i) \\ &= \frac{\delta_i}{1+\delta_i F_t^i} dF_t^i - \frac{1}{2}\frac{\delta_i^2}{1+\delta_i F_t^i} d\langle F^i\rangle_t \\ &= \frac{\delta_i}{1+\delta_i F_t^i}\sigma_t^i (F_t^i)^\beta dW_t^i - \frac{1}{2}\frac{\delta_i^2}{1+\delta_i F_t^i}(\sigma_t^i)^2 (F_t^i)^{2\beta} dt. \end{aligned} \tag{3.18}$$

According to Girsanow's Theorem [40] is X a local $\mathbb{P}^i$-martingale if and only if

$$Y := X - \langle X, \ln\Big(\frac{d\mathbb{P}^{i-1}}{d\mathbb{P}^{i}}\big|_{\mathcal{F}_\bullet}\Big)\rangle \tag{3.19}$$

is a local $\mathbb{P}^{i-1}$-martingale. A change of measure produces a drift that maintains the martingale property. In the finance literature this drift is often called *Convexity Correction*.

If we use (3.19) on F^i, we get, with the help of (3.18) and (3.10), the following dynamics under $\mathbb{P}^{i-1}$

$$\begin{aligned} d\widehat{F}^i &= dF^i - d\langle F^i, \ln\Big(\frac{d\mathbb{P}^{i-1}}{d\mathbb{P}^{i}}\big|_{\mathcal{F}_\bullet}\Big)\rangle \\ &= \sigma^i (F^i)^\beta dW^i - \sigma^i (F^i)^\beta \frac{\delta_i}{1+\delta_i F^i}\sigma^i (F^i)^\beta. \end{aligned} \tag{3.20}$$

Now we want to calculate the dynamics of F^i under $\mathbb{P}^{i-2}$. In analogy (3.18) to follows

$$\begin{aligned} d\Big[\ln\Big(\frac{d\mathbb{P}^{i-2}}{d\mathbb{P}^{i-1}}\big|_{\mathcal{F}_t}\Big)\Big] &= d\ln(1+\delta_{i-1} F_t^{i-1}) \\ &= \frac{\delta_{i-1}}{1+\delta_{i-1} F_t^{i-1}}\sigma_t^{i-1} (F_t^{i-1})^\beta dW_t^{i-1} \\ &\quad - \frac{1}{2}\frac{\delta_{i-1}^2}{1+\delta_{i-1} F_t^{i-1}}(\sigma_t^{i-1})^2 (F_t^{i-1})^{2\beta} dt. \end{aligned} \tag{3.21}$$

Together with (3.20) follows for F^i under $\mathbb{P}^{i-2}$ by considering $\frac{d\mathbb{P}^{i-2}}{d\mathbb{P}^i} = \frac{d\mathbb{P}^{i-2}}{d\mathbb{P}^{i-1}}\frac{d\mathbb{P}^{i-1}}{d\mathbb{P}^i}$

$$\begin{aligned}
d\widetilde{F}^i &= dF^i - d\langle F^i, \ln\Big(\frac{d\mathbb{P}^{i-2}}{d\mathbb{P}^i}\big|_{\mathcal{F}_\bullet}\Big)\rangle \\
&= d\widehat{F}^i - d\langle \widehat{F}^i, \ln\Big(\frac{d\mathbb{P}^{i-2}}{d\mathbb{P}^{i-1}}\big|_{\mathcal{F}_\bullet}\Big)\rangle \\
&= \sigma^i (F^i)^\beta \Big(dW^i - \Big(\frac{\delta_i}{1+\delta_i F^i}\sigma^i (F^i)^\beta \rho_{i,i} + \frac{\delta_{i-1}}{1+\delta_{i-1}F_t^{i-1}}\sigma_t^{i-1}(F_t^{i-1})^\beta \rho_{i-1,i}\Big)\Big)
\end{aligned}$$

Now follows the theorem for F^j in the case of $j > l$. The case $l < j$ follows in the same way and we only note: It holds

$$\frac{d\mathbb{P}^i}{d\mathbb{P}^{i-1}}\big|_{\mathcal{F}_t} = \Big(\frac{d\mathbb{P}^{i-1}}{d\mathbb{P}^i}\big|_{\mathcal{F}_t}\Big)^{-1} = \Big(\frac{B(t,T_i)}{B(t,T_{i+1})}\frac{B(0,T_{i+1})}{B(0,T_i)}\Big)^{-1} \quad \mathbb{P}^{i-1}\text{-a.s. }.$$

The dynamics of the stochastic volatilies k^i under the forward measure $\mathbb{P}^i$ is through (3.9) as

$$dk^i = h^i k^i dZ^i.$$

In the same fashion as for the forward rates we first calculate the stochastic differential equation of k^i under the measure $\mathbb{P}^{i-1}$ and $\mathbb{P}^{i-2}$.
With Girsanow (3.19) follows, with the help of (3.18), for the dynamics of k^i under the measure $\mathbb{P}^{i-1}$

$$\begin{aligned}
d\widehat{k}^i &= dk^i - d\langle k^i, \ln\Big(\frac{d\mathbb{P}^{i-1}}{d\mathbb{P}^i}\big|_{\mathcal{F}_\bullet}\Big)\rangle \\
&= h^i k^i dZ^i - h^i k^i \frac{\delta_i}{1+\delta_i F^i}\sigma^i (F^i)^\beta R_{i,i}.
\end{aligned}$$

Therefore, we obtain by considering (3.21) for the dynamics of k^i under $\mathbb{P}^{i-2}$

$$\begin{aligned}
d\widetilde{k}^i &= dk^i - d\langle k^i, \ln\Big(\frac{d\mathbb{P}^{i-2}}{d\mathbb{P}^i}\big|_{\mathcal{F}_\bullet}\Big)\rangle \\
&= h^i k^i \Big(dZ^i - \Big(\frac{\delta_i}{1+\delta_i F^i}\sigma^i (F^i)^\beta R_{i,i} + \frac{\delta_{i-1}}{1+\delta_{i-1}F^{i-1}}\sigma^{i-1}(F^{i-1})^\beta R_{i,i-1}\Big)\Big).
\end{aligned}$$

Again, per induction follows the theorem for k^j in the case of $j > l$. The case $l < j$ follows analogously and therefore is omitted. $\square$

Remark. Notice that the calculated model dynamics in Theorem 1 don't agree with the ones in [36]. There we find in the dynamics of k^j instead of $R_{i,j}$ the Term $R_{i,i}\rho_{i,j}$ and the function g^j. However our version coincides with the dynamics in [17].

3.4. The SABR-LMM Dynamics under the Spot Measure $\mathbb{P}^{\text{spot}}$

Another measure under which we can calculate the dynamics of the Forwardrates is the Spot Measure $\mathbb{P}^{\text{spot}}$. In this measure processes of the form $(\frac{X_t}{G_t})_t$ are local martingales, where $G_t := \frac{B(t,T_{\gamma(t)-1})}{\prod_{1\leq i\leq\gamma(t)-1} B(T_{i-1},T_i)}$, and $\gamma(t) := \inf\big\{k \in \mathbb{N} \mid T_0 + \sum_{i=0}^{k-1}\delta_i > t\big\} = \inf\big\{k \in \mathbb{N} \mid T_k \geq t\big\}$.

Theorem 2 (SABR-LMM Dynamics under $\mathbb{P}^{\text{spot}}$)**.**
Under $\mathbb{P}^{spot}$ the SABR-LMM dynamics given in definition 8 are the following:

$$dF_t^j = \sigma_t^j (F_t^j)^\beta \Big(\sum_{\gamma(t)\leq i\leq j} \frac{\rho_{i,j}\delta_i\sigma_t^i(F_t^i)^\beta}{1+\delta_i F_t^i} + dW_t^j \Big), \tag{3.22}$$

and

$$dk_t^j = h_t^j k_t^j \Big(\sum_{\gamma(t)\leq i\leq j} \frac{r_{j,i}\delta_i g_t^i h_t^i k_t^i (F_t^i)^\beta}{1+\delta_i F_t^i} dt + dZ_t^j \Big), \tag{3.23}$$

where $\sigma_t^j = h_t^j k_t^j$ stays the same.

Proof. A proof can be found in [7]. Alternatively one can carry out the proof in analogy to Theorem 1. Since the numéraires of $\mathbb{P}^{\text{spot}}$ and $\mathbb{P}^l$ are known one can calculate the density for the change of measure, like in (3.17). Then just the drifts comming from Grisanovs theorem have to be calculated to obtain the dynamics under the spot measure. □

To interpret $\mathbb{P}^{\text{spot}}$ we write G_t in a different way. It holds

$$\begin{aligned} G_t &= \frac{B(t, T_{\gamma(t)-1})}{\prod_{1 \le i \le \gamma(t)-1} B(T_{i-1}, T_i)} \\ &= \prod_{1 \le i \le \gamma(t)-1} (1 + \delta_{i-1} F_{T_{i-1}}^{i-1}) B(t, T_{\gamma(t)-1}). \end{aligned}$$

So G_t can be seen as the time value process of a portfolio with the following strategy: The portfolio value in the beginning is exactly one. Then, from period to period, the portfolio reinvests its capital with the actual one period spot rate. To get the time value at time t the portfolio value is discounted by $B(t, T_{\gamma(t)-1})$.

The reason for considering different measures is the effort of calculating the drifts terms in simulations. Almost half of the simulation time comes from the drift calculation the other half comes from generating random numbers. In the spot measure the processes F_t^j and k_t^j have drifts consisting of $(j - \gamma(t) + 1)$ summands as shown in (3.22) and (3.23), respectively. In the terminal forward measure the processes F_t^j and k_t^j have drifts consisting of $(N - j)$ summands as shown in (3.14) and (3.15). It is natural to choose the measure with the minimal cost of drift calculation. We conclude the following thumb rule: If only forwards with short expiries have to be simulated, we choose the spot measure and if forwards with longer expiries are involved, we choose the terminal forward measure.

4. Swaps Rates in the SABR-LMM

Swap rates depend directly on underlying forward rates, since we can write them as sum of forwards as shown in Section 2.1.1 in equation (2.5). This structure particularly yields a direct dependence of the swap rate on the interplay of the forward rates. We want to analyze how the dependence of the interplay can be described in terms of the super correlation matrix P which we defined in (3.13).
To achieve this, we first give a way to model the swap rate dynamics in a SABR environment. Then we approximate the swap's SABR coefficients by taking the structure as a sum of forward rates into account. Here we assume that the forward rate dynamics are governed by the SABR-LMM. The approximated SABR coefficients will depend on P. By doing this we find a proper way to describe a swap rate dependent on P, which we will later use to estimate the matrix implicitly by using market quotes of swaption prices. More on this can be found in chapter 6 which covers calibration to swaptions.

4.1. A SABR model for Swap Rates

A swap rate depends directly on forward rates, because of (2.5). Since we chose SABR-like dynamics for all the forwards it makes sense to assume that a swap rate does not evolve in a completely different style and can be described by a SABR model under the swap measure $\mathbb{P}^{m,n}$ as well. For this section we depend on [36]. We define the swap rate dynamics as follows:

Definition 9 (The SABR model for Swap Rates)**.**
The SABR dyanamic of a swap rate $S^{m,n}$ with expiry T_m and tenor $T_n - T_m$ is under the swap measure $\mathbb{P}^{m,n}$ defined as

$$dS_t^{m,n} = \Sigma_t^{m,n}\left(S_t^{m,n}\right)^{\beta^{m,n}} dW_t^{m,n}, \quad S_0^{m,n} = S^{m,n}(0) \tag{4.1}$$

$$d\Sigma_t^{m,n} = \Sigma_t^{m,n} V^{m,n} dZ_t^{m,n}, \quad \Sigma_0^{m,n} = \Sigma^{m,n}(0) \tag{4.2}$$

$$d\langle W_\bullet^{m,n}, Z_\bullet^{m,n}\rangle_t = R^{m,n} dt, \tag{4.3}$$

where $V^{m,n}$, $\Sigma_0^{m,n} \in \mathbb{R}^+$ and $R^{m,n} \in [-1, 1]$. Further, $W^{m,n}$ and $Z^{m,n}$ are onedimensional Wiener processes.

Remark. Notice, we write for the swap SABR coefficients capital letters, whereas we write for the SABR-LMM coeffcients, except for the forward/volatility correlation matrix R, small letters.

4.2. Swap Rates Dynamics in the SABR-LMM

We want to estimate the swap rate dynamics in a SABR-LMM framework, where we assume that the swap rate evolves under the swap measure $\mathbb{P}^{m,n}$ governed by a simple SABR model as in definition 9 above. There the swap process has the deterministic volatility $V^{m,n}$. Now, the challenging part is the following: If in a LMM the forward rates have deterministic volatility under the forward measures $\mathbb{P}^i$ the swap rates have in general stochastic ones under any $\mathbb{P}^i$. This simply comes from the sum-weights $\omega_i^{m,n}(t)$ in the sum representation (2.5), because they are quotients of stochastic processes. This even happens in the case of a LMM with deterministic volatility only. To circumvent the problem we will simply freeze the weights to their initial values to make them deterministic again.
The approximation will be done step by step. First, we approximate the initial level of the swap volatility $\Sigma_0^{m,n}$ and the vol/vol $V^{m,n}$. Then, the correlation $R^{m,n}$ is approximated. In this section we rely on *Rebonato* [36], but have thoroughly revised the derivations.

For a start, we describe the swap rate $S^{m,n}$ dynamics as

$$dS_t^{m,n} = \Phi_t^{m,n}\left(S^{m,n}\right)^{\beta^{m,n}} dW_t^{m,n}, \tag{4.4}$$

where

$$\Phi_t^{m,n} := \Phi_t^{m,n}(\{\ F_t^0, \ldots, F_t^{N-1}\}, \{\sigma_t^0, \ldots, \sigma_t^{N-1}\}, (\rho_{ij})_{i,j}),$$

is a stochastic volatility depending on the SABR-LMM parameters. Now, our goal is to approximate $\Phi^{m,n}$ and get an idea of its general structure. If we calculate the swap rate dynamics using Ito's formula [9] we get by keeping in mind the SABR-LMM dynamics for forwards (3.7)

$$\begin{aligned} dS_t^{m,n} &= \sum_{l=m}^{n-1} \frac{\partial S_t^{m,n}}{\partial F_t^l} dF_t^l + \frac{1}{2} \sum_{j,l=m}^{n-1} \frac{\partial S_t^{m,n}}{\partial F_t^j \partial F_t^l} d\langle F_\bullet^j, F_\bullet^l \rangle_t \\ &\approx \sum_{l=m}^{n-1} \frac{\sum_{j=m}^{n-1} \omega_j^{m,n}(0) F_t^j}{\partial F_t^l} + \frac{1}{2} \sum_{j,l=m}^{n-1} \frac{\partial S_t^{m,n}}{\partial F_t^j \partial F_t^l} \sigma_t^j \sigma_t^l F_t^j F_t^l \rho_{jl} dt \qquad (4.5) \\ &= \sum_{l=m}^{n-1} \omega_l^{m,n}(0) dF_t^l + \frac{1}{2} \sum_{j,l=m}^{n-1} \frac{\partial S_t^{m,n}}{\partial F_t^j \partial F_t^l} \sigma_t^j \sigma_t^l F_t^j F_t^l \rho_{jl} dt. \qquad (4.6) \end{aligned}$$

Here we used in (4.5) the sum formula (2.5) and freezed the weights $\omega_j^{m,n}$ to their initial values. This is a common technique in financial mathematics and is a quite good approximation for flat underlying yield curves. The $[\ldots]dt$ term can be interpreted as the drift correction from Girsanow [40] due to the change of measures from the forward measures $\mathbb{P}^j$ to the swap measure $\mathbb{P}^{m,n}$. We calculate the quadratic covariation of (4.6) as

$$\begin{aligned} d\langle S_\bullet^{m,n} \rangle_t &= d\langle \sum_{l=m}^{n-1} \omega_l^{m,n}(0) dF_\bullet^l \rangle_t \\ &= d\langle \sum_{l=m}^{n-1} \omega_l^{m,n}(0) \sigma_\bullet^l (F_\bullet^l)^\beta dW_t^l \rangle_t \qquad (4.7) \end{aligned}$$

Now (4.7) is, because of (4.4), equivalent to

$$\begin{aligned} (S_t^{m,n})^{2\beta^{m,n}} (\Phi_t^{m,n})^2 &= \sum_{j,l}^{n-1} \omega_j^{m,n}(0) \omega_l^{m,n}(0) (F_t^j)^\beta (F_t^l)^\beta \sigma_t^j \sigma_t^l \rho_{jl} \\ \Leftrightarrow (\Phi_t^{m,n})^2 &= \sum_{j,l}^{n-1} \frac{\omega_j^{m,n}(0)}{(S_t^{m,n})^{\beta^{m,n}}} \frac{\omega_l^{m,n}(0)}{(S_t^{m,n})^{\beta^{m,n}}} (F_t^j)^\beta (F_t^l)^\beta \sigma_t^j \sigma_t^l \rho_{jl} \qquad (4.8) \end{aligned}$$

We define

$$W_l^{m,n}(t) := \frac{\omega_l^{m,n}(0)\big(F_t^l\big)^{\beta}}{\big(S_t^{m,n}\big)^{\beta^{m,n}}} \tag{4.9}$$

and rewrite (4.8) as

$$\begin{aligned}\big(\Phi_t^{m,n}\big)^2 &= \sum_{j,l}^{n-1} W_j^{m,n}(t) W_l^{m,n}(t) \sigma_t^j \sigma_t^l \rho_{jl} \\ &\approx \sum_{j,l}^{n-1} W_j^{m,n}(0) W_l^{m,n}(0) \sigma_t^j \sigma_t^l \rho_{jl},\end{aligned} \tag{4.10}$$

where we froze the ratios (4.9), which results only in a small lose of precision. This comes from the observation that the ratio

$$\frac{\big(F_t^l\big)^{\beta}}{\big(S_t^{m,n}\big)^{\beta^{m,n}}}$$

is only slowly varying over time due to the high correlation of swaps and forwards, as *Hull* and *White* argued in [21]. Now (4.10) leads to

$$\Phi_t^{m,n} \approx \sqrt{\sum_{j,l}^{n-1} W_j^{m,n}(0) W_l^{m,n}(0) \sigma_t^j \sigma_t^l \rho_{jl}}. \tag{4.11}$$

Therefore, the swap rate dynamics in (4.4) can be approximated as

$$dS_t^{m,n} \approx \sqrt{\sum_{j,l}^{n-1} W_j^{m,n}(0) W_l^{m,n}(0) \sigma_t^j \sigma_t^l \rho_{jl}}\ \big(S_t^{m,n}\big)^{\beta^{m,n}} dW_t^{m,n}.$$

With the help of this representation we plan to approximate the SABR coefficients $\Sigma^{m,n}(0)$ and $V^{m,n}$. From (4.11) we obtain, by writing $\mathbb{E}^{m,n}$ for the expected value under $\mathbb{P}^{m,n}$

$$\mathbb{E}^{m,n}\Big[\int_0^{T_m} \big(\Phi_t^{m,n}\big)^2 dt\Big] \approx \mathbb{E}^{m,n}\Big[\int_0^{T_m} \sum_{j,l}^{n-1} W_j^{m,n}(0) W_l^{m,n}(0) \sigma_t^j \sigma_t^l \rho_{jl} dt\Big]. \tag{4.12}$$

For the right hand side of this equation we obtain

$$\mathbb{E}^{m,n}\big[\int_0^{T_m}\big(\Phi_t^{m,n}\big)^2 dt\big] \approx \mathbb{E}^{m,n}\big[\int_0^{T_m}\big(\Sigma_t^{m,n}\big)^2 dt\big]. \tag{4.13}$$

Now, from the definition of the quadratic variation

$$\langle X_\bullet \rangle_t = X_t^2 + 2\int_0^t X_t dX_t$$

we get, by knowing $\Sigma_t^{m,n}$ has a bounded variation since $V^{m,n} \in \mathbb{R}^+$ and our time horizon is finite what implies that the expected value of $\int_0^t X_t dX_t$ vanishes,

$$\begin{aligned}\frac{d}{dt}\mathbb{E}^{m,n}\big[\big(\Sigma_t^{m,n}\big)^2\big] &= \frac{d}{dt}\big(V^{m,n}\big)^2\int_0^t \mathbb{E}^{m,n}\big[\big(\Sigma_s^{m,n}\big)^2\big]ds \\ &= \big(V^{m,n}\big)^2\mathbb{E}^{m,n}\big[\big(\Sigma_t^{m,n}\big)^2\big].\end{aligned}$$

Hence,

$$\mathbb{E}^{m,n}\big[\big(\Sigma_t^{m,n}\big)^2\big] = \big(\Sigma_0^{m,n}\big)^2\exp\big(V^{m,n}t\big)$$

and therefore

$$\begin{aligned}\int_0^{T_m}\mathbb{E}^{m,n}\big[\big(\Sigma_s^{m,n}\big)^2\big]ds &= \int_0^{T_m}\big(\Sigma_0^{m,n}\big)^2\exp\big(V^{m,n}t\big)dt \\ &= \Big(\frac{\Sigma^{m,n}}{V^{m,n}}\Big)^2\Big(\exp\big((V^{m,n})^2T_m\big)-1\Big),\end{aligned} \tag{4.14}$$

which gives us the right hand side of (4.13). Now, we come back to (4.12) and use our last equation (4.14) together with the definition of the σ_t^l in the SABR-LMM (3.8). This leads to

$$\begin{aligned}\Big(\frac{\Sigma^{m,n}}{V^{m,n}}\Big)^2\Big(\exp\big((V^{m,n})^2T_m\big)-1\Big) \approx &\sum_{j,l=m}^{n-1}\rho_{jl}W_j^{m,n}(0)W_l^{m,n}(0) \\ &\times\int_0^{T_m} g^j(t)g^l(t)\mathbb{E}^{m,n}\big[k_t^jk_t^l\big]dt.\end{aligned} \tag{4.15}$$

Further, the definition of the quadratic covariation gives in the same fashion as above

$$\mathbb{E}^{m,n}\big[k_t^jk_t^l\big] \approx k_0^jk_0^l\exp\big(r_{jl}\widehat{h}_{jl}t\big), \tag{4.16}$$

where we neglected any drift terms for the k^l coming from the change of measures from $\mathbb{P}^l$ to $\mathbb{P}^{m,n}$ and

$$\widehat{h}_t^{jl} := \sqrt{\frac{1}{t}\int_0^t h_s^j h_s^l ds}.$$

Overall we get with (4.16) together with (4.15)

$$\begin{aligned}\Big(\frac{\Sigma^{m,n}}{V^{m,n}}T_m\Big)^2\Big(\exp\big((V^{m,n})^2\big)-1\Big) \approx& \sum_{j,l=m}^{n-1}\rho_{jl}W_j^{m,n}(0)W_l^{m,n}(0)k_0^j k_0^l\\ &\times\int_0^{T_m} g^j(t)g^l(t)\exp\big(r_{jl}\widehat{h}_{jl}t\big)dt.\end{aligned}$$

A Taylor approximation from second order of both sides and equating the terms of the same order gives

$$\Sigma^{m,n} \approx \sqrt{\frac{1}{T_m}\sum_{j,l=m}^{n-1}\rho_{jl}W_j^{m,n}(0)W_l^{m,n}(0)k_0^j k_0^l\int_0^{T_m} g_t^j g_t^l dt} \tag{4.17}$$

and

$$V^{m,n} \approx \frac{1}{\Sigma^{m,n}T_m}\sqrt{2\sum_{j,l=m}^{n-1}\rho_{jl}r_{jl}W_j^{m,n}(0)W_l^{m,n}(0)k_0^j k_0^l\int_0^{T_m} g_t^j g_t^l\big(\widehat{h}_t^{jl}\big)^2 t dt}. \tag{4.18}$$

This gives a good approximation for two of the four SABR parameters in the swap model in definition 9. Further, those equation will be from major importance in the calibration part, when it comes to estimating the two correlation ρ and r of the SABR-LMM model.

To describe the swap correlation R^{mn} in an environment of a SABR-LMM *Rebonato* approximates in [36]

$$R^{m,n} \approx \sum_{j,l}^{n-1}\Omega_{jl}R_{jl}, \tag{4.19}$$

where he defined the matrix $(\Omega_{jl})_{jl}$ as

$$\Omega_{jl} := \frac{2\rho_{jl} r_{jl} W_j^{m,n}(0) W_l^{m,n}(0) k_0^j k_0^l \int_0^{T_M} g_t^j g_t^l \big(\widehat{h}_{jl}\big)^2 t dt}{\big(V^{m,n} \Sigma^{m,n} T_m\big)^2}.$$

Notice that, $\sum_{j,l=m}^{n-1} \Omega_{jl} \approx 1$ due to equation (4.18). Further we define

$$\beta^{m,n} := \sum_{l=m}^{n-1} \omega_l^{m,n}(0)\beta. \tag{4.20}$$

This choice is reasonable since is exact for $\beta = 0$ and $\beta = 1$. In the case $\beta \in (0, 1)$ [4] implies that the error we produce is very small and can be neglected.

In his book [36] *Rebonato* showed that the approximations in the equations (4.18) to (4.19) works with great precision. He tested his approach by evaluating swaptions with different strikes. The accuracy gets better the longer the tenor of the swap is. In the case of swaptions on swaps with expiry 5 years and tenor 15 years or with expiry 10 years and tenor 10 years the approximations are working almost perfect for all strikes. If the expiry becomes shorter the derivations in terms of the volatility smile grow slightly for higher strikes.

5. Parametrization of the SABR-LMM Model

The acceptance of a market model stands and falls with its tractability, the quality of its produced prices and the valuation time needed for pricing. One huge factor for the tractability is the parametrization of the model, since it determines the amount of required parameters and, in terms of calibration, the calibration time and the calibration stability.

A parametrization reduces the describing model parameters drastically by using idealized functions to catch the relevant characteristics. In our case for the SABR-LMM, the underlying functions that parametrize certain model coefficients or structures, like correlation, belonging to a certain forward rate F^i is the same for all forwards and is individualized by using some dependency on the expiry or similar. For example, this techniques allows us to reduce the needed parameters for the correlation matrix ρ from $\frac{(N-1)N}{2}$ to 5, or even 2, numbers. Another example is the use of the same underlying function g_t to describe all the $N-1$ functions g_t^i by shifting the time parameter t depending on the expiry of F^i.

In the following chapters we will first describe how to chose the parameterizations for the functions g^i and h^i, respectively. Second, we give parameterizations for the correlation structure consisting of ρ, r and R.

5.1. The Volatility Structure

If we want to parametrize the volatility functions g^i and h^i in a reasonable way, we have to take into account the basic properties and general shapes volatilities empirically have depending on time. According to [35] the parametrization for the volatility should be time-homogeneous, because empirically the volatilities of forward rates develop all in the same way when expiry gets smaller. That means if we are currently at time point T_i and go one time step to T_{i+1} the k-th forward rate should have roughly the same volatility as the $(k-1)$-th forward rate one step back at T_i. Therefore, we want to parametrize the deterministic volatility and the vol-vol of the SABR-LMM as

$$g_t^i := g(T_i - t) \tag{5.1}$$

and

$$h_t^i := \zeta_i h(T_i - t) \tag{5.2}$$

for some functions g and h and coefficients ζ_i will are close to 1. Further, a typical volatility structure over time either has its maximum in a range of 1.5 years to 4 years to expiry or falls monotonously and concavely with rising times to expiry. In addition, it is observable that the volatilities have a certain terminal level. To describe all this behavior we define as in [35] the underlying functions g and h for g^i and h^i, respectively, as

$$g(t) := (a^g + b^g t)\exp(-c^g t) + d^g \tag{5.3}$$

and analogously

$$h(t) := (a^h + b^h t)\exp(-c^h t) + d^h, \tag{5.4}$$

where $a^g + d^g, a^h + b^h > 0$ and $d^g, d^h > 0$ some real numbers. Note that, the instantaneous volatilities are given as

$$\lim_{t\to 0} g(t) = a^g + d^g \tag{5.5}$$

and

$$\lim_{t\to 0} h(t) = a^h + d^h. \tag{5.6}$$

Further, the terminal volatilities are given as

$$\lim_{t\to \infty} g(t) = d^g \tag{5.7}$$

and

$$\lim_{t\to \infty} h(t) = d^h. \tag{5.8}$$

The extrema of g and h are given at

$$t = \frac{1}{c^g} - \frac{a^g}{b^g} \tag{5.9}$$

and

$$t = \frac{1}{c^h} - \frac{a^h}{b^h}, \tag{5.10}$$

respectively.
Notice, that the g^i and h^i are square integrable, as required in definition 8 for the SABR-LMM. Furthermore, closed form solutions exists for those integrals, which will valuable when it comes to calibration since we can solve the integrals analytically rather then by cumbersome numerical integration.
We will use the knowledge about instantaneous and terminal volatilities to choose initial value for the calibration later on. In addition, we will incorporate the stylized fact about extrema occurring in a range of 1.5 years to 4 years to expiry.

5.2. The Correlation Structure

The heart of the SABR-LMM is the super correlation matrix P. The correlations are describing the direct dependence of the forward rates on each other and the volatilities. A part of P describes the *cross skew* of the model, that is the correlation between forward rates and volatilities. Altogether, the super correlation matrix is the main difference and biggest advantage over the standard LMM model [2], [11]. The matrix does not carry the level of the model volatility, but most of the other informations over the shape of the volatility surface, like its skew and how strongly it is pronounced.
As in definition 8 of the SABR-LMM we write

$$P = \begin{pmatrix} \rho & R \\ R^T & r \end{pmatrix},$$

and notice that the Matrix consists of $4(N-1)^2$ parameters, from which we *only* have to estimate $N(N-1)$ due to symmetry. Since this number is way to big we want give some stylized parametrization depending on maximal 9 parameters and minimal 6, for the whole super correlation matrix P. We we do this in the same fashion as for the volatilities in chapter 5.1.

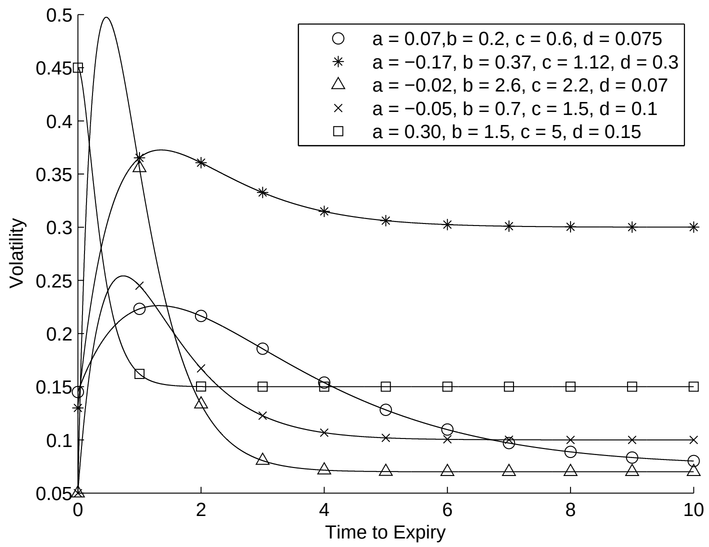

Figure 5.1.: Here a range of possible shapes for the function g and h are shown. The parametrization can be classified in two groups. In the first group are the ones that produce a real humped shape and in the second group are those parametrization where the volatility falls strictly. According to *Rebonato* in [36] and [35] the humped shaped volatility functions are characteristic for normal market situations and the falling volatility functions occur in excited market. The terminal volatility is clearly visible and agrees with the parameter d. The parameters are from [36] and the plot was done in Matlab.

We will give for each sub matrix of P a parametrization and glue them together in the end. For the gluing we will need an optimization algorithm that gives us the nearest correlation matrix, since just sticking together three parameterizations for ρ, r and R do not give a well-defined correlation matrix because the eigenvalues of the resulting matrix can be negative.
The most important part of P is the sub matrix ρ, which directly governs the interplay between the forward rates. In this way ρ has the most impact on the pricing quality of our model. Therefore, we put special emphasis in modeling the forward/forward correlation.

In the later, we discuss correlation matrices in general and refer to a correlation matrix using the symbol ρ. Obviously, all the discussions will hold for the volatility/volatility correlation r as well.

To give proper parameterizations we first have to give some criteria which properties a correlation matrix *must* have and which it *should* have. According to Lutz [29] and [36] for the correlation matrix $(\rho_{ij})_{ij}$ has to hold

$$\begin{aligned}
&(A1) \quad \rho \text{ has to be real and symmetric,}\\
&(A2) \quad \rho_{i,i} = 1 \text{ for all } i \in \{1,\dots,N\},\\
&(A3) \quad \rho \text{ has to be positve semi-definite.}
\end{aligned}$$

In addition, we demand two further properties, which describe empirical observations and whose validity is market consents.

$$\begin{aligned}
&(B1) \quad j \mapsto \rho_{ij} \text{ should fall strong monotonously for } j > i,\\
&(B2) \quad i \mapsto \rho_{i+p,i} \text{ grows for fixed } p \in \{1,\dots,N-2\}.
\end{aligned}$$

The first property assures that two forward rates whose expiry is farther apart a less correlated then two rates who expire closely together. The second property assures that, if we have two pairs of assets and in each pair the distance between the expiries is the same, then the pair of assets which overall expiry is further in the future is stronger correlated then the other one. For example, lets consider two pairs of forwards. The first pair consists of forward rates expiring in 1 and 3 years and the second pair consists of forwards expiring in 20 and 22 years. Intuitively, it is clear that the last two forwards should be more strongly correlated than the first two.

The simplest matrix that fulfills the properties $(A1)$–$(A3)$ is given through

$$\rho_{ij} = \exp(-\beta|i-j|), \tag{5.11}$$

where $\beta > 0$ is the decorrelation coefficient. It is obvious that this matrix does not obey $(B1)$ and $(B2)$, since only the index distance of the tenor points $(T_i)_i$ matters. The simple structure and the dependence on only one parameter β is nevertheless attractive. This is especially useful in situations where we have to set up a matrix under high uncertainty or we believe the correlations behave uniformly. We want to further develop the approach in (5.11) in a trivial way. For this, we first notice that, if $(\rho_{ij})_{i,j}$ is a correlation matrix then $(\widetilde{\rho}_{ij})_{i,j}$ defined through

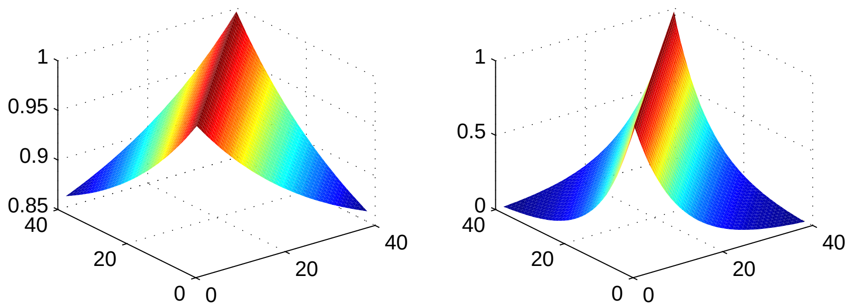

Figure 5.2.: Some examples for the correlation matrix in (5.13). The simple structure is obvious. On the left the parameters are $\beta^{(1)} = 0.03$ and $\rho_\infty^{(1)} = 0.8$ on the right the parameters are $\beta^{(2)} = 0.1$ and $\rho_\infty^{(2)} = 0.0$. The plot was done in Matlab.

$$\widetilde{\rho}_{ij} := \rho_\infty + (1-\rho_\infty)\rho_{ij} \tag{5.12}$$

is a correlation matrix as well, where $\rho_\infty \in [0,1)$. The coefficient ρ_∞ describes the terminal correlation and it holds

$$\widetilde{\rho}_{ij} \xrightarrow[j\to\infty]{} \rho_\infty.$$

Now we enhance (5.11) to

$$\rho_{ij} = \rho_\infty + (1-\rho_\infty)\exp(-\beta|i-j|), \tag{5.13}$$

where $\rho_\infty \in [0,1)$. Indeed, we will use the above parametrization to model the correlation between the volatility processes, hence we will use it for the submatrix r.
However the parametrization in (5.13) is way to simple and inflexible to model the correlation of forward rates. The reason for this is the uniform correlation coefficient β. Empirically, forward rates with shorter expiry are way less correlated with other forward rates then forwards with larger expiry. This means we need a way to model the correlations on the *short* end of the tenor structure more independently from the *long* end. The first step in this direction is the Doust parametrization [31]. This parametrization gives a general framework for matrices that fulfill $(A1)$–$(A3)$ and therefore for correlation matrices.
A matrix $(\rho_{ij})_{ij}$ obeys the Doust parametrization if there exists a set

$$\{a_k \mid a_k \in [-1,1],\ k \in \{1,\dots,N-2\}\}$$

such that

$$\begin{aligned} \rho_{i,i} &= 1,\ \text{for all } i \in \{1,\dots,N-2\} \\ \rho_{1,j} &= \prod_{k=1}^{j-1} a_k = \rho_{j,1} \\ \rho_{i,j} &= \frac{\rho_{1,j}}{\rho_{i,1}} = \prod_{k=i}^{j-1} a_k \end{aligned} \tag{5.14}$$

and therefore ρ can be written as

$$\rho = \begin{pmatrix} 1 & a_1 & a_1a_2 & \cdots & a_1 \dots a_{N-2} \\ a_1 & 1 & a_2 & \cdots & a_2 \dots a_{N-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_1 \dots a_{N-1} & \cdots & \cdots & \cdots & 1 \end{pmatrix}.$$

This implies ρ has the Cholesky decomposition

$$\rho = LL^T,$$

where

$$L := \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ a_1 & \sqrt{1-a_1^2} & 0 & \ddots & \cdots & \vdots \\ a_1 a_2 & a_2\sqrt{1-a_1^2} & \sqrt{1-a_2^2} & 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ a_1 \dots a_{N-1} & a_2 a_3 \dots a_{N-2}\sqrt{1-a_1^2} & \cdots & \cdots & \cdots & \sqrt{1-a_{N-2}^2} \end{pmatrix}.$$

which yields for ρ the correlation matrix property. However, this representation cannot grantee that either $(B1)$ or $(B2)$ holds. Another problem is the dependence on $N-2$ coefficients, which is simply too much since in practice N lies in the range of 20 to 40. *Schoenmakers* and *Coffey* further developed in [25] the approach from *Doust*. They gave a parametrization, which gives a correlation matrix that fulfills $(B1)$ and $(B2)$ as well. In its most common formulation it depends on N parameters, but the number can be efficiently reduced to two parameters. A correlation matrix $(\rho_{ij})_{ij}$ follows the Schoenmakers & Coffey parametrization if and only if there exists a growing sequence

$$1 = b_1 < b_2 < \cdots < b_N \tag{5.15}$$

such that

$$\frac{b_1}{b_2} < \frac{b_2}{b_3} < \cdots < \frac{b_{N-2}}{b_{N-1}} \tag{5.16}$$

and

$$\rho_{ij} = \frac{b_j}{b_i}, \text{ for all } 1 \leq j \leq i \leq N-1 \tag{5.17}$$

with

$$\rho_{ij} = \rho_{ji}.$$

Here the definition of the entries in ρ via fractions (5.17) corresponds to the definition of the Doust parametrization (5.14). The two additional requirements (5.15) and (5.16) yield the desired properties $(B1)$ and $(B2)$. Further, it was shown in [25] that a matrix ρ obeys the Schoenmakers & Coffey parametrization, if there exists a sequence

$(b_i)_{i\in\{1,\ldots,N-1\}}$ such that (5.17) holds and

$$b_i = \exp\big(\sum_{j=1}^{N-1}\min(j,i)\Delta_j\big), \tag{5.18}$$

for a real sequence $\Delta_1,\ldots,\Delta_N$.
With the help of equation (5.18) it is possible to generate correlation matrices with properties (5.16) and (5.17) without non linear bounded parameters b_i. This is an important fact for the implementation since it fastens up the computation time. Let's choose in (5.18) the following parameters

$$\Delta_i := \alpha\Big(\frac{N-i-2}{N-4}\Big), \text{ for } 1\le i< N-1 \text{ and } \Delta_{N-1} := \frac{\gamma}{N-2} - \frac{\alpha}{6}(N-3) \tag{5.19}$$

where

$$\alpha := \frac{6\eta}{(N-2)(N-3)},$$

This gives the *optimal* two parametric correlation matrix $(\rho_{ij})_{ij}$ from *Schoenmakers* & *Coeffey* [25] with

$$\rho_{ij} = \exp\Big[-\frac{|j-i|}{N-2}\Big(\gamma+\eta h(i,j)\Big)\Big] \text{ for all } i,j\in\{1,\ldots,N-1\}, \tag{5.20}$$

where

$$h(i,j) := \Big(\frac{i^2+j^2+ij-3(N-1)i-3(N-1)j+3i+3j+2(N-1)^2-N-5}{(N-3)(N-4)}\Big)$$

and

$$\eta\ge 0,\ \gamma\ge 0,\ \gamma-\eta\ge 0. \tag{5.21}$$

Here the parameter $\exp(-\gamma)$ is the terminal correlation just like in (5.12). From here on we will refer to the above parametrization as the (2SC) parametrization. *Schoenmakers* & *Coffey* claim in [25] that the representation in (5.20) is flexible enough to describe a wide range of different correlation matrices.
However, we found in the empirical work in chapter 8 that the parametrization works quit well, but may be too inflexible. This is due to the high dependency between the shape of the short end of the matrix – the area for assets with shorter maturity –

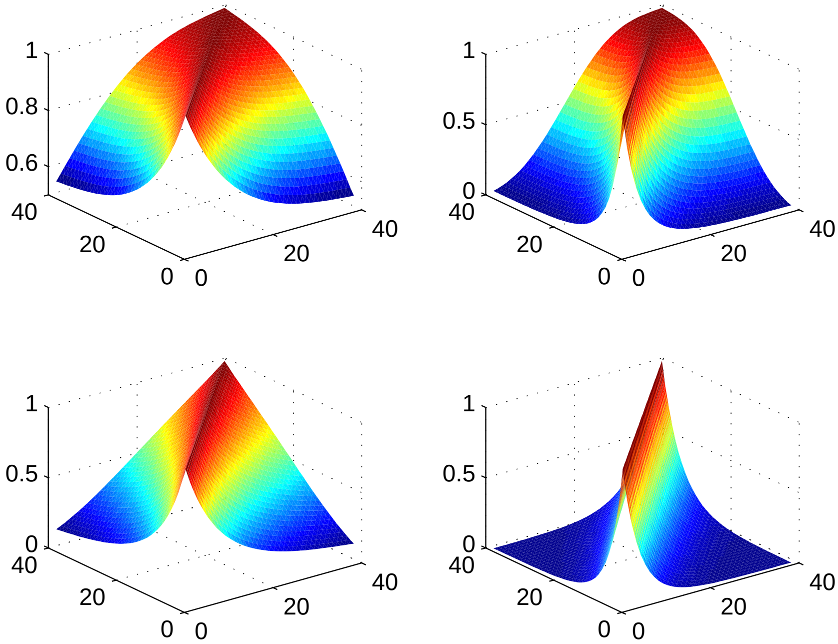

Figure 5.3.: Here different (2SC) parametrizations are shown. From the left to the right the coefficients are given in table 5.1. The limitation of the parametrization is clearly visible, since the back of the matrix heavily depends on the front. The plot was done in Matlab.

and the long end of the matrix – the area for assets with longer maturity. The strong dependency of those two areas comes from property (5.16), which says that the last row of the matrix $\rho_{i,N} = \frac{b_i}{b_N}$ is almost inverse proportional to the first row of the matrix $\rho_{i,1} = \frac{1}{b_i}$. We believe that the two parameters of the (2SC) representation are simply not enough to give the flexibility which we desire. Especially not if, we try to describe the correlation of assets whose prices highly depend on the correlations between the underlying assets, like Constant Maturity Swaps Spreads (CMS spreads). Nevertheless, the (2SC) parametrization is very good since it works highly efficient with only two parameters, which allows for a very fast calibration and, particularly, a very stable one. Further, it reduces the risk of over-fitting, because it does not react to every minor disturbance in the data and rather keeps a general textbook shape.

Table 5.1.: Parameters for Figure 5.3

Plot No.	1	2	3	4
γ	0.6	3.5	2.0	8.0
η	0.5	3.4	1.0	1.0

To get a more flexible parametrization we chose to incorporate more parameters. A good parametrization which can reproduce a wide range of stylized shapes is given by *Lutz* in [29]. His correlation matrix depends on 5 parameters and the representation is derived by its Cholesky decomposition. For this, he first gave a general framework for the decomposition:

Theorem 3 (Lutz' Cholesky Decomposition for Correlation Matrices)**.**
Let ρ be a $N \times N$ correlation matrix with full rank. Further define $I_2 := \{2, \dots, N\}$ and $I_1 := \{1, \dots, N\}$. Then there exists functions $f : I_1 \longrightarrow [-1, 1]$ and $g : I_2 \times I_2 \longrightarrow \mathbb{R}$, such that

$$\rho_{ij} = f(i)f(j) + \frac{(1 - f(i)^2)(1 - f(j)^2)}{a_i a_j} \sum_{k=2}^{min(i,j)} h(i,j)h(j,k) \text{ for all } i, j \in I_1$$

and the Cholesky decomposition L of ρ can be written as

$$L_{ij} = \begin{cases} f(i) & \text{, for } j = 1 \\ h(i,j)\sqrt{\frac{1-f(i)^2}{a_i}} & \text{, for } 1 < j \leq i \text{ ,} \\ 0 & \text{, otherwise} \end{cases}$$

with

$$a_1 = 1, \ a_i = \sum_{k=2}^{i} h(i,k)^2, i \in I.$$

Proof. The proof is simple, but doesn't give any new insights. We refer to [29]. □

Remark. The correlation matrix of Schoenmakers & Coffey can be retained by setting

$$f(i) = \frac{1}{b_i} \text{ and } h(i,j) = \sqrt{b_j^2 - b_{j-1}^2}, \ i, j \in I_2,$$

where we choose b_i according to (5.18) and with Δ_i given in (5.19).

If we choose in Theorem 3

$$f(i) = \exp(-\beta i^{\alpha}), \text{ for } \alpha, \beta > 0, i \in I_1$$

and

$$h(i,j) = \exp\Big[-\Big(\frac{i-1}{N-2}\gamma + \frac{N-1-i}{N-2}\delta\Big)\Big(\frac{j}{i}-1\Big)\Big]$$

and combine the resulting matrix with the terminal correlation extension as in (5.12) we get Lutz' 5 parametric form (5L). The parametrization is the following

$$\rho_{ij} = \rho_\infty + (1-\rho_\infty)\Big[\exp(-\beta(i^\alpha + j^\alpha)) + H(i,j,\alpha,\beta)\Big], \ \rho_\infty \in [0,1) \tag{5.22}$$

with

$$H(i,j,\alpha,\beta) := \frac{\theta_{ij}}{\sqrt{\theta_{ii}\theta_{jj}}}\sqrt{(1-\exp[-2\beta i^\alpha])(1-\exp[-2\beta j^\alpha])}$$

for $\alpha, \beta > 0$ and $i, j \in I_1$ and with

$$\theta_{ij} := \begin{cases} 1 & , \text{ if } \min(i,j) = 1 \\ \min(i,j) - 1 & , \text{ if } \min(i,j) > 1, \xi_i\xi_j = 1 \ , \\ \frac{(\xi_i\xi_j)^{\min(i,j)-1} - 1}{1 - 1/(\xi_i\xi_j)} & , \text{ if } \min(i,j) > 1, \xi_i\xi_j \neq 1 \end{cases}$$

where

$$\xi_i := \exp\Big(-\frac{1}{i}\Big(\frac{i-1}{N-2}\gamma + \frac{N-i-1}{N-2}\delta\Big)\Big), \ \gamma, \delta \in \mathbb{R}.$$

Notice, the parametrization for ρ given above in (5.22) need not necessarily fulfill the desired requirements $(B1)$ and $(B2)$ since Lutz' Theorem only affects the general correlation matrix properties $(A1)$–$(A3)$. This could be a possible problem, when it comes to calibration, because we lose some control over the general shape of the matrix. In [29] *Lutz* argues that, given good enough input data for the calibration procedure the resulting shape should fulfill the properties $(B1)$ and $(B2)$. His opinion comes from empirical studies of correlation matrices for forward rates from the years

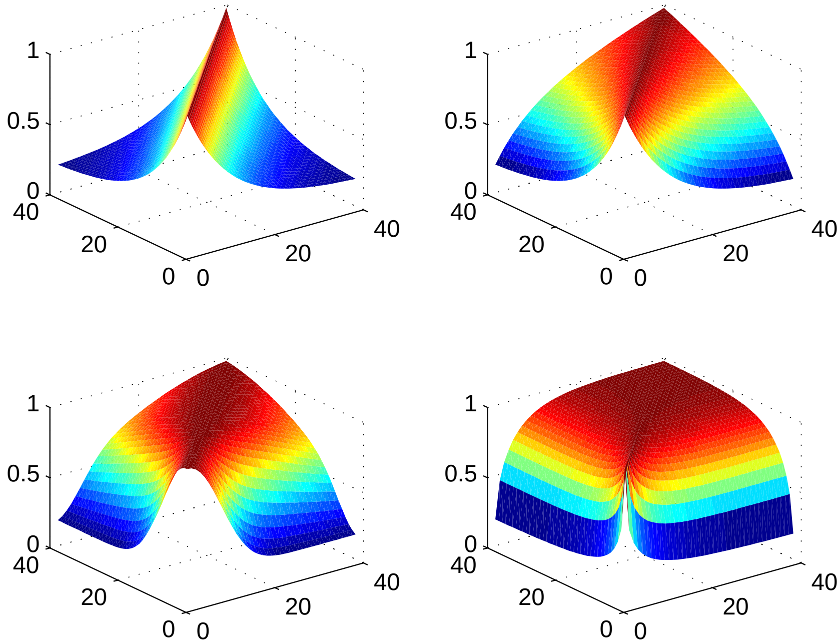

Figure 5.4.: Here are possible shapes of the (5L) parametrization shown. We kept ρ_∞ fixed at 0.2. Observe the high independency of the back end of the matrix from the front end and flexibility comparing with the (2SC) parametrization in figure 5.3. The parameters for this plot are from [29] and can be found in table 5.2 . The plot was done in Matlab.

2004 to 2008. There, without given any parametrization for the matrix, the average resulting shape fulfilled ($B1$) and ($B2$). In addition, in our empirical work we gain the same opinion, because all our estimated correlation matrices for different dates and parametrized through (5L) obeyed all desired properties. Further in this work we will use the parameterizations (2SC) (5.20) and (5L) (5.22) to model the correlation ρ of the forward rates in the SABR-LMM. We will compare both approaches visually and by the accuracy of produced prices.

Further, we model the correlation of the volatilities with the simple parametrization given in (5.13).

The last structure we have to define is the one of the cross skew R, which describes the correlation between forward rates and volatilities. Here, we follow *Rebonato* [36]

Table 5.2.: Parameters for Figure 5.4

Plot No.	1	2	3	4
α	1.0	1.0	2.2	0.5
β	0.1	0.1	0.006	0.8
γ	-3.7	0.0	0.95	3.0
η	-0.3	0.0	3.6	-0.2

and choose

$$R_{ij} = \operatorname{sgn}(R_{ii})\sqrt{|R_{ii}R_{jj}|}\exp\big(-\lambda_1(T_i - T_j)^+ + \lambda_2(T_j - T_i)^+\big),\ \ \lambda_1, \lambda_2 > 0. \tag{5.23}$$

We will see in the calibration part that, the R_{ii} are the individual forward/volatility correlations for each forward rate F^i. Overall, we have given a piecewise parametrization of whole super correlation matrix

$$P = \begin{pmatrix} \rho & R \\ R^T & r \end{pmatrix}.$$

5.3. Approximation of P through a proper Correlation Matrix

In chapter 5.2 we have shown how to parametrize the pieces of the super correlation matrix P. But, by simply sticking those pieces together it is not guaranteed that we really obtain a correlation matrix by definition, since the eigenvalues can be negative. This is extremely problematic, because we need to be able to do a Cholesky decomposition of P for a Monte Carlo simulation. To fix this problem, we approximate the matrix P through the nearest correlation matrix $\widehat{P}$. Hence, we solve the problem

$$\widehat{P} = \arg\min_{A \in \mathcal{C}} \sqrt{\sum_{ij} \omega_{ij}(A_{ij} - P_{ij})^2}, \tag{5.24}$$

where $\mathcal{C}$ is the set of all correlation matrices and the $\omega_{ij} \geq 0$ are weights. In most cases we won't retain P as a solution, but the weights enable us to put special emphasis on some areas of P to maintain the most crucial characteristics.

Obviously, in the SABR-LMM the matrix for the forward/forward correlation ρ is of

major importance because it describes the direct interplay of the rates. Another important part of P is the diagonal of R, since it consists of the individual forward/volatility correlations, which have an huge impact on the individual model smile of a the rate F^i. *Rebonato* stressed in [36] the importance of the R_{ii} and our empirical studies confirmed him. The impact of the changes in the skew were visualized in figure 3.1. So we decided to use the ω_{ij} to overweight ρ by 8 and $(R_{ii})_i$ by 80.
The minimization algorithm is behind the scope of this work. More on this topic can be found in the *Cresnik*'s thesis [28], where he used the majorization technique, described in [34]. Further, the algorithm in [34] can solve the problem (5.24) extended to

$$\widehat{P} = \underset{A \in \mathcal{C}_k}{\arg\min} \sqrt{\sum_{ij} \omega_{ij} (A_{ij} - P_{ij})^2}, \tag{5.25}$$

where now $\mathcal{C}_k$ is the set of all correlation matrices of rang $k \leq 2(N-1)$.
This enables us to reduce the simulation time immensely since there is a one to one relation between rank and number processes we have to simulate for pricing. In this work we will use $k = 10$.

6. Calibration of the SABR-LMM to Swaptions

Now that the theoretical foundation for the SABR-LMM is given, we want to show how the model can be calibrated to the market. We will give a method to calibrate implicitly to current market prices and which is not based on historical prices for some past period. Historic prices have a certain backwards character and do not carry more information than current market prices. These already incorporate all relevant expectation of future price developments. Therefore, historic prices cannot help to estimate the future any better.

The basis for the calibration will be the prices in Black volatilities of caps and swaptions. The goal is to calibrate the induced model prices as close as possible to the observed market prices. To calculate the induced model prices fast and with satisfactory accuracy, we will use the formula for the implied Black volatility for the simple SABR model (3.4) together with the results in chapter 4. There we demonstrated how to transfer the swap dynamics from the SABR-LMM to the SABR model, where we can use the implied volatility formula.

This chapter is split into three parts. In the first part we explain what kind of data we need and how it can be stripped. The second describes the calibration of the volatility structure, that is the estimation of the parametrization for the g^i and h^i, which is done solely based on caplets. This part of the calibration will be exactly the same when we calibrate the model to CMS spread options later in chapter 7.3. The third part of this chapter concerns the calibration of the correlation structure to swaption prices.

In the following, by using four parameters in our notation we will only slightly rely on the parameterizations of the g^i and h^i. The parametrization of the super correlation matrix does not matter either, because we will give all formulas only depending on the entries of P. So the described procedures can easily used with other parametrizations then the ones introduced in chapter 5.

6.1. Preparation for the Calibration

For the calibration we need certain basic data. Here, I want explain in short how to obtain it. First, we need the initial values F_0^i for the forward processes. Those can be obtained by a forward bootstrapping as explained in chapter 2.2.1. Further, we have to get the bond prices. Those are necessary to calculate the swap numéraires $A_0^{m,n}$ when it comes to calibration on swaps. We calculate the bond prices by using the forward rates as

$$B(0,T_i) = \prod_{j=0}^{i-1}(1+\delta_j F_0^j)^{-1},$$

where we obtained the formula by simple iteration of the definition for forward rates (2.1).

6.2. The Calibration of the Volatility Structure

One of the big objectives of the calibration of the SABR-LMM is to keep the dynamics of the forward rates as close to the simple SABR model as possible. If this works the SABR-LMM's acceptance will be highly strengthened since the SABR model has already been approved in practice. In particular, the following calibration procedure will assure that we can use the formula for implied volatility (3.4) of the SABR model to approximate prices of European options on forward rates and swaps even in the SABR-LMM as well.
The first step in the calibration of the volatility functions is to bootstrap the implied caplet volatilities from quoted cap volatilities as described in the preliminary chapter 2.2.2. Then we calibrate for each forward rate F^i a simple SABR model by using the formula from (3.5)

$$(\sigma_i, \rho_i, \nu_i) = \underset{\alpha,\rho,\nu}{\arg\min} \sqrt{\sum_j \left[\sigma_{\mathrm{M}}(F_0^i, K_j^i) - \sigma_{\mathrm{I}}(F_0^i, K_j^i, \beta_j, \alpha, \nu, \rho, T_i)\right]^2}, \tag{6.1}$$

where the $\sigma_{\mathrm{M}}(F_0^i, K_j^i)$ are the stripped caplet volatilities, the $(K_j^i)_j$ the available strikes for the i-th caplet and the β_j are the betas of forward rates in the simple SABR models. In our case holds $\beta_j \equiv 0.5$, but other choices are possible, as explained in [36]. We would like to emphasize that, the set

$$(\sigma_i, \rho_i, \nu_i)_i \tag{6.2}$$

will play a central role in all calibration procedures from here on.
The functions g^i can be interpreted as the deterministic volatility of F^i. Since we chose h^i square integrable the stochastic volatility k^i is on average k_0 as can be seen in (3.8). Therefore, we follow [36] and calibrate the g^i like in the simple LMM with deterministic volatility [3]. We define

$$\widehat{g}_i := \sqrt{\frac{1}{T_i}\int_0^{T_i} \left(g_t^i\right)^2 dt}$$

as the average squared volatility and then use the minimization problem from (3.6)

$$\min_{(a^g,b^g,c^g,d^g)\in\mathcal{C}} \sqrt{\sum_{i=1}^{N} \omega_i \left(\sigma_i - \widehat{g}_i\right)^2}, \tag{6.3}$$

where $\mathcal{C}$ is defined as

$$\mathcal{C} := \left\{(a,b,c,d) \in \mathbb{R}^4 \mid a+b > \epsilon_1, d > \epsilon_2\right\}, \tag{6.4}$$

with ϵ_1 and ϵ_2 representing some believes about instantaneous and terminal volatility, as explained in (5.5) and (5.7). The weights ω_i are defined as

$$\omega_i := (0.2 + 0.5T_i)\exp(-0.2T_i) + 0.2$$

and overweight the σ_i for the range 1.5 years to 5 years. If the parametrization is able to fit today's volatility structure perfectly the weights shouldn't have any impact.
With (6.3) we are able to calibrate the deterministic volatility averaged over all forward rates. The derivation for the validity of the formula is exactly as in the LMM with deterministic volatility, if we approximate $k_t^i \approx k_0^i \approx 1$. The last approximation is valid, because to achieve that the SABR volatility σ_i agrees with the SABR-LMM volatility σ_t^i on average for each forward rate, we calculate the initial value k_0^i of the stochastic volatility as

$$k_0^i = \frac{\sigma_i}{\widehat{g}_i}.$$

So the k_0^i should be near to one if the fits of the g^i are adequate. To estimate the parameters for h we follow [36] and define

$$\widehat{h}_t^i := \sqrt{\frac{1}{t}\int_0^t \left[(a^h + b^h(T_i - s))\exp(-c^h(T_i - s)) + d^h\right]^2 ds}$$

and set up the optimization problem

$$\min_{(a^h,b^h,c^h,d^h)\in\mathcal{C}} \sqrt{\sum_{i=1}^{N} \nu_i - \frac{k_0^i}{\sigma_i T_i}\left[2\int_0^{T_i} \left(g_t^i\right)^2 \left(\widehat{h}_t^i\right)^2 t dt\right]^{1/2}}. \tag{6.5}$$

In addition, we go a step further then in [36] and calculate the correction factors ζ_i, in order to hit the vol-vol for each forward rate exactly, as

$$\zeta_i = \nu_i / \Big(\frac{k_0^i}{\sigma_i T_i}\left[2\int_0^{T_i} \left(g_t^i\right)^2 \left(\widehat{h}_t^i\right)^2 t dt\right]^{1/2}\Big).$$

Again, the ζ_i should be near to one if the fit is good. Notice, in (6.5) the calibration of the vol-vol functions h^i depends on the g^i as well. So the vol-vol implied by the h^i is higher weighted if the general volatility level, given by the g^i, is higher and vice versa.

6.3. The Calibration of the Correlation to Swaps

Similar to the case of the calibration of the volatility structure in chapter 6.2, we use for the calibration some coefficient from the simple SABR model as target values. The correlation is calibrated on a set of coterminal swaps $\{S^{i,N}\}_i$. Those swaps have different expiry dates T_i and tenors $T_N - T_i$, but all mature in T_N. So a set of coterminal swaps can be seen as a set of swaps in which all swaps end at the same date.

We choose coterminal swaps, because, depending on the picked expiry dates, we can take almost every entry of correlation matrices ρ, r and R into account during the calibration. To realize this we remember the sum structure of swaps from (2.5)

$$S_t^{i,N} = \sum_{i=1}^{N-1} \omega^{m_i,N} F^i.$$

Hence, in a swap rate the forward rates with index in $\{i, \ldots, N-1\}$ and, especially, all the correlations

$$d\langle W^l_\bullet, W^j_\bullet \rangle_t = \rho_{lj},\ d\langle Z^l_\bullet, Z^j_\bullet \rangle_t = r_{lj} \text{ and } d\langle W^l_\bullet, Z^j_\bullet \rangle_t = R_{lj}.$$

for $l, j \in \{i, \ldots, N-1\}$ are involved. If we consider all possible expiries for coterminal swap rates we imply for all entries of the super correlation matrix P a condition. However in practice it is enough to choose only the quoted expiry dates.

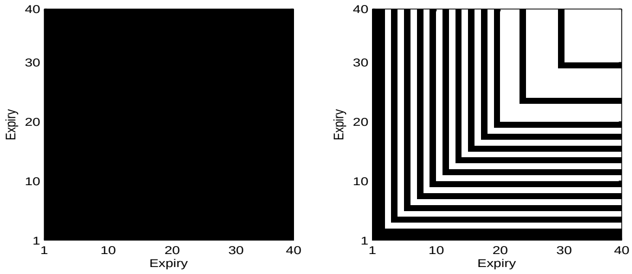

Figure 6.1.: This graphic shows the entries of ρ which are involved in coterminal swap dynamics. Black means involved and white means not involved. On the right we see the case if we use every possible coterminal swap and on the left we see the case in which we only used the coterminal swaps whose expires are quoted in the market. Here the tenor structure is $(T_0, \ldots, T_{40})$ with $\delta_i \equiv 0.5$. The plot was done in Matlab.

The first step in the calibration is to calibrate the SABR model for swaps from definition (9) in chapter 4.1 to swaption prices of coterminal swaps. Those prices quoted in implied volatility. Here we use an extended version of the minimization problem in (3.5)

$$\begin{aligned}(\sigma^S_i, \rho^S_i, \nu^S_i) = \underset{\alpha,\rho,\nu}{\arg\min} \Bigg(& \sqrt{\sum_j \mathbb{1}_{(\epsilon_3,\infty)} \left[\sigma_{\mathrm{M}}(S^{i,N}_0, K^i_j) - \sigma_{\mathrm{I}}(S^{i,N}_0, K^i_j, \beta^S_i, \alpha, \nu, \rho, T_i)\right]^2} \\ & + \varphi(\sigma, \rho, \nu) \Bigg),\end{aligned} \tag{6.6}$$

where $\epsilon_3 > 0$ is the required accuracy of the general optimization problem and φ a penalty function for the possible solutions. The $\sigma_M(S^{i,N}_0, K^i_j)$ are the observed swaption volatilities in the market, the strikes $(K^i_j)_j$ are the available strikes for the i-th coterminal swap and the swap betas β^S_i are calculated as in (4.20). The penalty

function controls the freedom given through $\epsilon_3 > 0$. In practice ϵ_3 is in dimension of 10^{-6} to 10^{-5}, so we disturb the optimization problem only slightly.
Basing on the estimated parameters in (6.6) we will setup an optimization problems for each submatrix of P. In chapter 4.2 we approximated the dynamics of swaps in the SABR-LMM assuming that the swaps follow a SABR model. Now, we will use the results therein. Approximation (4.17) leads to

$$\widehat{\rho} = \arg\min_{\rho} \Bigg(\sum_i \Big(\sigma_i^S - \Big[\frac{1}{T_{m_i}} \sum_{k,l=m_i}^{N-1} \rho_{kl} W_k^{m_i,N}(0) W_l^{m_i,N}(0) \\ \times k_0^k k_0^l \int_0^{T_{m_i}} g_t^k g_t^l dt \Big]^{1/2} \Big)^2 \Bigg)^{1/2} \tag{6.7}$$

as a way to estimate ρ. Next we use (4.18) to set up the following minimization problem, where we assume that ρ was already estimated:

$$\widehat{r} = \arg\min_{r} \Bigg(\sum_i \Big[\nu_i^S - \frac{1}{\sigma_i^S T_{m_i}} \Big[2 \sum_{k,l=m_i}^{N-1} \rho_{kl} r_{kl} W_k^{m_i,N}(0) W_l^{m_i,N}(0) k_0^k k_0^l \\ \times \int_0^{T_{m_i}} g_t^k g_t^l \big(\widehat{h}_t^{kl}\big)^2 t dt \Big]^{1/2} \Big]^2 \Bigg)^{1/2}, \tag{6.8}$$

where

$$\widehat{h}_t^{kl} := \sqrt{\frac{1}{t} \int_0^t h_s^k h_s^l ds}.$$

Our last optimization problem to estimate the cross skew matrix R is obtained by using (4.19) and the idea of we wanting to regain the SABR dynamics as close as possible, is the following

$$\widehat{R} = \arg\min_{R} \sqrt{\sum_i \Big(\rho_i^S - \sum_{k,l=m_i}^{N-1} \Omega_{kl}^i R_{kl} \Big)^2}, \tag{6.9}$$

where

$$\Omega_{kl}^i := \frac{2 \rho_{kl} r_{kl} W_j^{m_i,N}(0) W_l^{m_i,N}(0) k_0^j k_0^l \int_0^{T_{m_i}} g_t^j g_t^l \big(\widehat{h}_{jl}\big)^2 t dt}{\big(\nu_i^S \sigma_i^S T_{m_i}\big)^2}.$$

and we force

$$R_{ii} = \rho_i, \tag{6.10}$$

where ρ_i is the skew from (6.2) of the forward rate F^i. The step (6.10) forces the correlation of each forward rate and its volatility exactly to the same level as in the simple SABR model. By doing this, we once more try to keep the dynamics of the F^i in the SABR-LMM as closely as possible to the simple SABR model.

7. CMS Spread Options and Swaps

A constant maturity swap (CMS) is a swap consisting of floating legs given by swap rates of constant length and some fixed leg. In those contracts the long party is obligated to pay a fixed amount F at each settlement date T_i and to receive a swap rate S fixed at time T_j over some predefined period. From the long party's point of view, and thus similar to a simple swap, the CMS over the period $[T_i, T_j]$ has a value of

$$\sum_{l=i}^{j} \tau_l B(0, T_l) \mathbb{E}^l \left[S_{T_l}^{l,l+c} - F \right],$$

where c is the constant length of all swaps. A CMS can be seen as a normal swap where swap rates with constant length are exchanged instead of forwards. Constant maturity swaps, and options on those, enable market participants to hedge for certain periods against risks due to changes in the general shape of the underlying forward-curve for the swaps. For example, CMS can be used to hedge against changes of the 10 year swap rates by buying a call, put or straddle on a CMS.
An extension of CMS is a CMS spread, where one swap rate is exchanged for another one. The exchanged swap rates have the same expiry and underlying forward curve but different lengths. If the CMS spread consists of two swap rates of length c_1 and c_2, both expiring in T_i, the contract has the following payoff profile for the long-party

$$S_{T_i}^{i,c_1,c_2} := S_{T_i}^{i,i+c1} - S_{T_i}^{i,i+c2} - F, \tag{7.1}$$

where F is some fixed rate and has to be paid by the long-party. Therefore, today's abstract, undiscounted value under the CMS measure $\mathbb{P}^{\mathrm{CMSs}}$, under which the spread S^{i,c_1,c_2} is a martingale, is

$$\mathbb{E}^{CMSs} \left[S_{T_i}^{i,c_1,c_2} \right] - F = \mathbb{E}^{CMSs} \left[S_{T_i}^{i,i+c1} - S_{T_i}^{i,i+c2} - F \right].$$

CMS spread prices depend more strongly on the correlation between the underlying forward rates than ordinary swaps [29]. If we are able to evaluate spreads in the SABR-LMM context, we are capable of calibrating the correlation structure which is induced by spread market prices. We hope that this calibration is qualitatively better than the calibration on swaps due to the higher sensitivity of CMS spread prices to correlation forwards .
First we observe: The CMS spread dynamics of S^{i,c_1,c_2} can be written in direct dependency on forward rates. By assuming without limitation $c_1 > c_2$ we write

$$\begin{aligned} S_t^{i,c_1,c_2} &= S_t^{i,i+c1} - S_t^{i,i+c2} - F \\ &= \sum_{l=i}^{i+c_1-1} \omega_l^{i,i+c_1}(t) F_t^l - \sum_{l=i}^{i+c_2-1} \omega_l^{i,i+c_2}(t) F_t^l - F, \end{aligned}$$

where $\omega_l^{i,i+c_1}(t) := \frac{\tau_k P(t,T_{l+1})}{A_t^{i,i+c1}}$ are the well-known stochastic weights from (2.6). Now we define

$$\omega_l^{i,i+c_2} := 0 \text{ for all } l > i + c_2 - 1$$

and

$$v_l^{i,c1,c2}(t) := \omega_l^{i,i+c_1}(t) - \omega_l^{i,i+c_2}(t) \text{ for all } l \in \{i, \dots, i + c1 - 1\}.$$

Hence, for the CMS spread we can write

$$S_t^{i,c_1,c_2} = \sum_{l=i}^{i+c_1-1} v_l^{i,c_1,c_2}(t) F_t^l - F. \tag{7.2}$$

Therefore, a CMS spread is a portfolio of several forward rates and each rate is weighted with $v_l^{i,c1,c2}$. This notation further reveals the possibility of gaining insights into the correlation of forwards through CMS.
In this section we want to show how the correlation structure can be calibrated on prices of European options on CMS spreads. First, we analyze the dynamics of CMS and give approximations for those in a SABR-like environment under the abstract spread measure. Second, we show how to transform CMS prices evaluated under the spread measure to prices under a forward measure. Finally, we explain the calibration in detail.

7.1. Markovian Projections of CMS spreads

If we want to examine the CMS dynamics (7.2), we first have to simplify it. The reason for that is that the process in (7.2) is the sum of correlated SABR processes, which are weighted stochastically. Therefore, the resulting process may again be a SABR process, as foreshadowed in [4], but it is impossible to calculate the exact distribution or dynamics.
In the following we will project a portfolio dynamics driven by a number of random processes (7.2) on a *displaced* SABR-proces. Here an important tool is the technique of Markovian Projection, which was developed by *Gyöngy* in [19]. This technique was extended by *Brunick* and *Shreve* in [15]. We use the latter to give the following proposition

Proposition 5. *(Markovian Projection)*
Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbf{P})$ be a filtrated probability space fulfilling the usual conditions and on which $(W_t)_{t\geq 0}$ is a d-dimensional Wiener-process. Further, let b and $(\sigma_t)_t$ adapted processes with values in $\mathbb{R}^d$ and $\mathbb{R}^{d\times d}$, respectively.
Let X be a $\mathbb{R}^d$-valued process given through the SDE

$$dX_t = b_t dt + \sigma_t dW_t, \;\; t \geq 0, \;\; X_0 = x. \tag{7.3}$$

Then there exists a $\mathbb{R}^d$-valued measurable function $\widehat{b}$, a $\mathbb{R}^{d\times d}$-valued measurable function $\widehat{\sigma}$ and a Lebesgue-null-set N, so that

$$\widehat{b}(t, X_t) = \mathbb{E}\Big[b_t \;\Big|\; X_t\Big] \tag{7.4}$$

$$\widehat{\sigma}(t, X_t)\widehat{\sigma}^{tr}(t, X_t) = \mathbb{E}\Big[\sigma_t \sigma_t^{tr} \;\Big|\; X_t\Big] \; \mathbf{P}\text{-}a.s. \text{ and all } t \in N \tag{7.5}$$

holds. Here $\bullet^{tr}$ is the trace operator. Furthermore, there exits a filtrated probability space $(\widehat{\Omega}, \widehat{\mathcal{F}}, (\widehat{\mathcal{F}}_t)_{t\geq 0}, \widehat{\mathbf{P}})$ that supports a $\mathbb{R}^d$-valued adapted process $\widehat{X}$ and a d-dimensional Wiener-process $\widehat{W}$, satisfying

$$d\widehat{X}_t = \widehat{b}(t, \widehat{X}_t)d_t + \widehat{\sigma}(t, X_t)d\widehat{W}_t, \;\; t \geq 0, \widehat{X}_0 = x \tag{7.6}$$

and such that for all $t \geq 0$ the distribution of $\widehat{X}_t$ under $\widehat{P}$ agrees with the distribution of X_t under P.

Proof. A proof can be found in [15]. □

The proposition enables us to project a process of the form (7.3) on a process of the form (7.6) with coefficients given through (7.4) and (7.5). This new process is then in every case *weak* Markovian. Further, the distribution of X_t and $\widehat{X}_t$ is the same for each t. However, it is not the case that the distribution of individual paths coincide. So the distribution of $(X_t)_{t\in I}$ and $(\widehat{X}_t)_{t\in I}$ for some index set I may not agree. For this reason, the Markovian Projection is not suited to simplify a process first and then evaluating a path dependent options afterwards. In the case of European options we don't face those problems since they only depend on the distribution of the underlying process at some fixed point T. So, we don't run into any problems if we simplify a CMS spread to evaluate European options.

The coefficients in (7.4) and (7.5) obtained by using the proposition on CMS spreads can be easier understood and approximated then the original ones. Our goal is not to calculate the resulting process in (7.6) accurately. For that the process under which we have to condition will be too complicated. We rather want to approximate condition expectations in the frame of another model, the displaced SABR model. We will introduce this model in the following, but first let's take a look at a simple example for a Markovian Projection to get a better feeling for it. In addition, the example shows that not every projected process is *strong* Markovian.

Example (From Brunick and Shreve [15]).
Let in proposition 5 the parameters be $d = 1, b = 0$ and the initial process $dX_t = \sigma_t dW_t$ with

$$\sigma_t := \mathbb{1}_{(1,\infty)}\mathbb{1}_{\{W_1>0\}},$$

and $X_0 = 0$. Clearly it holds

$$X_t = \mathbb{1}_{(1,\infty)}(t)\mathbb{1}_{\{W_1>0\}}\big[W_t - W_1\big].$$

Now we calculate the Markovian Projection as

$$\widehat{\sigma}_t^2(t, X_t) = \mathbb{E}\Big[\sigma_t^2 \;\Big|\; X_t\Big] = \begin{cases} 1, & \text{falls } X_t \neq 0 \\ 0, & \text{falls } X_t = 0. \end{cases}$$

It follows

$$\widehat{X}_t^{(1)} \equiv 0$$

and

$$\widehat{X}_t^{(2)} = \mathbb{1}_{(1,\infty)}(t)\big[W_t - W_1\big]$$

agree with X on $\{X_1 < 0\}$ and $\{X_1 > 0\}$, respectively. However, $\widehat{X}_t^1$ and $\widehat{X}_t^2$ don't have the right distribution to be a projection of X on both subsets. To obtain a proper projection consider $Y \sim \mathrm{Ber}(0, \frac{1}{2})$ independent of $(W_t)_{t\geq 0}$ and $\mathcal{F}_0$ measurable. So $\mathcal{F}_0$ is non-trivial. Define

$$\widehat{X}_t := Y\widehat{X}_t^{(1)} + (1-Y)\widehat{X}_t^{(2)}.$$

Taking into account $W_1 \sim \mathcal{N}(0,1)$ and therefore $\mathbb{P}(W_1 < 0) = \frac{1}{2}$ follows $\widehat{X}_t \overset{d}{=} X_t$ for all $t \geq 0$.
However, $\widehat{X}$ is not *strong* Markovian. This can be seen by defining the stopping time

$$\tau_2 := \inf\{t \geq 2 \mid X_t > 0\}$$

This stopping time is not independent from $\mathcal{F}_0$ because of Y and therefore $(\widehat{X}_{\tau_2+s})_{s\geq 0}$ conditioned on $\tau_2 < \infty$ is not independent from $\mathcal{F}_0$. $\square$

Now that we we have a better understanding for the Markovian Projection we use the technique on CMS spreads. The following theorem gives a approximation of the spread dynamics (7.2) in a context of a *displaced* SABR model in dependence of forward rates F^i which follow a SABR-LMM model. The displaced SABR model coefficients depend on the correlation matrix P and therefore the theorem gives the link between forward correlations and spread dynamics.

Theorem 4 (The Displaced SABR Model for CMS spreads).
For a strike price $F = 0$ the dynamics of the CMS spread

$$S_t^{i,c_1,c_2} = \sum_{l=i}^{i+c_1-1} v_l^{i,c_1,c_2}(t) F_t^l. \tag{7.7}$$

can be approximated under the spread measure $\mathbb{P}^{CMSs}$ via

$$dS_t^{i,c_1,c_2} = u_t G(S_t^{i,c_1,c_2}) d\widehat{W}_t, \; S_0^{i,c_1,c_2} = \sum_{l=i}^{i+c_1-1} v_l^{i,c_1,c_2} F_0^l \tag{7.8}$$

$$du_t = A^{p,v}(\rho, r) d\widetilde{Z}_t, \; u_0 = 1 \tag{7.9}$$

$$d\langle S_\bullet^{i,c_1,c_2}, u_\bullet \rangle_t = \mathcal{X}(\rho, r, R) dt, \tag{7.10}$$

where G is defined via

$$G(x) = (x - S_0^{i,c_1,c_2} + \frac{p^{i,c_1,c_2}}{q^{i,c_1,c_2}}) q^{i,c_1,c_2}. \tag{7.11}$$

The coefficient a given in the proof at (7.40), (7.41), (7.36), (7.37), (7.42), (7.43), (7.33) and (7.34).

The proof is split into several parts. The first step is to project the dynamics (7.7) with the help of the Markovian Projection on a simpler process. The coefficients of the projection are then approximated such that we are able to express the spread dynamics in a *displaced* SABR model. This model is defined as

Definition 10. *(Displaced SABR model)*
The displaced SABR model is given through

$$\begin{aligned} dS_t &= u_t G(S_t) dW_t, \; \textit{für alle } t \geq 0, \; S_0 = x, \\ du_t &= \eta u_t dZ_t, \; \textit{für alle } t \geq 0, \; u_0 = 1, \\ d\langle W, Z \rangle_t &= \gamma dt, \end{aligned} \tag{7.12}$$

where $\eta > 0$ and $\gamma \in [-1, 1]$. The processes W and Z are one-dimensional Wiener processes. Further, $G : \mathbb{R} \longrightarrow \mathbb{R}$ is a Borel measurable function.

In the proof the form of G is given through the approximation of the projections parameters. Then the dynamics of u and therefore the parameter η and γ, respectively, are calculated. We will see that all parameters will depend on the correlation structure of the SABR-LMM model.

The theorem 4 was originally proven by *Kienitz* and *Wittke* in [22] for a simpler SABR-LMM model with constant parameters with methods from [41]. The version stated here is a real extension, since we approximate the spread dynamics in a SABR-LMM environment with time-dependent parameters g_t^i and h_t^i. Further we deal with the more complicated volatility processes $\sigma_t^i = g_t^i k_t^i$.

Remark. In the proof we will assume that the g^i and h^i are already calibrated to caplets as in chapter 6. So the parametrization of the g^i fulfill (6.3) and the one of the h^i (6.5). This doesn't limitate our approach, since we only want to calibrate the suppercorrelation matrix P based on CMS spreads. As in the case for the calibration to swaptions, the volatility structure is calibrated to caplets.

Proof of Theorem 4. First, we define

$$u_t^l := \frac{\sigma_t^l}{\sigma_0^l} \text{ for all } l \in \{1, \dots, N-1\}$$

and

$$\varphi(F_t^l) := \sigma_0^l (F_t^l)^\beta \text{ for all } l \in \{1, \dots, N-1\}.$$

Therefore $u_0^l = 1$ and $dF_t^l = \mu_t^l dt + u_t^l \varphi(F_t^l) dW_t^l$ under the measure $\mathbb{P}^{\text{CMSs}}$ - the measure under which the CMS spread is a local martingale - for all $l \in \{1, \dots, N-1\}$, where μ_t^l are drifts coming from the change of measures. Further we set

$$\begin{aligned} p_k &:= \varphi(F_0^l) = \sigma_0^k (F_0^l)^\beta \\ q_k &:= \varphi(F_0^l)' = \sigma_0^k \beta (F_0^l)^{\beta - 1}. \end{aligned}$$

Now, we use the dynamics in (7.7) and freeze the weights v_l^{i,c_1,c_2} at their initial values in $t = 0$. Under $\mathbb{P}^{\text{CMSs}}$ we obtain

$$\begin{aligned} dS_t^{i,c1,c2} &= d \sum_{l=1}^{i+c1-1} v_l^{i,c_1,c_2}(t) F_t^l \\ &\approx \sum_{l=1}^{i+c1-1} v_l^{i,c_1,c_2}(0) dF_t^l \\ &\approx \sum_{l=1}^{i+c1-1} v_l^{i,c_1,c_2}(0) u_t^l \varphi(F_t^l) dW_t^l, \end{aligned}$$

where in the last approximation the drifts μ_t^l are neglected. Now define

$$d\widehat{W}_t := \sigma_t^{-1} \sum_{l=i}^{i+c1-1} v_l^{i,c_1,c_2}(0) u_t^l \varphi(F_t^l) dW_t^l \tag{7.13}$$

and

$$\begin{aligned} \sigma_t^2 := &\sum_{l=i}^{i+c1-1} \big(v_l^{i,c_1,c_2}(0)\big)^2 \big(u_t^l\big)^2 \varphi(F_t^l)^2 \\ &+ 2 \sum_{k<l=i+1}^{i+c1-1} v_k^{i,c_1,c_2}(0) v_l^{i,c_1,c_2}(0) u_t^k u_t^l \varphi(F_t^k) \varphi(F_t^l). \end{aligned}$$

Hence, the Levy characterisation [20] gives us: $\widehat{W}$ is a Wiener-Process, since $\widehat{W}_0 = 0$, $\langle \widehat{W} \rangle_t = t$ and $\widehat{W}$ is continuous. The spread dynamics can now be written as

$$dS_t^{i,c_1,c_2} = \sigma_t d\widehat{W}_t. \tag{7.14}$$

Our goal is to describe the function G in (7.12) using (7.2). In order to achieve that we define

$$\begin{aligned} u_t^2 := \frac{1}{p^2} \Big[&\sum_{k=i}^{i+c_1-1} \big(v_k^l(0)\big)^2 p_k^2 \big(u_t^k\big)^2 \\ &+ 2 \sum_{k<l=i+1}^{i+c_1-1} p_k p_l u_t^k u_t^l v_k^{i,c_1,c_2}(0) v_l^{i,c_1,c_2}(0) \rho_{kl} \Big], \end{aligned} \tag{7.15}$$

with

$$p := \Big[\sum_{k=i}^{i+c_1-1} \big(v_k^l(0)\big)^2 p_k^2 + 2 \sum_{k<l=i+1}^{i+c_1-1} p_k p_l v_k^{i,c_1,c_2}(0) v_l^{i,c_1,c_2}(0) \rho_{kl} \Big]^{\frac{1}{2}}.$$

In particular, this implies $u_0 = 1$ as needed in the displaced SABR model. If we apply proposition 5 to (7.14), we get

$$\widehat{\sigma}(t,x) = \mathbb{E}^{\mathrm{CMSs}}\Big[\sigma_t^2 \;\Big|\; S_t^{i,c_1,c_2} = x\Big] \tag{7.16}$$

and on the other hand, if we assume that $G(S_t^{i,c_1,c_2})$ and u_t^2 are under $\sigma(S_t^{i,c_1,c_2})$ approximately stochastic independent, we get for the projection of our displaced SABR dynamics

$$\widehat{\sigma}(t,x) = \mathbb{E}^{\mathrm{CMSs}}\Big[u_t^2 \;\Big|\; S_t^{i,c_1,c_2} = x\Big] G^2(x). \tag{7.17}$$

For (7.16) and (7.17) being equivalent it has to hold

$$G^2(x) = \frac{\mathbb{E}^{\mathrm{CMSs}}\Big[\sigma_t^2 \;\Big|\; S_t^{i,c_1,c_2} = x\Big]}{\mathbb{E}^{\mathrm{CMSs}}\Big[u_t^2 \;\Big|\; S_t^{i,c_1,c_2} = x\Big]}, \tag{7.18}$$

which implies the exact structure of G. Our next goal will be to approximate the conditional expectations in (7.18) as well as possible. For that we will rewrite σ_t^2 and u_t^2. We define

$$f_{k,l} := \varphi(F_t^k)\varphi(F_t^l) u_t^k u_t^l$$

and

$$g_{k,l}(t) := \frac{p_k p_l u_t^k u_t^l}{p^2}.$$

Hence

$$\sigma_t^2 = \sum_{l=i}^{i+c_1-1} f_{l,l}(t)\big(v_l^{i,c_1,c_2}(0)\big)^2 + 2 \sum_{k<l=i+1}^{i+c_1-1} f_{k,l} v_k^{i,c_1,c_2}(0) v_l^{i,c_1,c_2}(0) \rho_{kl} \tag{7.19}$$

and

$$u_t^2 = \sum_{l=i}^{i+c_1-1} g_{k,k}(t)\big(v_l^{i,c_1,c_2}(0)\big)^2 + 2\sum_{k<l=i+1}^{i+c_1-1} g_{k,l} v_k^{i,c_1,c_2(0)} v_l^{i,c_1,c_2}(0)\rho_{kl}. \tag{7.20}$$

If we interpret $f_{k,l}$ as a function depending on F_t^k, F_t^l, u_t^k and u_t^l we get

$$f_{k,l}(t) = f(F_t^k, F_t^l, u_t^k, u_t^l),$$

with $f(x,y,z,w) = \varphi(x)\varphi(y)zw$. Therefore a Taylor series extension around the point $(F_0^l, F_0^k, 1, 1)$ in direction of $(F_t^l, F_t^k, u_t^l, u_t^k)$ gives

$$f_{k,l}(t) \approx p_k p_l\big(1 + \frac{q_k}{p_l}(F_t^k - F_0^k) + \frac{q_l}{p_l}(F_t^l - F_0^l) + (u_t^k - 1) + (u_t^l - 1)\big) \tag{7.21}$$

and it follows in the same manners that

$$g_{k,l} \approx \frac{p_k p_l}{p^2}\big(1 + (u_t^k - 1) + (u_t^l - 1)\big). \tag{7.22}$$

To calculate the conditional expectations in (7.18) we just need to handle expressions of the form

$$\mathbb{E}^{\text{CMSs}}\Big[F_t^l - F_0^l \;\Big|\; S_t^{i,c_1,c_2} = x\Big] \tag{7.23}$$

and

$$\mathbb{E}^{\text{CMSs}}\Big[u_t^l - 1 \;\Big|\; S_t^{i,c_1,c_2} = x\Big]. \tag{7.24}$$

Our plan is to calculate (7.23) via the definition of the conditional expectation as a orthogonal projection. For that we approximate our SABR-LMM like in [22] as follows

$$\begin{aligned} dF_t^i &= p_i d\widetilde{W}_t^i,\ F_0^i = F^i(0) \\ du_t^i &= \nu_i d\widetilde{Z}_t^i,\ u_0^i = 1 \\ d\langle \widetilde{W}_\bullet^i, \widetilde{W}_\bullet^j\rangle_t &= \rho_{ij} dt \\ d\langle \widetilde{Z}_\bullet^i, \widetilde{Z}_\bullet^j\rangle_t &= r_{ij} dt \\ d\langle \widetilde{W}_\bullet^i, \widetilde{Z}_\bullet^j\rangle_t &= R_{ij} dt, \end{aligned} \tag{7.25}$$

where the $(\nu_i)_{i\in\{1,\dots,N-1\}}$ are the vol-vol parameters for the forwards obtained from ordinary SABR fitting and the $(\widetilde{W}^i)_{i\in\{1,\dots,N-1\}}$, $(\widetilde{Z}^i)_{i\in\{1,\dots,N-1\}}$ Wiener processes. In addition, we approximate the spread dynamic as

$$dS_t^{i,c_1,c_2} = pd\widetilde{W}_t.$$

So we just freeze the coefficients in (7.14) at $t = 0$. That implies

$$\begin{aligned} d\widetilde{W}_t &= \sigma_0^{-1} \sum_{l=i}^{i+c1-1} v_l^{i,c_1,c_2}(0) u_0^l \varphi(F_0^l) d\widetilde{W}_t^l \\ &= p^{-1} \sum_{l=i}^{i+c1-1} v_l^{i,c_1,c_2}(0) p_l d\widetilde{W}_t^l. \end{aligned} \tag{7.26}$$

Now we get

$$\begin{aligned} d\langle \widetilde{W}_\bullet, \widetilde{W}_\bullet^i \rangle_t &= p^{-1} \sum_{l=i}^{i+c_1-1} \nu_l^{i,c_1,c_2} p_l \rho_{li} dt \\ =: \rho_i^W(\rho) dt \end{aligned} \tag{7.27}$$

and

$$\begin{aligned} d\langle \widetilde{W}_\bullet, \widetilde{Z}_\bullet^i \rangle_t &= p^{-1} \sum_{l=i}^{i+c_1-1} \nu_l^{i,c_1,c_2} p_l R_{li} dt \\ =: \rho_i^Z(R) dt. \end{aligned} \tag{7.28}$$

We estimate (7.23) with the help of (7.27) and while keeping (7.25) in mind as

$$\begin{aligned} \mathbb{E}^{\mathrm{CMSs}}\Big[F_t^l - F_0^l \;\Big|\; S_t^{i,c_1,c_2} = x\Big] &= \frac{\langle F_\bullet^i, S_\bullet^{i,c_1,c_2}\rangle_t}{\langle S_\bullet^{i,c_1,c_2}, S_\bullet^{i,c_1,c_2}\rangle_t}(x - S_t^{i,c_1,c_2}) \\ &= \frac{p_i \rho_i^W(\rho)}{p}(x - S_t^{i,c_1,c_2}) \end{aligned} \tag{7.29}$$

and with the help of (7.28) we get

$$\begin{aligned} \mathbb{E}^{\mathrm{CMSs}}\Big[u_t^l - 1 \;\Big|\; S_t^{i,c_1,c_2} = x\Big] &= \frac{\langle u_\bullet^i, S_\bullet^{i,c_1,c_2}\rangle_t}{\langle S_\bullet^{i,c_1,c_2}, S_\bullet^{i,c_1,c_2}\rangle_t}(x - S_t^{i,c_1,c_2}) \\ &= \frac{\nu_i \rho_i^Z(R)}{p}(x - S_t^{i,c_1,c_2}). \end{aligned} \tag{7.30}$$

With (7.19) and (7.21) in addition with the help of (7.29) and (7.30) it follows

$$
\begin{aligned}
\mathbb{E}^{\text{CMSs}}\Big[\sigma_t^2 \;\Big|\; S_t^{i,c_1,c_2} = x\Big] \approx & \sum_{l=i}^{i+c_1-1} v_l^{i,c_1,c_2}(0)^2 p_l^2 \Big[1 + 2\frac{q_l}{p_l}\frac{p_l\rho_l^W(\rho)}{p}(x - S_0^{i,c_1,c_2}) \\
& + 2\frac{\nu_l\rho_l^Z(R)}{p}(x - S_t^{i,c_1,c_2})\Big] \\
& + 2\sum_{j<l=i+1}^{i+c_1-1} \rho_{jl} v_l^{i,c_1,c_2}(0) v_j^{i,c_1,c_2}(0) p_j p_l \\
& \times \Big[1 + \frac{q_j}{p_j}\frac{p_j\rho_j^W(\rho)}{p}(x - S_0^{i,c_1,c_2}) \\
& + \frac{q_l}{p_l}\frac{p_l\rho_l^W(\rho)}{p}(x - S_0^{i,c_1,c_2}) + \frac{\nu_j\rho_j^Z(R)}{p}(x - S_t^{i,c_1,c_2}) \\
& + \frac{\nu_l\rho_l^Z(R)}{p}(x - S_t^{i,c_1,c_2})\Big] \\
= & \sum_{l=i}^{i+c_1-1} v_l^{i,c_1,c_2}(0)^2 p_l^2 \\
& + 2\sum_{j<l=i+1}^{i+c_1-1} \rho_{jl} v_l^{i,c_1,c_2}(0) v_j^{i,c_1,c_2}(0) p_j p_l \\
& + \Bigg[\sum_{l=i}^{i+c_1-1} 2p_l^2 v_l^{i,c_1,c_2}(0)^2 (A_l(\rho) + B_l(\rho)) \\
& + 2\sum_{j<l=i+1}^{i+c_1-1} \rho_{jl} v_j^{i,c_1,c_2}(0) v_l^{i,c_1,c_2}(0) p_j p_l \Big(A_j(\rho) \\
& + A_l(\rho) + B_j(\rho) + B_l(\rho)\Big)\Bigg](x - S_0^{i,c_1,c_2}),
\end{aligned} \tag{7.31}
$$

where we define

$$
A_l(\rho) := \frac{q_l\rho_l^W(\rho)}{p} \tag{7.32}
$$

and

$$
B_l(\rho) := \frac{\nu_l\rho_l^Z(R)}{p}(x - S_t^{i,c_1,c_2}).
$$

Further we set

$$\begin{aligned}A_0(\rho) := &\sum_{l=i}^{i+c_1-1} 2p_l^2 v_l^{i,c_1,c_2}(0)^2 (A_l(\rho) + B_l(\rho))\\ &+ 2\sum_{j<l=i+1}^{i+c_1-1} v_j^{i,c_1,c_2}(0) v_l^{i,c_1,c_2}(0) p_j p_l\\ &\times \big(A_j(\rho) + A_l(\rho) + B_j(\rho) + B_l(\rho)\big).\end{aligned} \tag{7.33}$$

Overall, we get

$$\mathbb{E}^{\mathrm{CMSs}}\Big[\sigma_t^2 \;\Big|\; S_t^{i,c_1,c_2} = x\Big] \approx p^2 + A_o(\rho)(x - S_0^{i,c_1,c_2}).$$

Now we have evaluated the enumerator of equation (7.18). To approximate the denominator we use (7.20) and (7.22) together with (7.30) and obtain

$$\begin{aligned}\mathbb{E}^{\mathrm{CMSs}}\Big[u_t^2 \;\Big|\; S_t^{i,c_1,c_2} = x\Big] \approx &\sum_{l=i}^{i+c_1-1} v_l^{i,c_1,c_2}(0)^2 \frac{p_l^2}{p^2}\Big(1 + 2\frac{\nu_i \rho_i^Z(R)}{p}(x - S_0^{i,c_1,c_2})\Big)\\ &+ 2\sum_{j<l=i+1}^{i+c_1-1} \rho_{jl} v_j^{i,c_1,c_2}(0) v_l^{i,c_1,c_2}(0) \frac{p_j p_l}{p^2}\\ &\times \left[1 + \frac{\nu_i \rho_i^Z(R)}{p}(x - S_0^{i,c_1,c_2}) + \frac{\nu_i \rho_i^Z(R)}{p}(x - S_0^{i,c_1,c_2})\right]\\ &= \sum_{l=i}^{i+c_1-1} v_l^{i,c_1,c_2}(0)^2 \frac{p_k^2}{p^2}\\ &+ 2\sum_{j<l=i+1}^{i+c_1-1} \rho_{jl} v_j^{i,c_1,c_2}(0) v_l^{i,c_1,c_2}(0) \frac{p_j p_l}{p^2}\\ &+ \left[\sum_{l=i}^{i+c_1-1} v_l^{i,c_1,c_2}(0)^2 \frac{p_k^2}{p^2} 2B_l\right.\\ &\left.+ 2\sum_{j<l=i+1}^{i+c_1-1} \rho_{jl} v_j^{i,c_1,c_2}(0) v_l^{i,c_1,c_2}(0) \frac{p_j p_l}{p^2}\Big[B_j(\rho) + B_l(\rho)\Big]\right]\\ &\times (x - S_0^{i,c_1,c_2})\\ &= 1 + A_u(\rho)(x - S_0^{i,c_1,c_2}),\end{aligned}$$

where

$$A_u(\rho) := \Bigg[\sum_{l=i}^{i+c_1-1} v_l^{i,c_1,c_2}(0)^2 \frac{p_k^2}{p^2} 2B_l(\rho) + 2 \sum_{j<l=i+1}^{i+c_1-1} \rho_{jl} v_j^{i,c_1,c_2}(0) v_l^{i,c_1,c_2}(0) \frac{p_j p_l}{p^2} \Big[B_j(\rho) + B_l(\rho) \Big] \Bigg]. \tag{7.34}$$

Now we can use (7.18) to get the form of G. We get

$$G^2(x) \approx \frac{p^2 + A_o(\rho, R)(x - S_0^{i,c_1,c_2})}{1 + A_u(\rho, R)(x - S_0^{i,c_1,c_2})}. \tag{7.35}$$

Hence

$$G^2(S_0^{i,c_1,c_2}) \approx p^2 \tag{7.36}$$

and for the derivative

$$\begin{aligned} \frac{d}{dx} G(S_0^{i,c_1,c_2}) &= \frac{A_o(\rho, R) - p^2 A_u(\rho)}{2p} \\ &= p^{-1} \Big(\sum_{l=i}^{i+c_1-1} p_l^2 v_l^{i,c_1,c_2}(0)^2 A_l(\rho) \\ &+ \sum_{j<l=i+1}^{i+c_1-1} v_j^{i,c_1,c_2}(0) v_l^{i,c_1,c_2}(0) p_j p_l \\ &\times \big(A_j(\rho) + A_l(\rho) \big) \Big). \end{aligned} \tag{7.37}$$

Now we will calculate the dynamic of u_t, while approximating $\frac{u_t^j u_t^l}{u_t}$ and $\frac{(u_t^l)^2}{u_t}$ via their expectation values as in [22]. The expectation values are $\mathbb{E}^{\text{CMSs}}\Big[\frac{u_t^j u_t^l}{u_t}\Big] \approx 1 \approx \mathbb{E}^{\text{CMSs}}\Big[\frac{(u_t^l)^2}{u_t}\Big]$. We obtain using

$$\begin{aligned} du_t^l &= \frac{1}{\sigma_0^l} d\sigma \approx \frac{g_t^l}{\sigma_0^l} dk_t^l \\ &= \frac{g_t^l}{\sigma_0^l} k_t^l \zeta_l h_t^l dZ_t^l \approx u_t^l \nu_l dZ_t^l \end{aligned}$$

the following

$$\begin{aligned}
du_t^2 &= \frac{1}{p^2}\left(2\sum_{j<l=i+1}^{i+c_1-1} p_j p_l v_j^{i,c_1,c_2}(0) v_l^{i,c_1,c_2}(0)\rho_{jl}\left[u_t^j du_t^l + u_t^l du_t^k\right]\right.\\
&\left.+\,2\sum_{l=i}^{i+c_1-1}\left(v_l^{i,c_1,c_2}(0)\right)^2 p_l^2 u_t^l du_t^l\right)\\
&\approx \frac{1}{p^2}\left(2\sum_{j<l=i+1}^{i+c_1-1} \rho_{jl} p_j p_l v_j^{i,c_1,c_2}(0) v_l^{i,c_1,c_2}(0) u_t^j u_t^l\left[\nu_l dZ_t^l + \nu_j dZ_t^j\right]\right.\\
&\left.+\,2\sum_{l=i}^{i+c_1-1}\left(v_l^{i,c_1,c_2}(0)\right)^2 p_l^2 \left(u_t^l\right)^2 \nu_l dZ_t^l\right)\\
&= \frac{2}{p^2}\left[\sum_{l=i}^{i+c_1-1}\left[\left(v_l^{i,c_1,c_2}(0)\right)^2 p_l^2 + \sum_{l\neq j=i}^{i+c_1-1} p_j p_l v_j^{i,c_1,c_2}(0) v_l^{i,c_1,c_2}(0)\rho_{jl} u_t^j u_t^l\right]\nu_l dZ_t^l\right].
\end{aligned}$$

Since $u_0 > 0$ and u^l never changes its sign we obtain

$$\begin{aligned}
\frac{du_t}{u_t} &= \frac{d\sqrt{u_t^2}}{u_t}\\
&\approx \frac{1}{2}\frac{1}{u_t^2} du_t^2\\
&\approx \frac{1}{p^2}\left[\sum_{l=i}^{i+c_1-1}\left[\left(v_l^{i,c_1,c_2}(0)\right)^2 p_l^2\right.\right.\\
&\left.\left.+\sum_{l\neq j=i}^{i+c_1-1} p_j p_l v_j^{i,c_1,c_2}(0) v_l^{i,c_1,c_2}(0)\rho_{jl}\right]\nu_l dZ_t^l\right]. \qquad (7.38)
\end{aligned}$$

Now (7.38) leads to

$$\frac{du_t}{u_t} \approx A^{p,v}(\rho, r) d\widetilde{Z}_t,$$

where

$$\begin{aligned}\widetilde{Z}_t &= \frac{1}{A^{p,v}(\rho,r)p^2}\Bigg[\sum_{l=i}^{i+c_1-1}\bigg[\Big(v_l^{i,c_1,c_2}(0)\Big)^2 p_l^2 \\ &\quad+\sum_{l\neq j=i}^{i+c_1-1} p_j p_l v_j^{i,c_1,c_2}(0) v_l^{i,c_1,c_2}(0)\rho_{jl}\bigg]\nu_l dZ_t^l\Bigg] \\ &= \frac{1}{A^{p,v}(\rho,r)p^2}\sum_{l=i}^{i+c_1-1}\lambda_l(\rho)\nu_l dZ_t^l, \end{aligned} \tag{7.39}$$

with

$$\lambda_l(\rho) := \Big(v_l^{i,c_1,c_2}(0)\Big)^2 p_l^2 + \sum_{l\neq j=i}^{i+c_1-1} p_j p_l v_j^{i,c_1,c_2}(0) v_l^{i,c_1,c_2}(0)\rho_{jl}.$$

The variable $A^{p,v}(\rho,r)$ is calculated such that $\widetilde{Z}$ is a Wiener process

$$\begin{aligned}\Big(A^{p,v}(\rho,r)\Big)^2 &:= \frac{1}{t}\langle\frac{1}{p^2}\sum_{l=i}^{i+c_1-1}\varepsilon_l(\rho)\nu_l dZ_\bullet^l\rangle_t \\ &= \frac{1}{p^4}\sum_{j,l=i}^{i+c_1-1}\lambda_j(\rho)\lambda_l(\rho)r_{jl}\nu_j\nu_l. \end{aligned} \tag{7.40}$$

Our last step is to calculate the correlation between S_t^{i,c_1,c_2} and u_t. We get using (7.13), (7.39) and $\sigma_0 = p$

$$\begin{aligned} d\langle\widehat{W},\widetilde{Z}\rangle_t &= \frac{1}{A^{p,v}(\rho,r)}\frac{1}{\sigma_t}\frac{1}{p^2}\langle\sum_{l=i}^{i+c_1-1}v_l^{i,c_1,c_2}(0)u_\bullet^l\varphi(F_\bullet^l)dW_\bullet^l,\sum_{l=i}^{i+c_1-1}\lambda_l(\rho)\nu_l dZ_\bullet^l\rangle_t \\ &\approx \frac{1}{A^{p,v}(\rho,r)p^3}\langle\sum_{l=i}^{i+c_1-1}v_l^{i,c_1,c_2}(0)\sigma_l F_0^l dW_\bullet^l,\sum_{l=i}^{i+c_1-1}\lambda_l(\rho)\nu_l dZ_\bullet^l\rangle_t \\ &= \frac{1}{A^{p,v}(\rho,r)p^3}\sum_{j,l=i}^{i+c_1-1}v_j^{i,c_1,c_2}(0)\sigma_j F_0^j\lambda_l(\rho)\nu_l R_{jl} \\ &=: \mathcal{X}(\rho,r,R). \end{aligned} \tag{7.41}$$

Therefore we have approximative for the CMS spread dynamics with the results in

(7.40), (7.41), (7.36), (7.37), (7.33) and (7.34)

$$\begin{aligned}
dS_t^{i,c_1,c_2} &= u_t G(S_t^{i,c_1,c_2}) d\widehat{W}_t,\ S_0^{i,c_1,c_2} = \sum_{l=i}^{i+c_1-1} v_l^{i,c_1,c_2} F_0^l \\
du_t &= A^{p,v}(\rho, r) d\widetilde{Z}_t,\ u_0 = 1 \\
d\langle S_\bullet^{i,c_1,c_2}, u_\bullet \rangle_t &= \mathcal{X}(\rho, r, R) dt,
\end{aligned}$$

where

$$G(S_0^{i,c_1,c_2}) = p =: p^{j,c_1,c_2} \tag{7.42}$$

and

$$\begin{aligned}
\frac{d}{dx} G(S_0^{i,c_1,c_2}) = p^{-1} \Big(& \sum_{l=i}^{i+c_1-1} p_l^2 v_l^{i,c_1,c_2}(0)^2 A_l(\rho) \\
& + \sum_{j<l=i+1}^{i+c_1-1} v_j^{i,c_1,c_2}(0) v_l^{i,c_1,c_2}(0) p_j p_l \\
& \times \big(A_j(\rho) + A_l(\rho) \big) \Big) =: q^{j,c_1,c_2}.
\end{aligned} \tag{7.43}$$

So G can be approximated via

$$G(x) = (x - S_0^{i,c_1,c_2} + \frac{p^{j,c_1,c_2}}{q^{j,c_1,c_2}}) q^{j,c_1,c_2}.$$

□

Remark. In Theorem 4 the coefficients p^{j,c_1,c_2} and q^{j,c_1,c_2} are obtained by using a freezing procedure to approximate the expected values as in Theorem 5. For short terms the approximation is valid, but the longer the expiry of the spread the more inaccurate the approximation gets. Further, the parameters $A_t^{p,v}$ and $\mathcal{X}(\rho, r, R)$ are tailored to our SABR-LMM with time-dependent coefficients and a volatility structure calibrated to caplets.
Similar to the case of the approximation of the swap dynamics in chapter 4.2 we froze in the beginning of the proof the CMS spread weights $v_j^{i,c_1,c_2}(0)$. The freezing is needed

since the volatility of CMS spreads in the SABR-LMM is stochastic again, because the weights are clearly random, and we want to have deterministic volatility for the displaced SABR model in the end.

7.2. Convexity Correction for CMS spreads

In Theorem 4 we approximate the dynamics of CMS spreads under the spread measure $\mathbb{P}^{\text{CMSs}}$ under which we assume a spread as drift-free. However our goal is to estimate the correlation of forward rates under the forward measure $\mathbb{P}^i$ to be consistent with our estimations. Since we don't assume deterministic volatilities for the processes of interest, the measure under which we calculate the correlations indeed matters. In addition, a spread option fixed in T_i pays in T_{i+1}. Therefore we want to discount this payment by the numéraire $B(t, T_{i+1})$ which belongs to $\mathbb{P}^i$.
The market quotes prices of CMS spread options and we model the underlying processes as drift-less under $\mathbb{P}^{\text{CMSs}}$. Hence, the quoted prices are not calculated under $\mathbb{P}^i$ and we have to do a so-called convexity correction, which helps us to transfer prices gained under one measure to another measure.
Historically a convexity correction is understood as an adjustment to the risk exposure of an asset to changes of the forward or yield curve. The idea was: If the forward curve changes over time, the value of the asset somehow comoves and therefore the expected future price of an asset depends on the yield curve of interest. We will use this idea to calculate the convexity correction for CMS spreads. But first we introduce the mathematical way of thinking about those corrections.
A convexity correction enables us to transform prices in one measure into prices in an other measure by adding a correction term. The correction term is then called *convexity correction.* The added term can be interpreted as some drift correction due to a measure change, just as in Girsanovs Theorem [40]. To describe the convexity correction in formulas we remember that if we have to random variables X and Y then and all integrals are well defined, it holds that

$$\begin{aligned}
\text{Cov}(X, Y) &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\
\Leftrightarrow \mathbb{E}[XY] &= \mathbb{E}[X]\mathbb{E}[Y] + \mathbb{C}\text{ov}(X, Y).
\end{aligned}$$

Now, let's consider two different measures $\mathbb{P}$ and $\mathbb{Q}$ with numéraire B and A, respectively. Then it holds due to the change of numéraire technique, like in the proof of the forwardrate dynamics under different measures in Theorem 1, for a claim S_T which pays in T that

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}\big[S_T\big] &= \frac{A_0}{B_0}\mathbb{E}^{\mathbb{Q}}\big[S_T\frac{B_T}{A_T}\big] \\ &= \frac{A_0}{B_0}\mathbb{E}^{\mathbb{Q}}\big[\frac{B_T}{A_T}\big]\mathbb{E}^{\mathbb{Q}}\big[S_T\big] + \frac{A_0}{B_0}\mathrm{Cov}^{\mathbb{Q}}\big[S_T, \frac{B_T}{A_T}\big] \\ &= \mathbb{E}^{\mathbb{Q}}\big[S_T\big] + \frac{A_0}{B_0}\mathbb{C}\mathrm{ov}^{\mathbb{Q}}\big[S_T, \frac{B_T}{A_T}\big]. \end{aligned} \tag{7.44}$$

Here the term

$$\frac{A_0}{B_0}\mathbb{C}\mathrm{ov}^{\mathbb{Q}}\big[S_T, \frac{B_T}{A_T}\big]$$

is the convexity correction.
To calculate the convexity correction for an CMS spread S^{j,c_1,c_2} we first want to set up a general framework. For a spread S^{j,c_1,c_2} with zero strike we can write

$$S^{j,c_1,c_2} = S^{(1)} - S^{(2)},$$

where the $S^{(i)}$ are swap rates over the time interval $[T_j, T_{j+c_i}]$, where $j \in \mathbb{N}$ is some index and $c_i \in \mathbb{N}$ the tenor of the j-th swap rate. Let $\mathbb{P}^{\mathrm{CMSs}}$ be the spread measure with numéraire SN, let $\mathbb{P}^{j,j+c_i}$ the swap measures for the i'-th swap rate with numéraire $A^{j,j+c_i}$ and define $\mathbb{P}^j$ as the forward measure belonging to the rate F^j. We want to evaluate the spread under $\mathbb{P}^j$ at time T_j, since we have to price European options expiring at T_j later on. By following [22] we write

$$\begin{aligned}
\mathbb{E}^j\left[S_{T_j}^{j,c_1,c_2}\right] &= \mathbb{E}^{\mathrm{CMSs}}\left[S_{T_j}^{j,c_1,c_2}\frac{B(T_j,T_{j+1})}{SN_{T_j}}\frac{SN_0}{B(0,T_{j+1})}\right] \\
&= S_0^{(1)} - S_0^{(2)} \\
&+ \Bigg(\mathbb{E}^{j,j+c_1}\left[S_{T_j}^{(1)}\left(\frac{B(T_j,T_{j+1})}{A_{T_j}^{j,j+c_1}}\frac{A_0^{j,j+c_1}}{B(0,T_{j+1})}-1\right)\right] \\
&- \mathbb{E}^{j,j+c_2}\left[S_{T_j}^{(2)}\left(\frac{B(T_j,T_{j+1})}{A_{T_j}^{j,j+c_2}}\frac{A_0^{j,j+c_2}}{B(0,T_{j+1})}-1\right)\right]\Bigg) \\
&- \Bigg(\mathbb{E}^{j,j+c_1}\left[S_{T_j}^{(1)}\left(\frac{SN_{T_j}}{A_{T_j}^{j,j+c_1}}\frac{A_0^{j,j+c_1}}{SN_0}-1\right)\right] \\
&- \mathbb{E}^{j,j+c_2}\left[S_{T_j}^{(2)}\left(\frac{SN_{T_j}}{A_{T_j}^{j,j+c_2}}\frac{A_0^{j,j+c_2}}{SN_0}-1\right)\right]\Bigg) \\
&\approx \mathbb{E}^{j,j+c_1}\left[S_{T_j}^{(1)}\right] - \mathbb{E}^{j,j+c_2}\left[S_{T_j}^{(2)}\right] \\
&+ \Bigg(\mathbb{E}^{j,j+c_1}\left[S_{T_j}^{(1)}\left(\frac{B(T_j,T_{j+1})}{A_{T_j}^{j,j+c_1}}\frac{A_0^{j,j+c_1}}{B(0,T_{j+1})}-1\right)\right] \\
&- \mathbb{E}^{j,j+c_2}\left[S_{T_j}^{(2)}\left(\frac{B(T_j,T_{j+1})}{A_{T_j}^{j,j+c_2}}\frac{A_0^{j,j+c_2}}{B(0,T_{j+1})}-1\right)\right]\Bigg),
\end{aligned} \tag{7.45}$$

where we used the martingale property of $S^{(i)}$ under $\mathbb{P}^{j,j+c_i}$ and neglected in the last approximation the difference

$$\mathbb{E}^{j,j+c_1}\left[S_{T_j}^{(1)}\left(\frac{SN_{T_j}}{A_{T_j}^{j,j+c_1}}\frac{A_0^{j,j+c_1}}{SN_0}-1\right)\right] - \mathbb{E}^{j,j+c_2}\left[S_{T_j}^{(2)}\left(\frac{SN_{T_j}}{A_{T_j}^{j,j+c_2}}\frac{A_0^{j,j+c_2}}{SN_0}-1\right)\right]. \tag{7.46}$$

According to [22] the term (7.46) is almost zero, since the swap measure and the spread measure evaluate the spreads nearly identical. In order to calculate $\mathbb{E}^j\left[S_{T_j}\right]$ we just have to approximate the convexity corrections

$$\mathbb{E}^{j,j+c_i}\left[S_{T_j}^{(i)}\left(\frac{B(T_j,T_{j+1})}{A_{T_j}^{j,j+c_i}}\frac{A_0^{j,j+c_i}}{B(0,T_{j+1})}-1\right)\right] \tag{7.47}$$

as well as possible. Using the definition of $(\bullet)^+$ we obtain a alternative form for (7.47). We get

$$\begin{aligned}\mathbb{E}^{j,j+c_i}\Big[S_{T_j}^{(i)}\Big(\frac{B(T_j,T_{j+1})}{A_{T_j}^{j,j+c_i}}\frac{A_0^{j,j+c_i}}{B(0,T_{j+1})}-1\Big)\Big] &= S_0^{(i)}\Big(\frac{B(0,T_{j+1})}{A_0^{j,j+c_i}}\frac{A_0^{j,j+c_i}}{B(0,T_{j+1})}-1\Big)\\ &+\mathbb{E}^{j,j+c_i}\Big[\Big(S_{T_j}^{(i)}-S_0^{(i)}\Big)^+\Big(\frac{B(T_j,T_{j+1})}{A_{T_j}^{j,j+c_i}}\frac{A_0^{j,j+c_i}}{B(0,T_{j+1})}-1\Big)\Big]\\ &-\mathbb{E}^{j,j+c_i}\Big[\Big(S_0^{(i)}-S_{T_j}^{(i)}\Big)^+\Big(\frac{B(T_j,T_{j+1})}{A_{T_j}^{j,j+c_i}}\frac{A_0^{j,j+c_i}}{B(0,T_{j+1})}-1\Big)\Big]\\ &=\mathbb{E}^{j,j+c_i}\Big[\Big(S_{T_j}^{(i)}-S_0^{(i)}\Big)^+\Big(\frac{B(T_j,T_{j+1})}{A_{T_j}^{j,j+c_i}}\frac{A_0^{j,j+c_i}}{B(0,T_{j+1})}-1\Big)\Big]\\ &-\mathbb{E}^{j,j+c_i}\Big[\Big(S_0^{(i)}-S_{T_j}^{(i)}\Big)^+\Big(\frac{B(T_j,T_{j+1})}{A_{T_j}^{j,j+c_i}}\frac{A_0^{j,j+c_i}}{B(0,T_{j+1})}-1\Big)\Big].\end{aligned} \tag{7.48}$$

So we find there a call part and a put part. Further, if we apply the call-put parity with strike $S_0^{(i)}$ on the asset $S_{T_j}^{(i)}$, we get

$$\mathbb{E}^{j,j+c_i}\Big[S_{T_j}^{(i)}\Big]=S_0^{(i)}+\mathbb{E}^{j,j+c_i}\Big[(S_{T_j}^{(i)}-S_0^{(i)})^+\Big]-\mathbb{E}^{j,j+c_i}\Big[(S_0^{(i)}-S_{T_j}^{(i)})^+\Big]. \tag{7.49}$$

Together with (7.45) and (7.48) we can conclude the following: To evaluate the convexity correction we just have to calculate the expected values of call/put like options. The call like options are

$$\begin{aligned}&\mathbb{E}^{j,j+c_i}\Big[\Big(S_{T_j}^{(i)}-S_0^{(i)}\Big)^+\Big(\frac{B(T_j,T_{j+1})}{A_{T_j}^{j,j+c_i}}\frac{A_0^{j,j+c_i}}{B(0,T_{j+1})}\Big)\Big]\\ &=\mathbb{E}^{j,j+c_i}\Big[(S_{T_j}^{(i)}-S_0^{(i)})^+\Big]\\ &+\mathbb{E}^{j,j+c_i}\Big[\Big(S_{T_j}^{(i)}-S_0^{(i)}\Big)^+\Big(\frac{B(T_j,T_{j+1})}{A_{T_j}^{j,j+c_i}}\frac{A_0^{j,j+c_i}}{B(0,T_{j+1})}-1\Big)\Big]\end{aligned} \tag{7.50}$$

and the put like options are

$$\begin{aligned}
&\mathbb{E}^{j,j+c_i}\Big[\Big(S_0^{(i)} - S_{T_j}^{(i)}\Big)^+ \Big(\frac{B(T_j, T_{j+1})}{A_{T_j}^{j,j+c_i}} \frac{A_0^{j,j+c_i}}{B(0, T_{j+1})}\Big)\Big] \\
&= \mathbb{E}^{j,j+c_i}\Big[(S_0^{(i)} - S_{T_j}^{(i)})^+\Big] \\
&+ \mathbb{E}^{j,j+c_i}\Big[\Big(S_0^{(i)} - S_{T_j}^{(i)}\Big)^+ \Big(\frac{B(T_j, T_{j+1})}{A_{T_j}^{j,j+c_i}} \frac{A_0^{j,j+c_i}}{B(0, T_{j+1})} - 1\Big)\Big],
\end{aligned} \tag{7.51}$$

where

$$\frac{B(T_j, T_{j+1})}{A_{T_j}^{j,j+c_i}} \frac{A_0^{j,j+c_i}}{B(0, T_{j+1})} - 1$$

vanishes on average and goes, according to *Hagan* [32], to zero linearly with the variance of the swap rate $S^{(i)}$. Therefore the second expected value in (7.50) and (7.51), respectively, should be much smaller than the first one. To evaluate the expressions in (7.50) and (7.51) we have to do two things. First, we have to express the moving of the underlying forward curve in terms of the swap rate $S^{(i)}$. Then we have to use this dependency to evaluate the expected values. Here, the SABR formula for implied volatility will play a major role.

To tackle the first problem we want to write

$$\frac{B(t, T_{j+1})}{A_t^{j,j+c_i}} = H(S_t^{(i)}) \tag{7.52}$$

for some function H. To obtain a proper form for H we assume that there are only parallel shifts in the yield curve and, by following [32], we obtain

$$\begin{aligned}
A_t^{j,j+c_i} &= \sum_{k=j+1}^{j+c_i} \delta_k B(t, T_k) \\
&= B(t, T_j) \sum_{k=j+1}^{j+c_i} \frac{\delta_k B(t, T_k)}{B(t, T_j)}.
\end{aligned}$$

If we assume $\delta_k = \delta$ for all k and interpret $S_t^{(i)}$ as the average discount rate over the period $[T_j, T_{j+c_1}]$, we get

$$
\begin{aligned}
A_t^{j,j+c_i} &\approx B(t,T_j) \sum_{k=j+1}^{j+c_i} \frac{\delta}{(1+\delta S_t^{(i)})^{k-j}} \\
&= B(t,T_j) \sum_{k=1}^{c_i} \frac{\delta}{(1+\delta S_t^{(i)})^{k}} \ .
\end{aligned}
$$

Hence with the formula for geometric sum and some calculations it follows that

$$
A_t^{j,j+c_i} \approx \frac{B(t,T_j)}{S_t^{(i)}} \Big(1 - \frac{1}{(1+\delta S_t^{(i)})^{c_i}}\Big).
$$

In a similar way we estimate

$$
B(t, T_{j+1}) \approx \frac{B(t,T_j)}{1+\delta S_t^{(i)}}.
$$

To achieve (7.52) it makes now sense to define

$$
H(S_t^{(i)}) := \frac{S_t^{(i)}}{(1+\delta S_t^{(i)})(1 - \frac{1}{(1+\delta S_t^{(i)})^{c_i}})}. \tag{7.53}
$$

Hence

$$
H'(S_t^{(i)}) = \frac{(\delta S_t^{(i)}+1)^{c_i-2}((\delta S_t^{(i)}+1)^{c_i} - c_i \delta S_t^{(i)} - 1)}{((\delta S_t^{(i)}+1)^{c_i} - 1)^2}.
$$

Using (7.53) we are able to evaluate the expected values in (7.50). We write

$$
\begin{aligned}
&\mathbb{E}^{j,j+c_i}\Big[\Big(S_{T_j}^{(i)} - S_0^{(i)}\Big)^+ \Big(\frac{B(T_j,T_{j+1})}{A_{T_j}^{j,j+c_i}} \frac{A_0^{j,j+c_i}}{B(0,T_{j+1})} - 1\Big)\Big] \\
&= \mathbb{E}^{j,j+c_i}\Big[\Big(S_{T_j}^{(i)} - S_0^{(i)}\Big)^+ \Big(\frac{H(S_{T_j}^{(i)})}{H(S_0^{(i)})} - 1\Big)\Big] \qquad (7.54)
\end{aligned}
$$

To evaluate (7.54) we define the smooth function

$$f(x) := (x - S_0^{(i)})\Big(\frac{H(S_{T_j}^{(i)})}{H(S_0^{(i)})} - 1\Big) \tag{7.55}$$

and calculate by using partial integration

$$f'(S_0^{(i)})(S_{T_j}^{(i)} - S_0^{(i)})^+ + \int_{S_0^{(i)}}^{\infty} (S_{T_j}^{(i)} - x)^+ f''(x)dx = \begin{cases} f(S_{T_j}^{(i)}) & \text{, if } S_{T_j}^{(i)} \geq S_0^{(i)} \\ 0 & \text{, if } S_{T_j}^{(i)} < S_0^{(i)} \end{cases}.$$

If we use this on (7.54), we obtain

$$\begin{aligned} &\mathbb{E}^{j,j+c_i}\Big[\Big(S_{T_j}^{(i)} - S_0^{(i)}\Big)^+ \Big(\frac{H(S_{T_j}^{(i)})}{H(S_0^{(i)})} - 1\Big)\Big] = \\ &f'(S_0^{(i)})\mathbb{E}^{j,j+c_i}\Big[(S_{T_j}^{(i)} - S_0^{(i)})^+\Big] + \int_{S_0^{(i)}}^{\infty} \mathbb{E}^{j,j+c_i}\Big[(S_{T_j}^{(i)} - x)^+\Big] f''(x)dx. \end{aligned} \tag{7.56}$$

Now we define

$$C(K) := \mathbb{E}^{j,j+c_i}\Big[(S_t^{(i)} - K)^+\Big] \tag{7.57}$$

and get for (7.50) by using (7.56)

$$\begin{aligned} &\mathbb{E}^{j,j+c_i}\Big[\Big(S_{T_j}^{(i)} - S_0^{(i)}\Big)^+ \Big(\frac{B(T_j, T_{j+1})}{A_{T_j}^{j,j+c_i}} \frac{A_0^{j,j+c_i}}{B(0, T_{j+1})}\Big)\Big] \\ &= \Big(1 + f'(S_0^{(i)})\Big)C(S_0^{(i)}) + \int_{S_0^{(i)}}^{\infty} C(x) f''(x)dx. \end{aligned} \tag{7.58}$$

Therefore (7.58) gives us the expected value of a call under P^j with strike $S_0^{(i)}$ on the underlying $S_t^{(i)}$ by integration of all possible call prices greater then the strike. To see that, just remember

$$\mathbb{E}^{j,j+c_i}\Big[\Big(S_{T_j}^{(i)} - S_0^{(i)}\Big)^+ \Big(\frac{B(T_j, T_{j+1})}{A_{T_j}^{j,j+c_i}} \frac{A_0^{j,j+c_i}}{B(0, T_{j+1})}\Big)\Big] = \mathbb{E}^j\Big[\Big(S_{T_j}^{(i)} - S_0^{(i)}\Big)^+\Big].$$

In the same fashion follows for (7.51)

$$\begin{aligned}
&\mathbb{E}^{j,j+c_i}\Big[\Big(S_0^{(i)} - S_{T_j}^{(i)}\Big)^+\Big(\frac{B(T_j,T_{j+1})}{A_{T_j}^{j,j+c_i}}\frac{A_0^{j,j+c_i}}{B(0,T_{j+1})}\Big)\Big] \\
&= \mathbb{E}^j\Big[\Big(S_0^{(i)} - S_{T_j}^{(i)}\Big)^+\Big] \\
&= \Big(1 + f'(S_0^{(i)})\Big)P(S_0^{(i)}) - \int_{-\infty}^{S_0^{(i)}} P(x)f''(x)dx, \qquad (7.59)
\end{aligned}$$

where

$$P(K) := \mathbb{E}^{j,j+c_i}\Big[(K - S_t^{(i)})^+\Big]. \qquad (7.60)$$

Up to now the function f depends through H on the swap rate $S^{(i)}$ at time T_i. This is highly problematic when it comes to implementation since $S_{T_i}^{(i)}$ is random. To circumvent this issue we use a Tailor-expansion on H at $S_0^{(i)}$ in direction of x. Hence

$$H(x) \approx H(S_0^{(i)}) + H'(S_0^{(i)})(x - S_0^{(i)}) + \dots \qquad (7.61)$$

and therefore

$$f(x) \approx \frac{H'(S_0^{(i)})}{H(S_0^{(i)})}(x - S_0^{(i)})^2. \qquad (7.62)$$

Further we obtain

$$f'(x) \approx 2\frac{H'(S_0^{(i)})}{H(S_0^{(i)})}(x - S_0^{(i)}) \qquad (7.63)$$

and

$$f''(x) \approx 2\frac{H'(S_0^{(i)})}{H(S_0^{(i)})}. \qquad (7.64)$$

According to *Hagan* [32] is the approximation in (7.61) fairly good since H is a smooth function which varies very slowly.

The following Theorem summarizes the results in this chapter.

Theorem 5 (Convexity Correction for CMS spreads)**.**
Let S^{j,c_1,c_2} be a CMS spread with zero strike defined via

$$S_t^{j,c_1,c_2} = S_t^{j,j+c_1} - S_t^{j,j+c_2}.$$

Then the expected value $\mathbb{E}^j\left[S_{T_j}^{j,c_1,c_2}\right]$ under the forward measure $\mathbb{P}^j$ can be approximated as

$$\begin{aligned}
\mathbb{E}^j\left[S_{T_j}^{j,c_1,c_2}\right] &= S_0^{j,c_1,c_2} + \mathbb{E}^j\left[\left(S_{T_j}^{j,j+c_1} - S_0^{j,j+c_1}\right)^+\right] - \mathbb{E}^j\left[\left(S_0^{j,j+c_1} - S_{T_j}^{j,j+c_1}\right)^+\right] \\
&\quad - \mathbb{E}^j\left[\left(S_{T_j}^{j,j+c_2} - S_0^{j,j+c_2}\right)^+\right] + \mathbb{E}^j\left[\left(S_0^{j,j+c_2} - S_{T_j}^{j,j+c_2}\right)^+\right] \\
&= S_0^{j,c_1,c_2} + \left(1 + f'(S_0^{j,j+c_1})\right) C(S_0^{j,j+c_1}) + \int_{S_0^{j,j+c_1}}^{\infty} C(x) f''(x) dx \\
&\quad - \left(1 + f'(S_0^{j,j+c_1})\right) P(S_0^{j,j+c_1}) + \int_{-\infty}^{S_0^{j,j+c_1}} P(x) f''(x) dx \\
&\quad - \left(1 + f'(S_0^{j,j+c_2})\right) C(S_0^{j,j+c_2}) - \int_{S_0^{j,j+c_2}}^{\infty} C(x) f''(x) dx \\
&\quad + \left(1 + f'(S_0^{j,j+c_2})\right) P(S_0^{j,j+c_2}) - \int_{-\infty}^{S_0^{j,j+c_2}} P(x) f''(x) dx,
\end{aligned}$$

where f is defined in (7.55).

Proof. The general structure of the approximation is given at (7.45) and the call/put splits are shown in (7.48) and (7.49). The approximation through the integration over expected call and put values for different strikes is given in (7.58) and (7.59), where we approximate f through (7.62), (7.63) and (7.64). □

The convexity correction can be seen as a measure for the skew difference of $\mathbb{P}^j$ and $\mathbb{P}^{\text{CMSs}}$. The skew of a probability measure describes where it has most of its mass. Under martingale measure $\mathbb{P}^{\text{CMSs}}$ belonging to S^{j,c_1,c_2} it holds

$$\begin{aligned}
&\mathbb{E}^{\text{CMSs}}\left[(S_{T_j}^{j,j+c_1} - S_0^{j,j+c_1})^+\right] - \mathbb{E}^{\text{CMSs}}\left[(S_{T_j}^{j,j+c_1} - S_0^{j,j+c_1})^-\right] \\
=&\mathbb{E}^{\text{CMSs}}\left[S_{T_j}^{j,j+c_1} - S_0^{j,j+c_1}\right] = 0.
\end{aligned} \tag{7.65}$$

If $\mathbb{P}^j$ and $\mathbb{P}^{CMSs}$ have the same skew and shape we would obtain $\mathbb{E}^j\left[S_{T_j}^{j,c_1,c_2}\right] = S_0^{j,c_1,c_2}$, since the integrals in theorem 5 would cancel each other out as well. However in most cases the size of the expected values under the two measures is nonzero. This means that $\mathbb{P}^j$ values undiscounted calls or put different from $\mathbb{P}^{\text{CMSs}}$. So under $\mathbb{P}^j$ high payoffs are more likely or less likely then under $\mathbb{P}^{\text{CMSs}}$. Therefore the measures differ in their skew, which is expressed by the value of the correction.

For completeness we mention that is possible to calculate the drift correction the intuitive way given at (7.44), but it is rather difficult and unknown if the accuracy can be improved significantly. However the interested reader can find more about this approach in [5].

7.3. Calibration to CMS Spread Options

With Theorem 4 from chapter 7.1 and Theorem 5 from chapter 7.2 we have everything together to calibrate our supercorrelation matrix P – which contains all the correlations of the processes in the SABR-LMM model – on CMS spread options. Theorem 4 gives us a proper dynamic for the spreads in a SABR-like environment, whereas Theorem 5 enables us to evaluate under the forward measures $\mathbb{P}^j$.
We will use the formula for implied volatility (3.4) to calculate the integrals in Theorem 5. To use the formula or some other workaround is necessary for two reasons. First, only call and put prices for strikes K in a certain range $[K_{\min}, K_{\max}]$ are quoted in the market. Second, there is only a finite amount of strikes, whereas we need a continuum.

In the model calibration based on CMS spreads the very first step is the same as in the model calibration based on Swaps. The volatility functions g_j and h_j are calibrated to caps, just as the coefficients k_0^j and ζ_j. We refer to chapter 6.2.

However, to calibrate the correlation we first focus on how to solve the integrals for the convexity correction in Theorem 5 and how to model the dynamics of the involved swaps. We assume that under the swap measure $\mathbb{P}^{j,j+c_i}$ the swap rate $S^{j,j+c_i}$ evolves like a SABR process. Therefore, the dynamics are

$$\begin{aligned} dS_t^{j,j+c_i} &= \sigma_t^{j,j+c_i}\left(S^{j,j+c_i}\right)^{\beta_{j,j+c_i}} dW_t^{j,j+c_i},\ S_0^{j,j+c_i} = S^{j,j+c_i}(0) \\ d\sigma_t^{j,j+c_i} &= \sigma_t^{j,j+c_i}\nu^{j,j+c_i} dZ_t^{j,j+c_i},\ \sigma_0^{j,j+c_i} = \sigma^{j,j+c_i}(0) \\ d\langle W_\bullet^{j,j+c_i}, Z_\bullet^{j,j+c_i}\rangle_t &= \rho^{j,j+c_i} dt, \end{aligned} \tag{7.66}$$

where as usually $\beta_{j,j+c_i} \in [0,1]$, $\nu^{j,j+c_i}, \sigma^{j,j+c_i}(0) > 0$ and $\rho^{j,j+c_i} \in [-1,1]$. From this we get the implied Black volatility (3.4) basing on the model parameters

$$\sigma_{\mathrm{B}}^{j,j+c_i}(K) := \sigma_{\mathrm{I}}(S^{j,j+c_i}(0), K, \beta_{j,j+c_i}, \sigma^{j,j+c_i}(0), \nu^{j,j+c_i}, \rho^{j,j+c_i}, T_j), \tag{7.67}$$

which gives us the right volatility to price calls and puts with strike K on the underlying $S^{j,j+c_i}$. With that we can write for the expected values in (7.57) and (7.60)

$$C(K) = S^{j,j+c_i}(0)\mathcal{N}(d_1) - K\mathcal{N}(d_2) \tag{7.68}$$

and

$$P(K) = K\mathcal{N}(-d_2) - S^{j,j+c_i}(0)\mathcal{N}(-d_1), \tag{7.69}$$

where

$$d_1 = \frac{\ln\left(\frac{S^{j,j+c_i}(0)}{K}\right) + \frac{1}{2}(\sigma_{\mathrm{B}}^{j,j+c_i}(K))^2 T_j}{\sigma_{\mathrm{B}}^{j,j+c_i}(K)\sqrt{T_j}}$$

and

$$d_2 = d_1 - \sigma_{\mathrm{B}}^{j,j+c_i}(K)\sqrt{T_j}.$$

As suggested and tested in [13] we can use (7.68) and (7.69) together with (7.67) to evaluate the integrals in Theorem 5 from $S^{j,j+c_i}(0)$ to ∞ and $-\infty$ to $S^{j,j+c_i}(0)$, respectively.
To incorporate the market data and to obtain the model parameters of the swap dynamics in (7.66) we solve an ordinary, unweighted least-square problem (3.5)

$$\min_{\sigma^{j,j+c_i}(0), \nu^{j,j+c_i}, \rho^{j,j+c_i}} \sqrt{\sum_K \left(\sigma_{\mathrm{B}}^{j,j+c_i}(K) - \sigma_{\mathrm{M}}(K)\right)^2}, \tag{7.70}$$

where $\sigma_{\mathrm{M}}(K)$ is the implied volatility quoted in the market for the strike K of a call or put. As in chapter 6, which treated the calibration of the SABR dynamics to swaps, we set

$$\beta_{j,j+c_i} = \sum_{k=j}^{j+c_i-1} \omega_k^{j,j+c_i} \beta_k, \tag{7.71}$$

where β_k are the betas in the SABR dynamics of the forward rates F^k. The techniques above enable us to calculate the convexity corrections properly.

The CMS spread dynamic for the spread S^{j,c_1,c_2} is modeled equivalent to Theorem 4 as

$$\begin{aligned} d\Big(S_t^{j,c_1,c_2} + B^{i,c_1,c_2}\Big) &= u_t(S_t^{j,c_1,c_2} + B^{i,c_1,c_2})dW_t^{i,c_1,c_2} \\ du_t &= u_t A^{p,\nu}_{j,c_1,c_2}(\rho, r)dZ_t^{i,c_1,c_2},\ u_0 = q^{j,c_1,c_2} \\ d\langle W_\bullet^{i,c_1,c_2}, Z_\bullet^{i,c_1,c_2}\rangle_t &= \mathcal{X}^{j,c_1,c_2}(\rho, r, R)dt, \end{aligned} \tag{7.72}$$

where

$$B^{j,c_1,c_2} := \frac{p^{j,c_1,c_2}}{q^{j,c_1,c_2}} - S^{j,c_1,c_2}(0)$$

and $p^{j,c_1,c_2}, q^{j,c_1,c_2}$ are given in the proof of the Theorem at (7.42) and (7.43). The parameterization of $A^{p,\nu}_{j,c_1,c_2}(\rho, r)$ is given in (7.40) whereas the one of $\mathcal{X}^{j,c_1,c_2}(\rho, r, R)$ is given in (7.41). To see the equivalence of both formulation for the spread dynamics in definition (4) and this one here just notice

$$d\Big(S_t^{j,c_1,c_2} + B^{i,c_1,c_2}\Big) = dS^{j,c_1,c_2},$$

since B^{i,c_1,c_2} is constant and that we just multiplied u_t by q^{i,c_1,c_2}.

Before we can start with the calibration of the correlation to CMS spreads we have to consider the way how the market quotes options on spreads. Different to swaptions, which are quoted in implied Black volatilities, options on CMS spreads are quoted in implied *normal* volatilities. This is because CMS spreads can get negative as differences of two swap rates. If the price is given in normal volatility σ^{n}, it is assumed that the underlying S follows a *Bachelier* model. That is

$$dS = \sigma^{\mathrm{n}} dW_t, \quad S_0 = S(0)$$

for $\sigma^{\mathrm{n}} > 0$ and a Wiener process $(W_t)_t$. Therefore a call on a underlying S with strike K, expiry $T_{\exp}$ can be evaluated as [42]

$$\mathbb{E}\big[(S-K)^+\big] = (S_0 - K)\mathcal{N}\Big(\frac{S_0 - K}{\sigma^{\mathrm{n}}\sqrt{T_{\exp}}}\Big) + \sigma^{\mathrm{n}}\sqrt{T_{\exp}}\varphi\Big(\frac{S_0 - K}{\sigma^{\mathrm{n}}\sqrt{T_{\exp}}}\Big),$$

where φ is the density function of the normal distribution. In the same fashion as for implied Black volatility there exists a analytic formula for the implied normal volatility σ^{n} of the SABR model (3.1). Again, the formula depends only on the current underlying price, strike of the call or put and the model parameters. In [17] *Hagan* approximates the implied normal volatility as follows

$$\begin{aligned}\sigma^{\mathrm{n}}_{\mathrm{I}}(F,K,\beta,\alpha,\nu,\rho,T_{\exp}) := {} & \alpha\big(FK\big)^{\beta/2} \\ & \times \frac{1+\frac{1}{24}\log^2\left(\frac{F}{K}\right)+\frac{1}{1920}\log^4\left(\frac{F}{K}\right)}{1+\frac{(1-\beta)^2}{24}\log^2\left(\frac{F}{K}\right)+\frac{(1-\beta)^2}{1920}\log^4\left(\frac{F}{K}\right)}\Big(\frac{\zeta}{\mathrm{X}(\zeta)}\Big) \\ & \times \Big[1+\Big(\frac{-\beta(2-\beta)\alpha^2}{24(FK)^{1-\beta}}+\frac{\rho\alpha\nu\beta}{4(FK)^{1-\beta}/2} \\ & + \frac{2-3\rho^2}{24}\nu^2\Big)T_{\exp}\Big],\end{aligned} \tag{7.73}$$

where

$$\zeta := \frac{\nu}{\alpha}(FK)^{(1-\beta)/2}\log(F/K)$$

and

$$\mathrm{X}(\zeta) := \log\Big(\frac{\sqrt{1-2\rho\zeta+\zeta^2}-\rho+\zeta}{1-\rho}\Big).$$

As in the case of the formula for the implied Black volatility the formula in (7.73) is purely analytic and highly tractable in regards of implementation.

Now, in order to fit the correlations under the forward measure $\mathbb{P}^j$ basing on European options on CMS we define

$$\begin{aligned}\sigma^j_{\mathrm{n}}(K) := \sigma^{\mathrm{n}}_{\mathrm{I}}\Big(& B(0,T_{j+1})\mathbb{E}^j[S^{j,c_1,c_2}_{T_j}] + B^{i,c_1,c_2}, K + B^{i,c_1,c_2}, \\ & 1, q^{j,c_1,c_2}, A^{p,\nu}_{j,c_1,c_2}(\rho,r), \mathcal{X}^{j,c_1,c_2}(\rho,r,R), T_j\Big).\end{aligned} \tag{7.74}$$

This makes sense since the price of a call option with expiry T_i and settlement date T_j on a CMS S^{i,c_1,c_2} is

$$\begin{aligned}
&B(0,T_j)\mathbb{E}^i\Big[(S^{i,c_1,c_2}_{T_i}-K)^+\Big] \\
&= B(0,T_j)\mathbb{E}^i\Big[((S^{i,c_1,c_2}_{T_i}+B^{i,c_1,c_2})-(K+B^{i,c_1,c_2}))^+\Big].
\end{aligned}$$

In (7.74) we calculate $\mathbb{E}^j[S^{j,c_1,c_2}_{T_j}]$ with the convexity correction as shown above. Now, we take calls and puts on CMS spreads for different expiry dates and strikes into account to calibrate to market data. Our goal is to calibrate the CMS spread dynamics (7.72) as close as possible to all market prices. We do this by an ordinary least-square problem and using (7.74)

$$\min_{q^{j,c_1,c_2},A^{p,\nu}_{j,c_1,c_2}(\rho,r),\mathcal{X}^{j,c_1,c_2}(\rho,r,R)} \sum_j \sqrt{\sum_{K_j}\Big(\sigma^j_{\mathrm{n}}(K_j)-\sigma^j_{\mathrm{M}}(K_j)\Big)^2}, \tag{7.75}$$

where we sum over all expiry dates T_j and available strikes K_j for this date. Here $\sigma^j_{\mathrm{M}}(K_j)$ is the implied normal volatility observed in the market for a call or put on S^{j,c_1,c_2} with underlying price S^{j,c_1,c_2}_0 and strike K_j. Since all the variables in the minimization problem depend on P through ρ, r or R we achieve our goal to calibrate the supercorrelation matrix.

Remark. In the same fashion as above it is possible to calibrate the CMS spread dynamics on straddles on CMS spreads. A straddle is the sum of a call and put option with the same strike. A straddle with strike K on S^{i,c_1,c_2} has in T_{i+1} the payoff

$$|S^{i,c_1,c_2}_{T_i}-K|. \tag{7.76}$$

Now let's denote by $C(S^{i,c_1,c_2}_0,K)$ the expected value of a call and $P(S^{i,c_1,c_2}_0,K)$ the expected value of a put with underlying price S^{i,c_1,c_2}_0 and strike K. From the definition we know that the expected value of a straddle can be calculated as

$$\begin{aligned}
\mathbb{E}\big[\text{Straddle}\big] &:= C(S^{i,c_1,c_2}_0,K)+P(S^{i,c_1,c_2}_0,K) \\
&= 2C(S^{i,c_1,c_2}_0,K)-S^{i,c_1,c_2}_0+K
\end{aligned} \tag{7.77}$$

where we used the call put parity. By assuming that S^{i,c_1,c_2} follows a *Bachlier* model with volatility $\sigma^{\mathrm{n}}>0$, the representation in (7.77) leads to the following expected value

$$\mathbb{E}\big[\text{Straddle}\big] = 2(S_0^{i,c_1,c_2} - K)\mathcal{N}\Big(\frac{S_0^{i,c_1,c_2} - K}{\sigma^{\mathrm{n}}\sqrt{T_i}}\Big) + 2\sigma^{\mathrm{n}}\sqrt{T_i}\varphi\Big(\frac{S_0^{i,c_1,c_2} - K}{\sigma^{\mathrm{n}}\sqrt{T_i}}\Big) - S_0^{i,c_1,c_2} + K. \tag{7.78}$$

Now, CMS spread straddles are quoted in implied normal volatility consistent with formula (7.78) and the implied normal volatility belongs to a call/put on S^{i,c_1,c_2} with strike K and exipiry T_i. This enables us indeed to calibrate on CMS straddle prices in the same way as we calibrate to calls and puts to CMS spreads.

8. Implementation and Empirical Study

8.1. Implementation

The implementation of the calibration on swaps and CMS spreads was carried out by the author himself, for which the programming language F# was used. F# is a free, functional .NET language, which is supported by a great number of routines provided by packages like Math.Net and Microsoft's Solver Foundation. All minimization problems from chapter 6 and chapter 7.3 were solved by the Nelder-Mead algorithm from the Solver Foundation package and for numerical integrations as in the case of generating random numbers the Numerics package from Math.Net was used.

8.2. Empirical Study

The calibration was tested for the Euribor as underlying yield curve. We used data from two different dates and carried out the calibration for a SABR-LMM which covers 20 years. We set $\delta_i \equiv 0.5$ and used the tenor structure $(T_i)_{i\in\{0,\dots,40\}}$. Hence, we calibrated the dynamic for half year forward rates whose first expiring date is in 0.5 years and the last in 19.5 years. Afterwards we tested the quality of the calibration by doing Monte Carlo simulations. The simulation was implemented by *Cresnik* [28], where discretization of the model SDE's was done by using the Milstein scheme.

8.2.1. The Data

The calibration of the SABR-LMM was tested for the calibration on swaptions and CMS spread options. In both cases we compared the (2SC) and (5L) parametrization for ρ. Further, for the swaptions we compared two dates. The first data set is from the 04.06.2006 and represents quiet market situation in an environment of low volatility and high interest rates. The second data set is from the 21.07.2014 and represents a rough market with high uncertainty and low interest rates. The calibration on CMS spreads was carried out on data as of 21.07.2014. For the calibration we followed the methods described in chapter 6 and chapter 7.3. The swaption prices are from Bloomberg. We thank Peter Krierer for providing the CMS spread option prices from BGC Market Data.

8.2.2. Calibration to Swaption Prices as of 04.06.2006

For this date we used a cap cube with 18 strikes ranging from ATM over 1% to up to 14% and maturities given in full years starting from 1 year and ranging up to 20 years. We interpolated the missing half-year maturities by using spline interpolation for the volatilities and strikes.
The given swaption cube for coterminal swaps consists of 22 strikes, ranging from ATM-2% to ATM+3%, and expiry dates in $\frac{1}{2}$, 1, 2,. . . , 10, 12 and 15 years. However not all tenors for coterminal swaps were available. For example, there are no options on 9x11 swaptions. We got the missing tenors by spline interpolation for strikes and volatilities.
After setting up the data we followed chapter 6 to calibrate first the volatility structure to caplet volatilities and second the correlations, based on the swaption prices. The parameters of the calibration and plots of the supercorrelation matrices can be found in the appendix.

To describe the distinction between the supercorrelation matrices P by using the two different parameterizations for ρ, lets focus first on the difference in the ρ's itself. The general shape of ρ using the different parameterizations almost the same. Nevertheless, there are differences, which are visualized in figure 8.1. First, the *wings* of the matrices differ. The (5L) matrix goes in areas for the correlation of rates more then 20 years apart down to 50% and therefore around 10% lower then the (2SC) parametrized matrix. Second, the correlation for rates which are less than 20 years apart is a little bit higher.
However, the other submatrices of P under the different approaches for ρ cannot be distinguished, since the parameters are almost the same. That is a bit surprising since the differences in the outer areas of ρ are pronounced.

To analyze the quality of our calibration we repriced the target products, which were used in the calibration process, and calculated the relative pricing errors. First, we used the *gluing* algorithm from chapter 5.3 to estimate the nearest correlation matrix $\widehat{P}$ of rang $k = 10$. Then, in each simulation we did 300.000 runs with 1000 Steps for the whole time interval $[T_0, T_{39}]$. Plots of the relative errors can be found in the appendix.

We repriced all caplets on all strikes. It did not matter, if we used the supercorrelation matrix basing on (5L) or (2SC). In both cases, the picture was the same. Overall the results are satisfying and getting more accurate with higher expiry, as can be seen on the mean errors in table 8.1. For caplets with any expiry we observed, that the relative error for strikes greater then 6% explodes. In this area the errors begin to sky rock starting at 20% and went up to 50%. Therefore, we excluded strikes greater then 6% from table 8.1.
One way to explain the errors is as follows: We tried to preserve the simple SABR model for forward rates in the SABR-LMM as well as possible. Further, caplets are almost not effected by the correlations of forward rates given through the submatrices ρ and r of P, but they rely on the skews given through R. The diagonal of R plays the role of the ρ's from the simple SABR model, since it holds $R_{ii} = \rho_i$. If ρ_i in the simple SABR is slightly disturbed the implied model volatility for this forward rate is disturbed. As figure 3.1 shows the relative error gets bigger with greater strikes and the induced relative pricing error through Black's formula is even higher in this area. Now, to be able to do the Monte Carlo simulation we need to approximation model supercorrelation matrix P by a real correlation matrix $\widehat{P}$. In general P and $\widehat{P}$ will deviate by a small amount and therefore the diagonals of R will slightly differ as well. This leads to differences of the individual skews for each forward rate in the simple SABR compared with the SABR-LMM. The approximation of P was introduced in chapter 5.3.
A similar explanation holds for caplets with small expiry and strikes away from the money. The market prices of those options are in the area from 0.001bp to 10bp and therefore really low. If the implied model volatility is disturbed slightly, the effect on the prices is tremendous and explains the relative errors in a area of 100%.

To verify the correlation structure we repriced all used coterminal swaptions. We compared the results we get by using the two different parameterizations (5L) and (2SC) for ρ. As in the case for caplets the results are overall satisfying. The absolute relative mean-pricing errors in table 8.2 show, that the (5L) parametrization works slightly better then the (2SC) parametrization. However, in both cases we face huge

Table 8.1.: Absolute Relative Pricing Errors for Caplets with Strikes in $\{1\%, \dots, 6\%\}$ as of 04.06.2006.

i	max	min	mean	i	max	min	mean
1	1.0058	0.0000	0.1484	21	0.0242	0.0066	0.0139
2	1.1960	0.0142	0.2023	22	0.0174	0.0022	0.0105
3	0.1196	0.0208	0.0600	23	0.0140	0.0046	0.0091
4	0.1860	0.0046	0.0654	24	0.0416	0.0000	0.0127
5	0.3150	0.0273	0.0912	25	0.0359	0.0052	0.0129
6	0.2507	0.0256	0.0798	26	0.0292	0.0073	0.0139
7	0.3381	0.0253	0.0940	27	0.0260	0.0069	0.0136
8	0.0542	0.0004	0.0353	28	0.0243	0.0058	0.0128
9	0.1369	0.0137	0.0422	29	0.0278	0.0039	0.0109
10	0.1667	0.0114	0.0470	30	0.0920	0.0046	0.0183
11	0.2541	0.0076	0.0595	31	0.0792	0.0049	0.0165
12	0.0914	0.0003	0.0194	32	0.0768	0.0038	0.0155
13	0.0722	0.0039	0.0204	33	0.0736	0.0029	0.0142
14	0.0912	0.0071	0.0258	34	0.0712	0.0014	0.0130
15	0.0921	0.0090	0.0271	35	0.0714	0.0003	0.0121
16	0.0693	0.0108	0.0278	36	0.0700	0.0001	0.0116
17	0.0684	0.0114	0.0287	37	0.0647	0.0000	0.0106
18	0.0522	0.0112	0.0262	38	0.0652	0.0004	0.0112
19	0.0429	0.0095	0.0218	39	0.0573	0.0001	0.0100
20	0.0290	0.0010	0.0148				

pricing errors for swaptions with small expiry and strikes far out-the-money – this visible in the plots in the appendix. In the end of chapter 4.2 we mentioned that *Rebonato* found that the used approximation for the swap dynamics breaks down in exactly those cases. The SABR-LMM does not reproduce the implied smile for the swap dynamics correctly which leads to tremendous pricing errors for high strikes, if the real prices and expiries of the options is low.

Table 8.2.: Absolute Relative Pricing Errors for Coterminal Swaptions with Strikes in $\{\text{ATM} - 2\%, \ldots, \text{ATM} + 3\%\}$ as of 04.06.2006.

P	basing	on	(5L)	P	basing	on	(2SC)
i	max	min	mean	i	max	min	mean
1	1.0000	0.0013	0.4038	1	1.0000	0.0011	0.3987
2	0.9967	0.0018	0.3149	2	0.9557	0.0027	0.3178
4	0.6592	0.0035	0.1866	4	0.7009	0.0035	0.1885
6	0.2906	0.0027	0.0816	6	0.3499	0.0028	0.0891
8	0.0352	0.0002	0.0107	8	0.0628	0.0001	0.0116
10	0.0421	0.0017	0.0200	10	0.0490	0.0022	0.0237
12	0.0744	0.0053	0.0415	12	0.0821	0.0070	0.0469
14	0.1295	0.0090	0.0641	14	0.1329	0.0111	0.0696
16	0.1023	0.0112	0.0605	16	0.1033	0.0130	0.0639
18	0.0982	0.0134	0.0625	18	0.0948	0.0142	0.0618
20	0.1016	0.0151	0.0649	20	0.0946	0.0155	0.0615
24	0.0672	0.0114	0.0493	24	0.0590	0.0119	0.0445
30	0.0834	0.0036	0.0350	30	0.0897	0.0058	0.0393

8.2.3. Calibration to Market Prices as of 21.07.2014

Calibration of the Volatility Structure

To calibrate the volatility as explained in chapter 6.2 we again used a cap cube with 18 strikes ranging from ATM over 1% to up to 14% and maturities given in full years starting from 1 year and ranging up to 20 years. We interpolated the missing half year maturities by using spline interpolation for the volatilities and strikes, as well. The parameters for the volatility functions can be found in the appendix.

Calibration of the Correlation Structure to Swaption Prices

For the calibration of the correlations we used a swaption cube consisting of 13 strikes, ranging from ATM-0.3% to ATM+2.5% and swaptions with expiries in $\frac{1}{2}$, 1, 2,..., 10, 12 and 15 years. Again, not all tenors for coterminal swaps were available and calculated the missing tenors by spline interpolation for strikes and volatilities. The parameters for the correlation matrices and the plots can be found in the appendix.

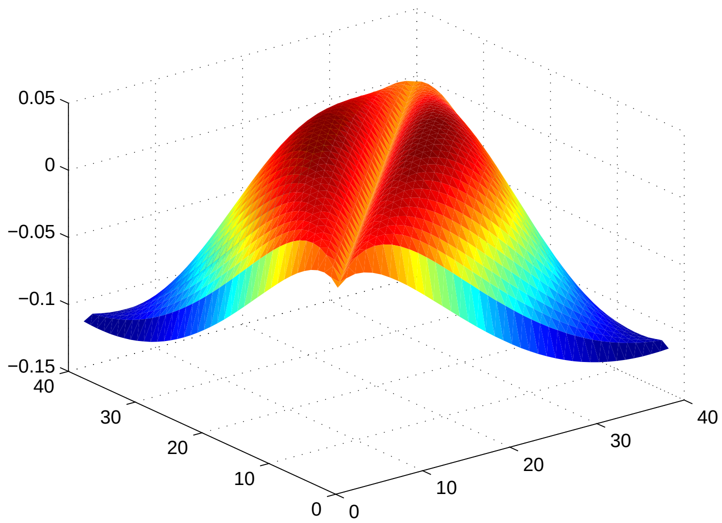

Figure 8.1.: Here we see the difference of Lutz' (5L) parameterization and Schoenmakers & Coffey's (2SC) parametrization for 04.06.2006. The (5L) parametrized matrix is lower at the wings and higher near the main diagonal. Notice, the curved difference surface shows the higher flexibility of the (5L) parametrization as well. The plot was done in Matlab.

In contrary to the calibration on the data from year 2006 the shapes of the correlation matrix ρ by a great amount depending on the parametrization. The (5L) matrix goes down to zero in the front and yields a tiny correlation only for forward rates with small expiries to any other forward rate. For these correlation the situation is differently under the (2SC) parametrization. Here the correlation of the short-term rates is bounded downwards by roughly 30%. This phenomena can be explained by the higher inflexibility of the parametrization. For the (2SC) parametrization holds: If the correlation in the front goes down the correlation in the back is drawn down as well. Therefore, in the calibration process some trade off between high correlations in the back and low in the front has to be found and the boundary is the result. In this case the impact is even visible in the submatrix r. In the case of the (5L) parametrization the minimum of r is 90%, whereas in case of the (2SC) parametrization r is almost constant one. But, R is in both cases the same.

Numerical Results for Caplets

To verify the calibrations we repriced again the target options via Monte Carlo Simulations. Here we doubled the number of runs to 600.000, because the simulations converged slower. We sticked to 1000 steps for the whole time interval $[T_0, T_{39}]$. Here we choose the parameterization for ρ and calibration approach whose resulting model supercorrelation matrix P with the lowest distance to the approximated supercorrelation matrix $\widehat{P}$ from chapter 5.3. Thus, we used the correlation matrix based on (5L) and calibrated to swaption prices.

We simulated caplet prices with strikes in a range of $\{1\%, \ldots, 14\%\}$. Differently from 2006, the relative error for strikes over 6% went up to *only* 20%. But again, a gradual increase of the error for those strikes was observable and the explanation is the same as for the previous data. For this reason, we only consider strikes of maximal 6% in table 8.3. The pricing errors for short expiries are more pronounced than in 2006. We believe this is a numerical issue resulting from the higher overall volatility and much lower initial forward rates F^0, which are below one percent. The problematic is further amplified by the missing drift or reflection barrier in the SABR-LMM. In 2006 the initial rates were above two percent whereas in 2014 there were by around 0.3%. A plot of the error surface can be found in the appendix.

Numerical Results for Swaptions when Calibrated to Swaps

The simulation of swaption prices shows that the parametrization of ρ really matters. In the case of a (5L) parametrization of ρ the relative errors are only a little bit higher than 2006 as can be seen in table 8.4 and is on a satisfying level if we consider the increase of pricing errors of the caplets. However, in the case of a (2SC) parametrization for ρ the relative error sky rocks. One explanation for this phenomena could be the higher correlations between forward rates with longer expiry, when we use the (5L) parametrization. Further, in the case of a (2SC) parametrization for ρ the distance of the approximating matrix $\widehat{P}$ for P and P was almost twice as high as for the (5L) parametrization for ρ. The approximation of P was explained in chapter 5.3.

Table 8.3.: Absolute relative Pricing Errors for Caplets with Strikes in $\{1\%, \ldots, 6\%\}$ as of 21.07.2014.

i	max	min	mean	i	max	min	mean
1	0.0000	0.0000	0.0000	21	0.0904	0.0349	0.0574
2	0.9578	0.0000	0.1709	22	0.0574	0.0315	0.0447
3	0.5413	0.0000	0.2160	23	0.0597	0.0297	0.0443
4	0.6608	0.1574	0.4379	24	0.0493	0.0118	0.0342
5	0.7654	0.1925	0.5397	25	0.0539	0.0063	0.0361
6	0.6191	0.0777	0.3832	26	0.0613	0.0303	0.0502
7	0.6957	0.0184	0.3385	27	0.0663	0.0516	0.0600
8	0.3275	0.0162	0.1520	28	0.0678	0.0523	0.0635
9	0.3764	0.0223	0.1717	29	0.0670	0.0476	0.0599
10	0.3583	0.0068	0.1464	30	0.1229	0.0064	0.0558
11	0.3688	0.0010	0.1311	31	0.0906	0.0073	0.0497
12	0.1537	0.0065	0.0571	32	0.0650	0.0064	0.0452
13	0.1098	0.0117	0.0434	33	0.0635	0.0133	0.0424
14	0.1173	0.0121	0.0470	34	0.0617	0.0031	0.0403
15	0.1316	0.0169	0.0551	35	0.0640	0.0108	0.0465
16	0.0629	0.0210	0.0340	36	0.0673	0.0388	0.0572
17	0.0516	0.0219	0.0320	37	0.0694	0.0581	0.0640
18	0.0541	0.0243	0.0359	38	0.0666	0.0582	0.0645
19	0.1026	0.0291	0.0553	39	0.0670	0.0557	0.0630
20	0.0581	0.0335	0.0459				

Calibration of the Correlation Structure to CMS Spread Option Prices

The calibration to CMS spread option was done by using European call and put prices in normal volatility on 10 year vs. 2 year (10y/2y) CMS spreads with strikes ranging from ATM-0.25% to ATM+1.5% and expiries of 1, 2, 3, 4, 5, 7 and 10 years. We interpolated the data linear in volatilities and strikes to get half year expiries starting at 1/2 years. To calibrate the two SABR models for swaps to calculate the convexity correction derived in chapter 7.2, we used quoted market data for swaptions. Those swaptions had the standard expiries $\frac{1}{2}$, 1, 2,. . . , 10, 12 and 15 years and were available for 13 strikes, ranging from ATM-0.3% to ATM+2.5% and tenors of 2 years and 10 years, respectively. We interpolated the market data by spline interpolation to get the volatilities and strikes for options with half yearly-expiry dates. For the calibration procedure we then followed chapter 7.3.

Table 8.4.: Absolute Relative Pricing Errors for Coterminal Swaptions with Strikes in $\{\text{ATM} - 0.3\%, \ldots, \text{ATM} + 2.5\%\}$ as of 21.07.2014., if the correlation structure is calibrated to swaps.

P	basing	on	(5L)	P	basing	on	(2SC)
i	max	min	mean	i	max	min	mean
1	0.9999	0.0253	0.3889	1	1.0000	0.0041	0.4640
2	0.2908	0.0162	0.1488	2	0.6093	0.0102	0.1467
4	0.0731	0.0027	0.0209	4	0.4179	0.0193	0.1654
6	0.5651	0.0028	0.1693	6	0.6877	0.0258	0.2606
8	0.5217	0.0015	0.1669	8	0.6289	0.0296	0.2435
10	0.0347	0.0004	0.0163	10	0.1865	0.0605	0.1034
12	0.0395	0.0154	0.0262	12	0.1651	0.0565	0.0929
14	0.0429	0.0194	0.0305	14	0.1564	0.0572	0.0909
16	0.0643	0.0569	0.0587	16	0.1902	0.0697	0.1150
18	0.0824	0.0655	0.0749	18	0.1991	0.0774	0.1269
20	0.0870	0.0730	0.0820	20	0.1976	0.0836	0.1327
24	0.1042	0.0744	0.0877	24	0.2240	0.0856	0.1405
30	0.1106	0.0789	0.0913	30	0.2479	0.0914	0.1514

Numerical Results for Swaptions when Calibrated to CMS spreads

As in the case for the calibration to swaption prices we calculated swaption prices via Monte Carlo simulations using 600000 runs and 1000 steps. For both possible parameterizations for ρ we got errors similar to the case of the calibration to swaptions, if we choose the (2SC) parametrization for ρ. This is a bit surprising since in the case of a (5L) parametrization for ρ the calibrated supercorrelation matrix has the same characteristics like the supercorrelation matrix obtained by calibrating to swaptions and using the same parametrization. However, as in the case of the calibration to swaptions the (5L) parametrization works much better then the (2SC) parametrization. A plot of the relative pricing errors can be found in the appendix.

One reason for the high errors may be the unusual high distance of the estimated correlation matrix $\widehat{P}$, which is used for the simulations, and the model correlation matrix P. Another reason could be that, different to coterminal swaps, CMS spread options alone do not induce a condition on each entry of P. In chapter 8.2.4 we discuss this issue thoroughly.

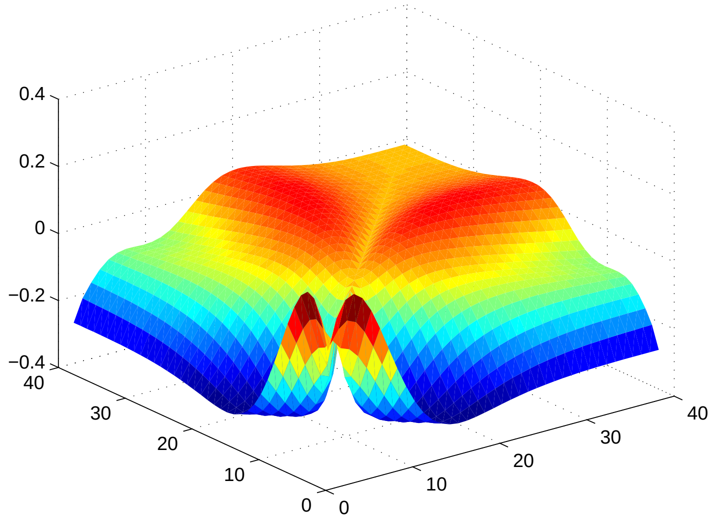

Figure 8.2.: This plot visualizes the difference between Lutz' and Schoenmakers & Coffey's (2SC) parametrization, if we calibrate on swaption prices. The difference shows: The (5L) matrix dictates a higher correlation of long-term forward rates and a much stronger decorrelation of short-term forwards. The plot was done in Matlab.

Table 8.5.: Absolute Relative Pricing Errors for Coterminal Swaptions with Strikes in $\{\text{ATM} - 0.3\%, \ldots, \text{ATM} + 2.5\%\}$ as of 21.07.2014, if P is calibrated to CMS spread options.

P	basing	on	(5L)	P	basing	on	(2SC)
i	max	min	mean	i	max	min	mean
1	1.0000	0.0066	0.4602	1	1.0000	0.0798	0.7178
2	0.6055	0.0003	0.1394	2	1.0000	0.1523	0.7012
4	0.3460	0.0082	0.1231	4	0.9818	0.2108	0.6639
6	0.6328	0.0062	0.2192	6	0.9750	0.2281	0.6339
8	0.5775	0.0076	0.2032	8	0.9423	0.2419	0.5992
10	0.0952	0.0400	0.0557	10	0.8000	0.2633	0.5171
12	0.0989	0.0410	0.0623	12	0.7339	0.2500	0.4684
14	0.0849	0.0447	0.0577	14	0.6488	0.2281	0.4112
16	0.1513	0.0592	0.0975	16	0.5830	0.2167	0.3742
18	0.1783	0.0674	0.1119	18	0.5045	0.1953	0.3276
20	0.1870	0.0776	0.1242	20	0.4272	0.1739	0.2820
24	0.2353	0.0850	0.1454	24	0.3431	0.1397	0.2217
30	0.2645	0.0939	0.1599	30	0.2802	0.1035	0.1700

8.2.4. Critique on Calibration Solely to CMS Spread Options

CMS spreads depend stronger on correlation between the underlying forward rates than swaps [29], [22] but nevertheless we do not recommend to calibrate to spreads alone. If we consider the general structure of a 10y/2y spread and $\delta_i \equiv 0.5$, which was given in (7.2) as

$$S_t^{i,20,4} = \sum_{l=i}^{i+19} v_l^{i,20,4}(t) F_t^l,$$

it is obvious that in each spread only 20 forward rates are involved. Now we assume that, we have for each tenor point T_i the prices for a range of European options on CMS spreads. Then, for each i it is only possible to estimate 20 of the $39 - i + 1$ possible correlations of the forward rates expiring in i or later. Therefore, only the first 20 left and right main-diagonals of the correlation matrices ρ and r are directly affected by the spread dynamics. This means during the calibration on spreads no conditions for the *wings* are given, that is a quarter of both matrices. The behavior of the outer areas of the matrix have to be *extrapolated* from the middle part. This can lead to overall lower correlations for forward rates which are further apart. This is because, the correlation for forward rates with expiries close to each other goes down rapidly in the beginning and induces, if we do not know any terminal minimal correlation, too low correlations for the out part of the matrix.

For calibration to a set of coterminal swaptions on $\{S^{i,40}\}_i$the situation is different. According to (2.5) a coterminal swap expiring in T_i can be written as

$$S_t^{i,40} = \sum_{k=i}^{39} \omega_k^{i,40}(t) F_t^k.$$

This shows, that the swap depends on all forward rates that expire in T_i or later and not, as in the case of a CMS expiring in T_i, on only a small part. Hence, using swaption prices on a coterminal swap $S^{i,40}$ we can induces a condition on each entry of both the i-th row and column of the correlation matrix ρ and r, respectively. If the correlation of the forward rates fall rapidly around the main diagonal, a terminal minimal correlation is still given through the condition on the outer entries.

To incorporate CMS spreads in a general calibration to dynamics implied by the mar-

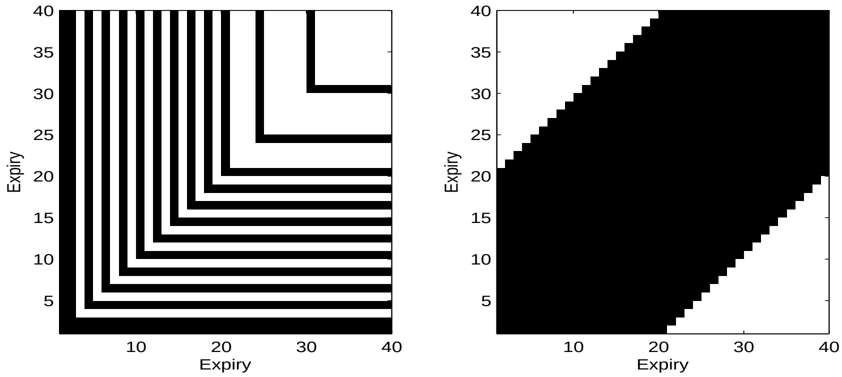

Figure 8.3.: Here we see which entries of ρ and r, respectively, are governed by options on coterminal swaps or CMS spreads. On the right, we see the entries which depend on a set of coterminal swaptions with the expiries which are quoted in the market. On the right, we see the entries which depend on a set of CMS spreads options with expires every half year. Notice the *wings* – the entries without any condition.

ket, we suggest a joint calibration to swaptions and CMS spread options, to grantee a reliable calibration. In the calibration process the options should be weighted, say 80% swaptions and 20% CMS spread options, since derivatives on swaps are more liquid then on CMS spreads.

9. Conclusion

We introduced the SABR-LMM to give a market model that can describe the dynamics of a number of forward rates in an arbitrage free framework. The Libor-Market model has its origin in the simple SABR model introduced by *Hagan* [17]. The goal of this work was to calibrate the SABR-LMM, effectively that means we had to estimate the volatility structures as well as the correlations between the volatility processes and the forward rates.

For the volatility calibration we kept the dynamics of the single forward rates as close to the SABR-models' as possible. We achieved this with special calibration techniques to calibrate the time-dependent volatility functions to caplet prices.
For the calibration of the correlation structure we followed two approaches. In one approach we used swaps and the other one we used CMS spreads as target products. In both cases we estimated the assumed SABR dynamics depending on the correlation matrix P. To estimate the swap dynamics we followed the classical approach of freezing swap weights [36], whereas in the case of CMS spreads we used Markovian Projection to simplify the target processes. Here we extended the work of *Kienitz & Wittke* [22] to our SABR-LMM with time-dependent coefficients.

In the empirical part we got supporting evidence for that the calibration to swap dynamics works quite well. This is because we almost obtained the real prices of swaptions by Monte Carlo Simulations. Our target was to capture the assumed simple SABR swap dynamics in the SABR-LMM as accurately as possible. The low deviation of the swaption prices show that we hit the smile for swaps very well. Now, since the smile is directly linked to SABR dynamics via formula for implied volatility, this shows we regain the swap dynamics in the SABR-LMM with the desired accuracy.
In contrast, the calibration to CMS spread option prices alone did not provide the desired accuracy, because the deviations of simulated swaption prices were too big. We explained possible reasons for this.
Further, we saw that, at least in an environment of high volatility, the parametriza-

tion of ρ can have a big impact on the quality of the calibration. This is what we can conclude from that the parametrization of *Lutz* worked much better than the one of *Schoenmakers* & *Coeffey*.

In this work we did not take a look at the hedging performance of the SABR-LMM or, in general, how hedging works. The SABR-LMM is strongly linked to simple SABR models due to its calibration. It would be interesting to compare the performance of SABR-LMM hedges with hedges done in the simpler SABR environment. For example hedges of derivatives on swaps, CMS or forwards could be researched. Note that for those underlyings we already gave an approximation for the SABR dynamics in a SABR-LMM world. So everything is provided.

A. Appendix

A.1. Parameters Obtained from the Calibration on Data as of 06.04.2006

Table A.1.: The Parameters for g and h

Function	a	b	c	d
g	-0.0067	0.0183	0.6990	0.0265
h	0.4544	0.0000	0.9098	0.1781

Table A.2.: The Parameters k_0^i

i	k_0^i	i	k_0^i	i	k_0^i	i	k_0^i
1	0.9128	11	1.0322	21	0.9634	31	0.9119
2	1.1153	12	1.0252	22	0.9469	32	0.8999
3	1.1010	13	1.0253	23	0.9396	33	0.8858
4	1.1028	14	1.0243	24	0.9404	34	0.8736
5	1.0848	15	1.0124	25	0.9471	35	0.8570
6	1.0537	16	0.9965	26	0.9481	36	0.8451
7	1.0391	17	0.9935	27	0.9487	37	0.8316
8	1.0541	18	1.0029	28	0.9456	38	0.8171
9	1.0650	19	1.0060	29	0.9373	39	0.8030
10	1.0514	20	0.9853	30	0.9269		

Table A.3.: The Parameters ζ^i

i	ζ^i	i	ζ^i	i	ζ^i	i	ζ^i
1	1.0297	11	1.1219	21	0.9930	31	0.9863
2	0.9631	12	1.0549	22	0.9958	32	0.9982
3	0.9450	13	0.9948	23	0.9944	33	1.0137
4	0.9678	14	1.0279	24	0.9848	34	1.0226
5	1.0485	15	1.0217	25	0.9608	35	1.0484
6	1.0450	16	0.9121	26	0.9594	36	1.0575
7	1.0035	17	0.8507	27	0.9476	37	1.0699
8	0.9868	18	0.9096	28	0.9474	38	1.0857
9	1.0185	19	0.9494	29	0.9587	39	1.1013
10	1.0852	20	0.9848	30	0.9649		

Table A.4.: The Parameters of the Submatrices of P, if the (5L) Parametrization is used for ρ

Submatrix	α	β	γ	η	ρ_∞
ρ	1.3551	0.0113	-0.1976	6.3000	0.3876
	β	ρ_∞			
r	0.35	0.4760166428			
	λ_1	λ_2			
R	15.30546563	1.611796712			

Table A.5.: The Parameters of the Submatrices of P, if the (2SC) Parametrization is used for ρ

Submatrix	γ	η
ρ	0.4660	0.0000
	β	ρ_∞
r	0.35	0.4831
	λ_1	λ_2
R	20.0	1.6349

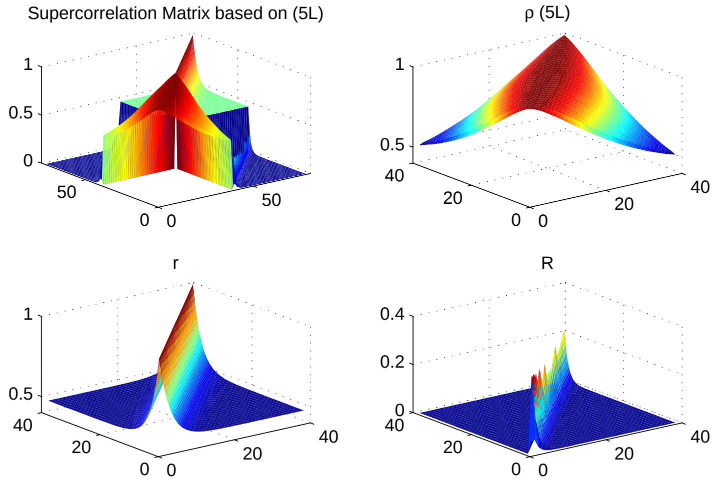

Figure A.1.: The correlation matrices from the calibration for 04.06.2006, if we use Lutz' (5L) parametrization for ρ. The plot was done in Matlab.

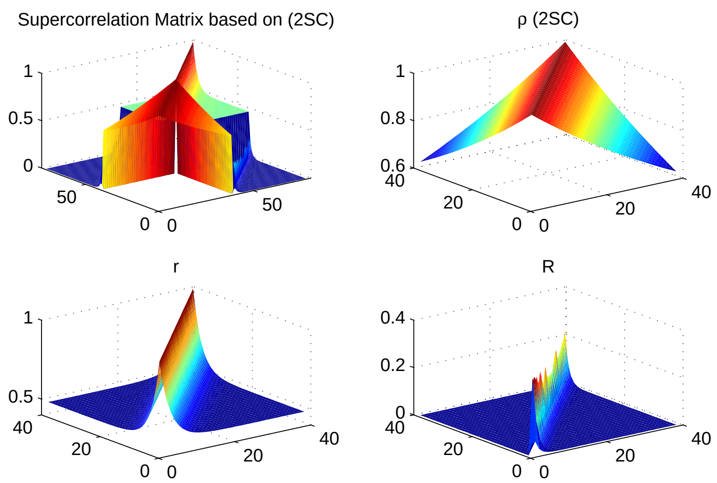

Figure A.2.: The correlation matrices from the calibration for 04.06.2006, if we use Schoenmakers & Coffey's (2SC) parametrization for ρ. The plot was done in Matlab.

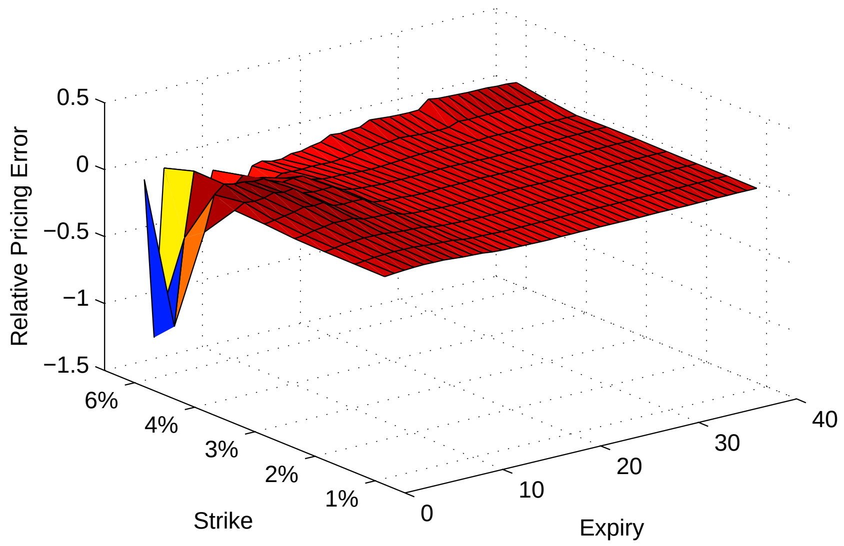

Figure A.3.: The relative pricing errors for caplets. The pricing error for high strikes and short expiries is clearly visible. The plot was done in Matlab.

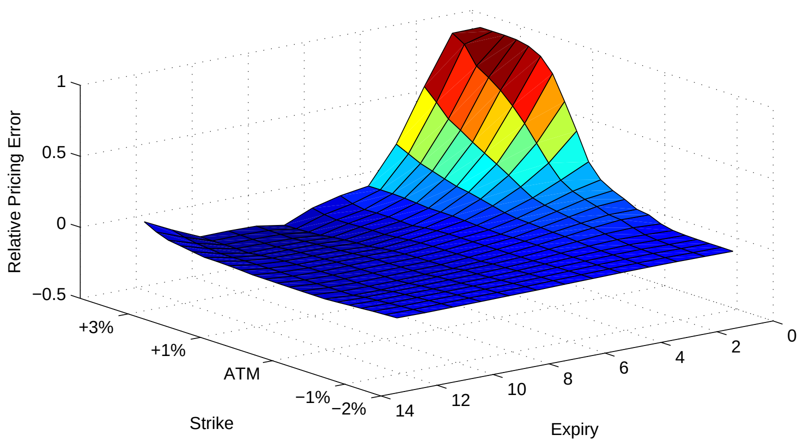

Figure A.4.: The relative pricing errors for swaptions of if for ρ the (5L) parametrization is used. The swaption prices are reproduced well. Only swpations with short expiry and high strikes are problematic. The plot was done in Matlab.

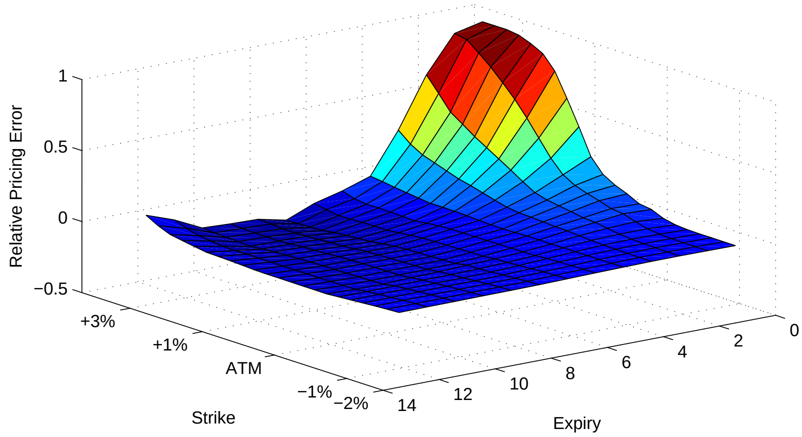

Figure A.5.: The relative pricing errors for swaptions of if for ρ the (2SC) parametrization is used. The errors look identical to the one obtained by using the (5L) parametrization. The plot was done in Matlab.

A.2. Parameters Obtained from the Calibration on Data as of 21.07.2014

Table A.6.: The Parameters for g and h

Function	a	b	c	d
g	-0.0473	0.0479	1.2295	0.0472
h	0.4494	0.3829	1.8611	0.2543

Table A.7.: The Parameters k_0^i

i	k_0^i	i	k_0^i	i	k_0^i	i	k_0^i
1	1.1519	11	1.2188	21	0.9361	31	0.8656
2	0.8285	12	1.2132	22	0.9186	32	0.8634
3	0.8166	13	1.1916	23	0.9178	33	0.8612
4	0.9520	14	1.1605	24	0.8867	34	0.8593
5	1.0989	15	1.1228	25	0.8912	35	0.8569
6	1.2382	16	1.0656	26	0.8970	36	0.8538
7	1.2767	17	1.0415	27	0.9031	37	0.8488
8	1.3575	18	1.0049	28	0.9085	38	0.8431
9	1.3617	19	0.9892	29	0.9121	39	0.8351
10	1.2820	20	0.9473	30	0.8685		

Table A.8.: The Parameters ζ^i

i	ζ^i	i	ζ^i	i	ζ^i	i	ζ^i
1	1.0664	11	0.3975	21	0.9408	31	0.9224
2	0.6803	12	0.7351	22	0.9494	32	0.9212
3	1.1407	13	0.8114	23	0.9587	33	0.9164
4	1.3548	14	0.8358	24	0.9511	34	0.9103
5	1.3911	15	0.8370	25	0.9599	35	0.9038
6	1.2991	16	0.8239	26	0.9635	36	0.8972
7	0.5837	17	0.8534	27	0.9623	37	0.8925
8	1.0161	18	0.8796	28	0.9593	38	0.8864
9	0.8904	19	0.9137	29	0.9540	39	0.8836
10	0.6476	20	0.9212	30	0.9231		

Table A.9.: The Parameters of the Submatrices of P for the Calibration to Swaptions, if the (5L) Parametrization is used for ρ

Submatrix	α	β	γ	η	ρ_∞
ρ	2.1313	0.0111	10.000	-5.3416	0.0000
	β	ρ_∞			
r	0.1000	0.8842			
	λ_1	λ_2			
R	20.0	20.0			

Table A.10.: The Parameters of the Submatrices of P for the Calibration to Swaptions, if the (2SC) Parametrization is used for ρ

Submatrix	γ	η
ρ	1.3310	1.3309
	β	ρ_∞
r	0.2000	0.9933
	λ_1	λ_2
R	20.0	20.0

Table A.11.: The Parameters of the Submatrices of P for the Calibration to CMS Spread Options, if the (5L) Parametrization is used for ρ

Submatrix	α	β	γ	η	ρ_∞
ρ	7.3128	0.0111	6.2121	-2.4125	0.0260
	β	ρ_∞			
r	0.03504	0.5767			
	λ_1	λ_2			
R	0.1	0.4017			

Table A.12.: The Parameters of the Submatrices of P for the Calibration to CMS Spread Options, if the (2SC) Parametrization is used for ρ

Submatrix	γ	η
ρ	9.4933	9.4933
	β	ρ_∞
r	0.3000	0.7500
	λ_1	λ_2
R	0.7294	0.3000

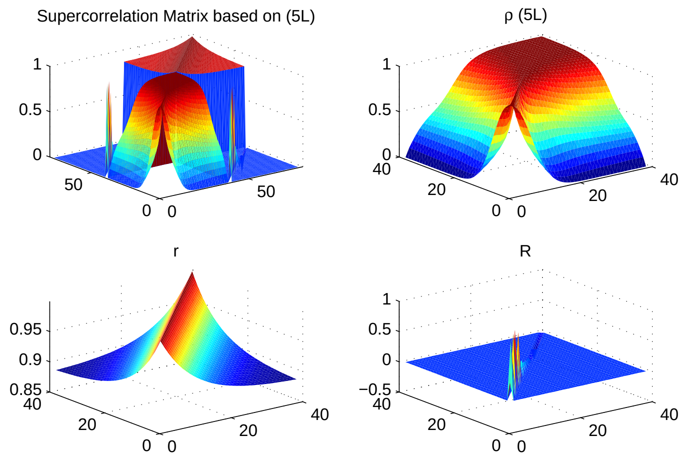

Figure A.6.: The calibrated correlation matrices, if we use Lutz' (5L) parametrization for ρ and calibrate to swaptions. Observe how the general shape relative to 2006 has changed. The short-term forward rates are now way less correlated then before. The plot was done in Matlab.

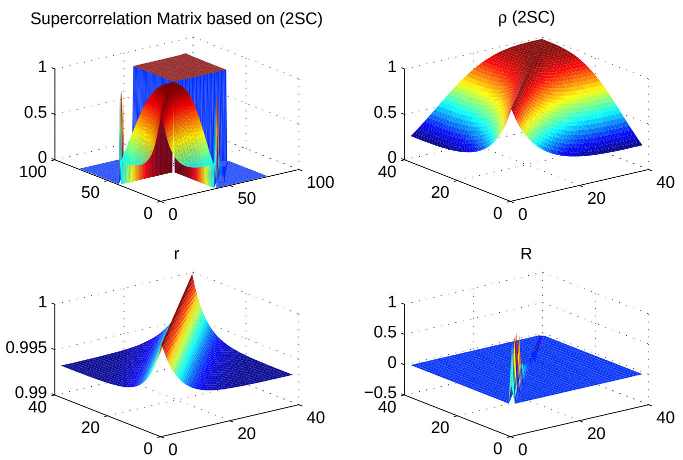

Figure A.7.: The calibrated correlation matrices, if we use the simpler Schoenmakers & Coffey's (2SC) parametrization for ρ and calibrate to swaptions. As for the (5L) parametrization the short term forward rates are less correlated then in 2006, but the level of correlation is around 30%. This is a result of the inflexibility of the (2SC) parametrization. The plot was done in Matlab.

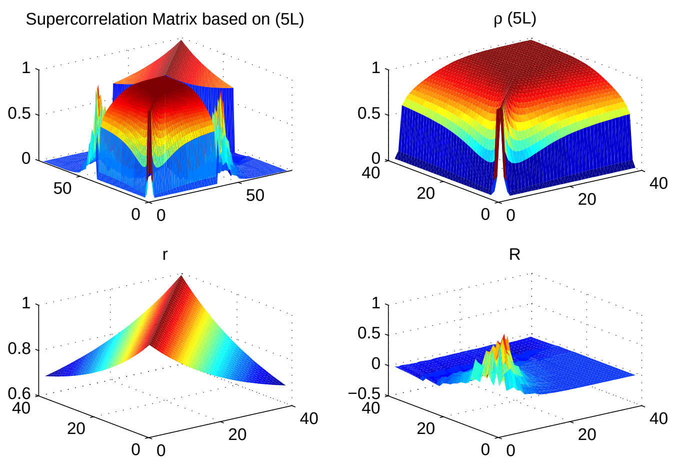

Figure A.8.: The calibrated correlation matrices, if we use Lutz' (5L) parametrization for ρ and calibrate to CMS spread options. The shape of ρ is almost identical to the shape from the calibration to swaptions. How ever the volatility/volatility correlation is in general lower and the surface of R seems a little bit smoother. The plot was done in Matlab.

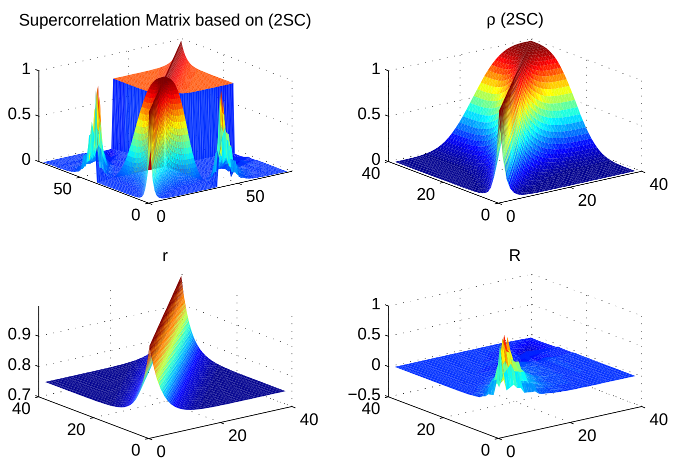

Figure A.9.: The calibrated correlation matrices, if we use the simpler Schoenmakers & Coffey's (2SC) parametrization for ρ and calibrate to CMS spread options. The forward rates in generally are less correlated compared with the (5L) parametrization. The plot was done in Matlab.

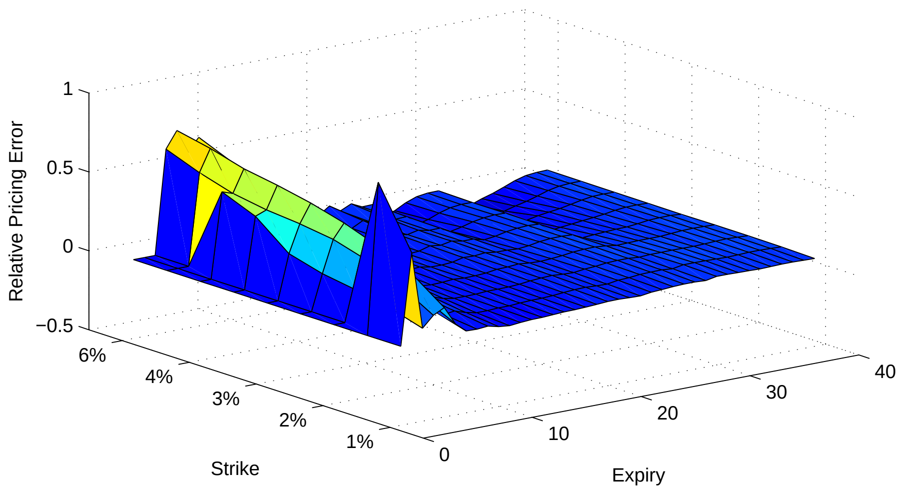

Figure A.10.: The relative pricing errors for caplets. Similar to 04.06.2006 the error becomes lower with growing expiry. The plot was done in Matlab.

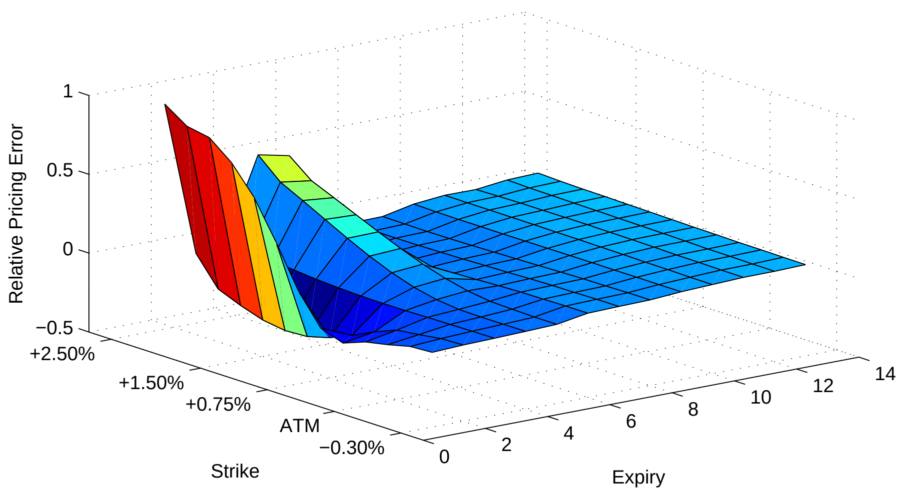

Figure A.11.: The relative pricing errors for swaptions, if the model is calibrated to swaptions and for ρ the (5L) parametrization is used. Besides the error peak for coterminal swaptions with expiries of 2 years and 3 years the surfaces looks similar to the one from 2006. The plot was done in Matlab.

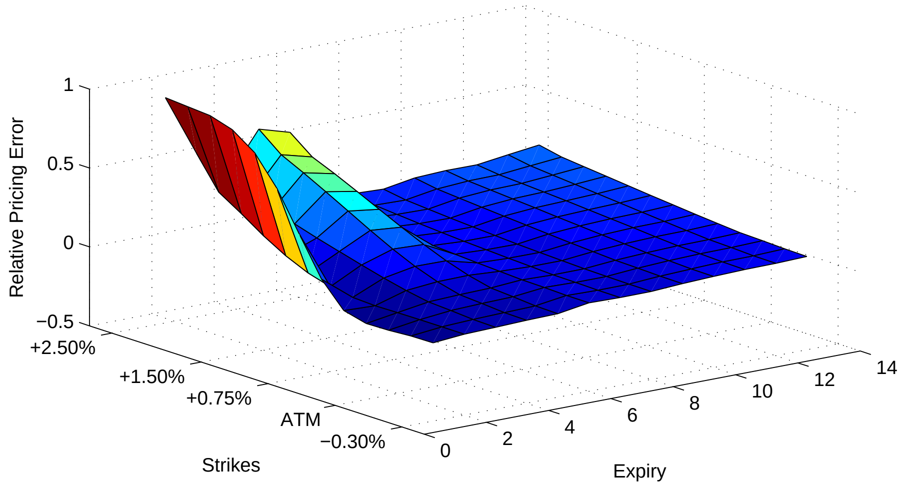

Figure A.12.: The relative pricing errors for swaptions, if the model is calibrated to swaptions and for ρ the (2SC) parametrization is used. The level of errors is higher than for the (5L) parametrization. The plot was done in Matlab.

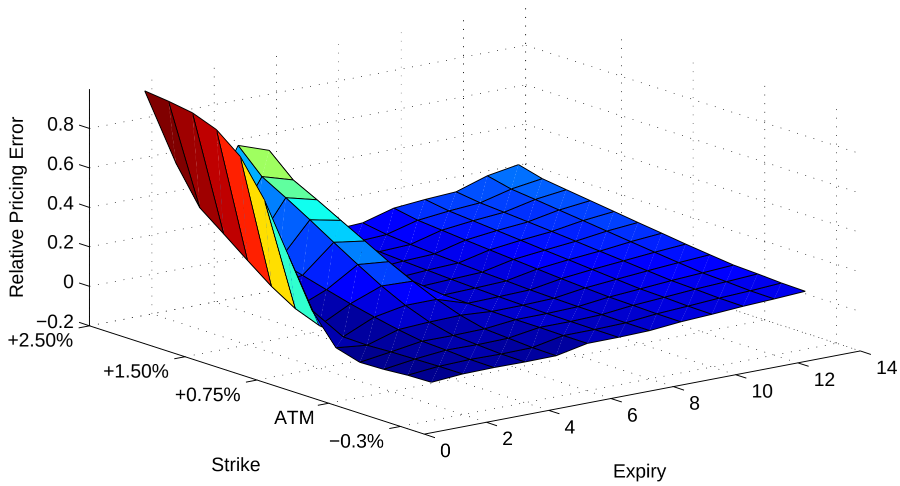

Figure A.13.: The relative pricing errors for swaptions, if the model is calibrated to CMS spread options and the (5L) parametrization is used for ρ. The surface looks almost like the one obtained by calibration to swaps, but the general level is a bit higher. The plot was done in Matlab.

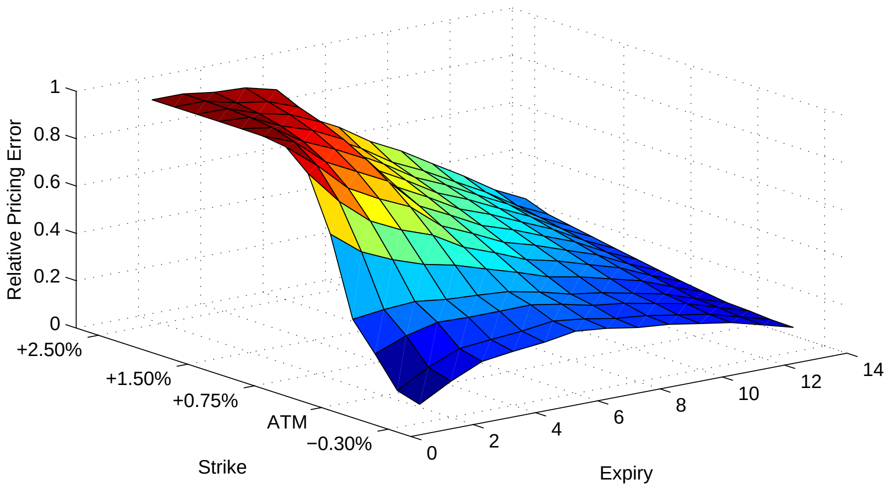

Figure A.14.: The relative pricing errors for swaptions, if the model is calibrated to CMS spread options and for ρ the (2SC) parametrization is used. The plot was done in Matlab.

Bibliography

[1] De Ceuste M. Annaert J., Claes B.A. and Zhang H. Estimating the Yield Curve Using the Nelson-Siegel Model. 2008.

[2] Dariusz Brace A., Gatarek D. and Marek M. The Market Model of Interest Rate Dynamics. *Mathematical Finance*, 7(2):127–155, 1997.

[3] Capitani C. Brigo D. and F. Mercurio. On the Joint Calibration of the Libor Market Model to Caps and Swaptions Market Volatilities. 2001.

[4] Lo C-F. The Sum and Difference of Two Constant Elasticity of Variance Stochastic Variables. *Journal of Applied Mathematics*, (4):1503–1511, 2013.

[5] Oosterlee C. Chen B. and van Weeren S. Analytic Approximation to Constant Maturity Swap Convexity Correction in a Multi-Factor SABR Model. *International Journal of Theoretical and Applied Finance*, 13(7):1019–1046, 2010.

[6] Oosterlee C. W. Chen B. and Van Der Weide H. Efficient Unbiased Simulation Scheme for the SABR Stochastic Volatility Model, 2011.

[7] Brigo D. and Mercurio F. Interest Rate Models - Theory and Practice: With Smile, Inflation and Credit (Springer Finance), 2006.

[8] Filipovic D. Fixed Income Models. *Princeton University*, 2002.

[9] Protter P. E. *Stochastic Integration and Differential Equations*. Springer, 2nd edition, 2005.

[10] Black F. The Pricing of Vommodity Contracts. *Journal of Financial Economics*, 3(1-2):167–179, 1976.

[11] Jamshidian F. LIBOR and Swap Market Models and Measures. *Finance and Stochastics*, 1(4):293–330, 1997.

[12] Mercurio F. A LIBOR Market Model with Stochastic Basis, 2010.

[13] Mercurio F. and Pallavicini A. Swaption Skews and Convexity Adjustments. 2006.

[14] Andersen L. B. G. and Piterbarg V. V. *Interest Rate Modeling. Volume 1: Foundations and Vanilla Models.* Atlantic Financial Press, 2010.

[15] Brunick G. and Shreve S. Mimicking an Itô Process by a Solution of a Stochastic Differential Equation. *The Annals of Applied Probability*, 23(4):1584–1628, 2013.

[16] Schumann E. Gilli M., Grosse S. and Zhang H. Calibrating the Nelson-Siegel-Svensson Model. 2010.

[17] Lesniewski A. Hagan P. LIBOR Market Model with SABR style stochastic volatility. 2008.

[18] Lesniewski A. Hagan P., Deep K. and Woodward D. Managing Smile Risk. *Wilmott magazine*, pages 84–108, 2002.

[19] Gyöngy I. Mimicking the One-Dimensional Marginal Distributions of Processes Having an Ito Differential. *Probability Theory and Related Fields*, (71):1503–1511, 1986.

[20] Karatzas I. and Shreve S. *Brownian Motion and Stochastic Calculus (Graduate Texts in Mathematics) (Volume 113).* Springer, 2nd edition, 8 1991.

[21] Hull J. and White A. Forward Rate Volatilities, Swap Rate Volatilities, And The Implementation of the Libor Market Model. *Journal of Fixed Income*, 2000.

[22] Kienitz J. and Wittke M. Option Valuation in Multivariate SMM/ Models - with an Application to the CMS Spread -. 2010.

[23] Obloj J. Fine-tune your Smile: Correction to Hagan et al. Science & finance (cfm) working paper archive, Science & Finance, Capital Fund Management, 2007.

[24] Schoenmakers J. *Robust Libor Modelling and Pricing of Derivative Products* . Chapman and Hall/CRC, 1 edition, 3 2005.

[25] Schoenmakers J. and Coffey B. *Stable Implied Calibration of a Multi-factor LIBOR Model Via a Semi-parametric Correlation Structure.* WIAS, 2000.

[26] Heston S. L. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*, 6:327–343, 1993.

[27] Paulot L. Asymptotic Implied Volatility at the Second Order with Application to the SABR Model. Technical report, 2010.

[28] Cresnik M. Master's Thesis: Title Unknown. *http://wwwmath.uni-muenster.de/statistik/paulsen/Abschlussarbeiten/Masterarbeiten/*, 2014.

[29] Lutz M. Extracting Correlations from the Market: New correlation parameterizations and the Calibration of a Stochastic Volatility LMM to CMS Spread Options. 2010.

[30] Islah O. Solving SABR in Exact Form and Unifying It with Libor Market Model. 2009.

[31] Doust P. Modelling Discrete Probabilities. 2007.

[32] Hagan P. Convexity Conundrums : Pricing CMS Swaps, Caps and Floors. 2006.

[33] Hagan P. and West G. Methods for Constructing a Yield Curve. *Wilmott magazine*, pages 70–81, 2008.

[34] Pietersz R. and Groenen P. Rank Reduction of Correlation Matrices by Majorization. *Quantitative Finance*, 4(6):649–662, 2004.

[35] Rebonato R. *Volatility and Correlation: The Perfect Hedger and the Fox.* Wiley, 2 edition, 9 2004.

[36] McKay K. Rebonato R. and White R. The SABR/LIBOR Market Model: Pricing, Calibration and Hedging for Complex Interest-Rate Derivatives, 2009.

[37] Bates D. S. Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in PHLX Deutschemark Options. 1993.

[38] Johnson S. and Nonas B. Arbitrage-free Construction of the Swaption Cube. 2009.

[39] Bank For International Settlements. OTC Derivatives Statistics at End-December 2013. 5 2014.

[40] Girsanov I. V. On Transforming a Certain Class of Stochastic Processes by Absolutely Continuous Substitution of Measures. 5(3):285–301, 1960.

[41] Piterbarg V. Markovian Projection Method for Volatility Calibration. 2006.

[42] Schachermayer W. and Teichmann J. How Close Are The Option Pricing Formulas of Bachalier And Black-Merton-Scholes? 11 2007.