Multicurrency Extension of a Multiple Stochastic Volatility Libor Market Model

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim
Fachbereich Mathematik und Informatik
der Goethe-Universität
in Frankfurt am Main

von
Stanley Sijan Mathew
Dipl. Math., M.A. Statistics
aus Sofia

Frankfurt

May 5, 2009

vom Fachbereich **Mathematik und Informatik** der Goethe-Universität als Dissertation angenommen.

Dekan : Prof. Dr.-Ing. Detlef Krömker

1. Gutachter: Prof. Dr. Peter E. Kloeden (Betreuer)

2. Gutachter: Dr. John G.M. Schoenmakers (Betreuer)

Datum der Disputation: 4. Mai 2009

Danksagung

Die vorliegende Dissertation enstand aus einer Kooperation der Goethe-Universität in Frankfurt und des Weierstraß-Instituts für Angewandte Analysis und Stochastik in Berlin. Daher möchte ich an erster Stelle meinen Betreuern Prof. Dr. Peter E. Kloeden und Dr. John G.M. Schoenmakers meinen ausdrücklichen Dank für die intensive und erfolgreiche Zusammenarbeit aussprechen. In unseren Diskussionen profitierte ich vor allem von ihrer Fachkenntnis und ihrem umsichtigen Führungsstil. Besonderer Dank gilt ihnen für das mir entgegengebrachte Vertrauen, meine Forschungstätigkeit im Rahmen des Promotionsprojektes weitgehend eigenverantwortlich zu gestalten.

Ebenso schulde ich der Vereinigung von Freunden und Förderern der Johann Wolfgang Goethe-Universität e.V. und der Willkomm Stiftung aufrichtigen Dank für die Finanzierung wesentlicher Forschungsaufenthalte.

Große Dankbarkeit empfinde ich auch für den Rückhalt und das Verständnis seitens meiner Familie, die mir mit Geduld und Zuspruch beistand.

Zusammenfassung

Das LIBOR Markt Modell (LMM) ist seit seiner Entwicklung in den Veröffentlichungen von Brace, Gatarek, Musiela (1997), einerseits, und unabhängig von diesen von Miltersen, Sandmann, Sondermann (1997), andererseits, zu dem anerkanntesten Instrument zur Modellierung der Zinsstruktur und der damit verbundenen Preisfindung für relevante Finanzderivate geworden. LIBOR steht dabei für *London Inter-Bank Offered Rate*, ein täglich in London fixierter Referenz-Zins für kurzfristige Anlagen. Drei- oder sechsmonatige Laufzeiten sind in Verbindung mit dem LMM üblich.

Die Forschung zur Verbesserung dieses Modells hat in den letzten Jahren an Zuwachs gewonnen. Beim Versuch den Fehler der Anpassung an die täglich beobachteten Preise von Zinsoptionen wie Caps und Swaptions zu verringern, erhält man in der Folge auch genauere Bewertungen für andere, exotischere, Derivate. Die zugrunde liegende und zentrale Idee des LMM besteht darin, die Forward (Termin) Zinsen direkt als primären (Vektor) Prozess mehrerer LIBOR Sätze zu betrachten und diese simultan zu modellieren, anstatt sie nur herzuleiten aus einem übergeordneten, unendlich dimensionalen Forward Zinsprozess, wie im zeitlich früher entwickelten Heath-Jarrow-Morton Modell. Das überzeugendste Argument für diese *Diskretisierung* ist, dass die LIBOR Sätze direkt im Markt beobachtbar sind und ihre Volatilitäten auf eine natürliche Weise in Beziehung gebracht werden können zu bereits liquide gehandelten Produkten, eben jenen Caps und Swaptions.

Dennoch beinhaltet das Modell eine gravierende Insuffizienz, indem es keine Krümmung der Volatilitätsoberfläche, im Hinblick auf Optionen mit verschiedenen Basiszinsen, abbildet. Wie im einfachen eindimensionalen Black-Scholes Modell prägen sich auch hier die Ungenauigkeiten der Verteilung in fehlenden *heavy tails* deutlich aus. Smile und Skew Effekte sind erkennbar. Im klassischen LIBOR Markt Modell wird in Richtung der Basiszins-dimension nur eine affine Struktur erzeugt, welche bestenfalls als Approximation für die erwünschte Oberfläche dienen kann. Die beobachteten Verzerrungen führen naturgemäss zu einer ungenauen Abbildung der Realität und fehlerhaften Reproduktion der Preise in Regionen, die ein wenig entfernt vom Bereich *am Geld* liegen. Derartig ungewollte Dissonanzen in Gewinn und Verlustzahlen führten z.B. in 1998 zu gravierenden Verlusten im Zinsderivateportfolio der heutigen Royal Bank of Scotland.

Diverse Versuche sind in den letzten Jahren unternommen worden, um eine bessere Anpassung an die beobachtete gekrümmte Fläche zu erlangen. Es wurde schliesslich offensichtlich, dass man nicht umhin kam, entweder eine

Sprungdiffusion zu integrieren oder aber geeignete Faktoren mit einer stochastischen Volatilität auszustatten. Erstere wurden bereits von Merton (1976) studiert, später jedoch für LIBOR Modellierung von Glasserman und Merener (2001), Glasserman und Kou (2003) und Belomestny und Schoenmakers (2006) wiederentdeckt. Die Klasse, welche die grösste Flexibilität aufzuweisen schien, waren jedoch die Modelle mit stochastischer Volatilität (SVM). Sie wurden unter anderem von Andersen und Brotherton-Ratcliffe (2001), Wu und Zhang (2002) und Piterbarg (2003) vorgeschlagen.

Noch bevor diesen sind sogenannte *local-volatility* Modelle erwogen worden. Die zwei Arme dieser Klasse sind die *constant elasticity of variance* (CEV) Modelle, mittlerweile mehrfach erweitert, z.B. von Wu (2003), aber ursprünglich vorgeschlagen von Andersen-Andreasen (2000), und die *displaced diffusion* (DD) Modelle, welche zwar auf die Arbeit von Rubinstein (1983) zurückgehen, aber zuerst von Rebonato und Joshi (2002) auf Zinsderivate angewendet wurden. Ein wesentlicher Kritikpunkt an beide Modelltypen ist, dass sie zwar einen monotonen Skew generieren können, aber keinen zufriedenstellenden Smile hervorbringen. Abhilfe brachten die oben erwähnten Alternativen.

Sprungdiffusionen fingen das *heavy tail* Phänomen erstaunlich gut ein. Es gibt jedoch keine zufriedenstellenden Ergebnisse hinsichtlich der Stabilität von Kalibrationen, d.h. die Anpassung der Modellparameter an die realen Marktdaten. Ob die Parameter von Sprungdiffusionen stetig variieren, wenn Input Daten perturbiert werden, ist eine offene Frage.

Stochastische Volatilitäts Modelle, andererseits, performen gut, wenn es um die Erzeugung von gekrümmten Flächen geht. Unglücklicherweise produzieren sie keinen Skew, wenn man Unkorreliertheit zwischen LIBOR Rate und stochastischem Volatilitätsprozess unterstellt. Unter den bekannten Modellen erlauben nur das bereits erwähnte Wu/Zhang und das SABR Modell die notwendigen nicht trivialen Korrelationskoeffizienten.

Ein Nachteil des Wu/Zhang Modells ist, dass der CIR Prozess, welcher als Volatilitätskomponente eingeführt wird, nur eindimensional ist. Instabilitäten sind somit bei der Kalibrierung zu erwarten, wenn dieser eindimensionale Prozess an $n-1$ LIBOR Sätze anzugleichen ist. Für verschiedene Maturitäten weisen diese nachweislich ein sehr verschiedenartiges Verhalten auf. So ist es nicht verwunderlich, dass Wu und Zhang in ihrem numerischen Teil ihr Modell nicht an Marktpreise kalibrieren, sondern lediglich mit exogen vorgegebenen Parametern arbeiten. SABR Modelle, andererseits, beinhalten zwar sehr allgemeine Ausdrücke für LIBOR Modelle mit stochastischer

Volatilität, jedoch bleibt auch hier unklar, wie eine stabile Kalibrierung erfolgen soll.

Beim direkten Vergleich der beiden zuletzt betrachteten Ansätze, unterstützen Experten und Teilnehmer des LIBOR Derivate Marktes die Idee, dass die Dynamik von Prozessen mit stochastischer Volatilität eine ausgeprägtere Flexibilität offeriert und einen besseren Fit gewährleistet, als die Dynamik von Sprungprozessen. Für Details siehe Chen and Scott (2001).

Die Arbeit gliedert sich wie folgt: Das erste Kapitel dient als Einführung in die Thematik und behandelt in späteren Kapiteln erforderliche Konzepte. Im zweiten Kapitel werden einige frühere Modelle erläutert und eine Methode eingeführt mit deren Hilfe die Optimierung der Parameter gelingt. Als Konsequenz aus den oben beschriebenen Ergebnissen, wird im dritten Kapitel ein multiples stochastisches Volatilitäts Modell vorgeschlagen, welches Korrelationen zwischen LIBOR Raten und Volatilitätsprozessen zulässt. Ferner wird dort eine Routine zur Kalibration der Parameter empfohlen, welche robuste Schätzungen verspricht. Im vierten Kapitel wird das Modell erweitert auf den Fall zweier Währungen. Der Algorithmus zur Kalibrierung wird übertragen.

Das in dieser Arbeit betrachtete Modell ist das Folgende:

$$\frac{dL_i}{L_i} = (...)dt + \sqrt{1-r^2}\gamma_i \cdot dW + r\overline{\gamma}_i \cdot dU, \quad 1 \leq i < n\,,$$

$$dU_k = \sqrt{v_k}d\widetilde{W}_k\,,$$

$$dv_k = \kappa_k(\theta_k - v_k)dt + \sigma_k\sqrt{v_k}\left(\rho_k d\widetilde{W}_k + \sqrt{1-\rho_k^2}d\overline{W}_k\right),$$

wobei $k = 1, ..., n-1$. $\widetilde{W}$ und $\overline{W}$ sind unabhängige (n-1)-dimensionale Standard Brownsche Prozesse, beide wiederum unabhängig von W. Somit sind $\gamma_i \in \mathbb{R}^{n-1}$ und $\overline{\gamma}_i \in \mathbb{R}^{n-1}$. Die Dynamik ist im zugrunde liegenden Wahrscheinlichkeitsmaß gegeben, daher ist der Drift zunächst nicht näher spezifiziert.

Für $r = 0$ erhält man das Standard LIBOR Markt Modell. Der Parameter r sollte somit als "Allokations"- oder "Proportions"-Faktor verstanden werden, welcher quantifiziert, wieviel vom originären Standard LIBOR Modell *im Spiel* bleibt. Für kleine Werte von r, lässt sich dieses erweiterte Modell als Perturbation des Standard Modells auffassen. Diese Perturbationen gehen auf Wurzel-Diffusions Prozesse zurück, auch CIR Prozesse genannt, weshalb die Konstruktion sehr an das bewährte Heston Modell bei eindimensionalen Prozessen erinnert.

Die obige Version dient der besseren Anschauung, aus technischen Gründen wird jedoch häufig die logarithmierte Version im terminalen Maß P_n verwendet

$$
\begin{aligned}
d\ln L_i &= -\frac{1}{2}|\Gamma_i|^2 dt - \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1+\delta L_j} \left(\sum_{k=1}^{2n-2} \Gamma_{jk}\Gamma_{ik} \right) dt + \Gamma_i \cdot d\mathcal{W}^{(n)} \\
&= -\frac{1}{2}|\Gamma_i|^2 dt - \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1+\delta L_j} \Gamma_j \cdot \Gamma_i dt + \Gamma_i \cdot d\mathcal{W}^{(n)},
\end{aligned}
$$

wobei

$$
\Gamma_i = \begin{pmatrix} \sqrt{1-r^2}\gamma_{i1} \\ \cdot \\ \cdot \\ \sqrt{1-r^2}\gamma_{i,n-1} \\ r\overline{\gamma}_{i,1}\sqrt{v_1} \\ \cdot \\ \cdot \\ r\overline{\gamma}_{i,n-1}\sqrt{v_{n-1}} \end{pmatrix} \quad d\mathcal{W} = \begin{pmatrix} dW_1 \\ \cdot \\ \cdot \\ dW_{n-1} \\ d\widetilde{W}_1 \\ \cdot \\ \cdot \\ d\widetilde{W}_{n-1} \end{pmatrix}.
$$

Nach Festlegung des Maßes ist der Driftterm bestimmbar. Nachdem durch Normierung, o.B.d.A., $\theta_k = 1$ gesetzt werden kann, besteht die Herausforderung nun darin die restlichen neuen Parameter κ_k, σ_k und ρ_k, für $1 \leq k \leq n-1$, somit $3n-3$ neue Parameter, auf eine stabile Weise aus den vorhandenen Marktpreisen zu schätzen. Der Parameter r kann entweder mitkalibriert oder auch nach Belieben vorab festgesetzt werden.

Es ist keine explizite Darstellungen für die Verteilung von $\ln L_i$ bekannt. Die charakteristische Funktion von $\ln L_i$ ist jedoch unter bestimmten Bedingungen an die Koeffizientenfunktion angebbar. Geschickte Maßwechsel ermöglichen zunächst die Elimination des Drifttermes. In den verbleibenden Ausdrücken schaffen die Wurzelterme der Form $r\overline{\gamma}_{i,k}\sqrt{v_k}d\widetilde{W}_k$ Probleme. Sollten $\overline{\gamma}_{i,k}$ jedoch konstant gewählt werden können, steht der Bestimmung der charakteristischen Funktion nichts mehr im Wege. Eine gute und vor allem konstante Approximation für $\overline{\gamma}_i$ wird angegeben.

Die Stabilität betreffend ist Folgendes festzuhalten. Die Einführung von weiteren $3n-3$ Parametern mag zunächst Zweifel an einer stabile Optimierung aufkommen lassen. Es gelingt jedoch die Korrelationsmatrix in einer Weise zu zerlegen, dass nur jeweils über drei Parameter pro Iterationsschritt optimiert werden muss. Dies wiederum *stabilisiert* die Kalibration erheblich. Um dies zu sehen, müssen die Capletpreise näher untersucht werden. Sie werden ermittelt via

$$C_i(K) = \delta B_{i+1}(0)\,(L_i(0)-K)^+ \\ + \frac{\delta B_{i+1}(0)L_i(0)}{2\pi}\int_{-\infty}^{\infty}\frac{1-\varphi_{i+1}(z-\mathfrak{i}\,;T_i)}{z(z-\mathfrak{i})}e^{-\mathfrak{i}\,z\ln\frac{K}{L_i(0)}}dz.$$

Alle Größen sind hierbei bekannt. B_{i+1} repräsentiert einen Zerobond mit Laufzeit T_i. Es wird gezeigt, dass die unter dem Integral vorzufindende (bedingte) charakteristische Funktion φ_{i+1} von $\ln L_i(T_i) - \ln L_i(0)$ unter dem Maß P_{i+1} gegeben ist durch:

$$\begin{aligned}\varphi_{i+1}(z\,;T,v) &= E_{i+1}\left[e^{\mathfrak{i}z\ln\frac{L_i(T)}{L_i(0)}}\middle|\, v_k(0)=v_k\,,k=1,...,n-1\right]\\ &= \varphi_{i+1,0}\,(z\,;T)\prod_{k=i}^{n-1}\varphi_{i+1,k}\,(z\,;T,v_k)\,,\end{aligned}$$

wobei

$$\varphi_{i+1,0}(z\,;T) = \exp\left(-\frac{1}{2}(1-r^2)\eta_i^2(T)\left(z^2+\mathfrak{i}\,z\right)\right),\quad \eta_i^2(T)=\int_0^T|\gamma_i|^2\,dt$$

und auch jedes $\varphi_{i+1,k}\,(z\,;T,v_k)$ bekannt ist – als Lösung einer parabolischen Differentialgleichung. Entscheidend ist nun, dass obige Produktbildung bei i beginnt und jedes $\varphi_{i+1,k}$ nur von den jeweiligen Parametern κ_k, σ_k und ρ_k abhängt. Somit wird eine rückwärtsgerichtete, bei $i=n-1$ beginnende, Iteration ermöglicht, welche in jedem Schritt *nur* über genau drei Parameter optimiert.

Eine erste Fallstudie, in welcher r in die Kalibration mit eingebunden ist, ergibt folgende Parameterwerte: $r=0.24$, ausserdem ρ, σ und κ entsprechender Laufzeiten, wie in Tabelle 1 angegeben.

Im vierten Kapitel wird das vorgeschlagene LIBOR Markt Modell mit multipler stochastischer Volatilität auf den Fall zweier Währungen erweitert. Die Ergebnisse aus dem dritten Kapitel werden erfolgreich auf diesen neuen Fall

Tenor	20	19	18	17
ρ	-0.7832	-0.7832	-0.7832	-0.7832
σ	7.4920	7.4920	6.2427	5.0198
κ	2.3376	2.3376	3.9385	4.5590

Table 1: Schätzungen der Parameter für ausgewählte Laufzeiten.

übertragen. Besondere Schwierigkeiten, die es zu bewältigen galt, waren die Einbindung des Währungskurses als eigenen stochastischen Prozess, sowie die Anpassung des Algorithmus' zur Kalibration beim Übergang auf die charakteristischen Funktionen der Fremdwährungszinsen.

Bemerkung

In dieser Arbeit wird eine ökonomisch motivierte Erweiterung eines gegebenen (vor-kalibrierten) LIBOR Markt Modells um multiple stochastische Volatilitätsprozesse vorgestellt. Es wird gezeigt, dass diese Erweiterung eine schnelle (approximative) Cap und Swaption Preisfindung unter Berücksichtigung von Smile und Skew Effekten erlaubt. Ein Algorithmus zur Kalibration an die Cap-Strike-Matrix wird angegeben und in einer Fallstudie erörtert. Bei der Analyse anderer Datensätze ist ein stabiles Verhalten der kalibrierten Parameter beobachtet worden. Es sei erwähnt, dass in der vorliegenden Arbeit der Fokus auf der Entwicklung und theoretischen Analyse der Struktur des präsentierten stochastischen Volatilitätsmodells und seiner Implementierung lag. Eine tiefergehende Analyse der Kalibrationseigenschaften des Modells und seiner Performance, z.B. anhand weiterer Fallstudien oder an anderen Produktgruppen, wie CMS-spreads, ist für folgende Arbeiten vorgesehen.

Contents

1 Introduction and Preliminaries

Since Brace, Gatarek, Musiela (1997), Jamshidian (1997), and Sandmann, Sondermann, Miltersen (1997), almost independently, initiated the development of an interest rate modeling and derivative pricing tool, research around the LIBOR market model, as it is more concisely known, has grown considerably. LIBOR stands short for London Inter-Bank Offered Rate, a reference short term interest rate determined daily. In an attempt to improve the models fit to daily observed prices of standard interest rate products like caps and swaptions, one also obtains more accurate valuations for other, more exotic, derivatives. The inherent idea is to directly model the forward LIBOR or swap rate as the primary process, rather than considering it a secondary process deduced from an instantaneous rate. Moreover, the successful way to avoid explosions of the underlying stochastic process was new at that time and entailed convincing results. A corresponding zerobond market, consisting of finitely many securities, could be embedded into an arbitrage-free setting.

Nevertheless, the model incorporates a serious drawback in ignoring the curvature of the volatility surface. In its classical form all that can be accounted for is an affine hyperplane that operates as an approximation to the desired surface. As known for some time, it simply could not explain the smile and skew effects constantly observed in markets. Occasionally this caused serious problems to risk management and trading desks regarding a proper reproduction of the current state. Unwanted mismatches in profit and loss figures obscured a clear sense for reality. In 1998 RBS incurred a serious loss in their interest books purely for this reason.

Several approaches have been made in order to incorporate a better fit to the observed curvature. Soon it became evident that one could not avoid to introduce either a jump-diffusion component or a factor endowed with a stochastic volatility. Jump-diffusion models were studied already by Merton (1976) but have been reinvented for LIBOR modeling by Glasserman and Merener (2001), Glasserman and Kou (2003) and Belomestny and Schoenmakers (2006). The class generally considered to offer the most flexibility and to be of exceptionally broad form are the stochastic volatility models (SVM). They were recommended by Andersen and Brotherton-Ratcliffe (2001), Wu and Zhang (2002), Piterbarg (2003) and others.

Local-volatility models have been proposed even earlier. The two main strands are the constant elasticity of variance (CEV) models, meanwhile repeatedly extended, for example by Wu (2003), but originally proposed by

Andersen-Andreasen (2000), and displaced diffusions (DD) going back to the work of Rubinstein (1983), but first applied to interest rates by Rebonato and Joshi (2002). The major draw back of both CEV and DD models, however, is that they can in fact generate a monotone skew for the implied volatilities, but fail to create a satisfactory smile. The remedy has been to adopt one of the following alternatives.

Jump-diffusions captured the heavier tails phenomenon remarkably well. There are, however, no satisfactory results that demonstrate stability of calibration procedures. Whether parameters of jump diffusion models vary continuously when input data are perturbed, is an open question.

Stochastic volatility models performed well too, when it came to creating a curvature. Unfortunately they could not produce a skew, unless one permitted correlation between LIBOR rate and stochastic volatility process. Among the above mentioned, only Wu/Zhang's model and SABR models allow for the necessary correlation coefficients. A disadvantage of the Wu/Zhang model is that the CIR process, introduced to enhance the models volatility component, is only one-dimensional. As such instabilities are to be expected when fitting it to $n-1$ LIBORS which for different maturities demonstrably behave diversely. In their numerical part they do not calibrate to market prices, but use exogeneously provided parameters. SABR models, on the other hand, consider general expressions for stochastic volatility LIBOR models, but again it remains unclear whether calibration routines are robust.

Regarding a direct comparison of the latter two approaches, LIBOR derivative markets support the idea that the dynamics of processes with stochastic volatility components offer more flexibility and provide a better fit than processes with jumps. For details see Chen and Scott (2001).

As a consequence of all these findings, we propose a multiple stochastic volatility model that admits correlation and recommend a calibration routine that delivers robust estimations. The latter is demonstrated by numerical tests and their results.

This work will be divided into four chapters. In the first chapter we recapitulate some preliminary subjects on derivatives pricing, the parametrization of LIBOR models and results on the Heston model. Thereby we focus on bond and LIBOR market perspectives and do not consider equivalent results in stock, commodity or currency markets. Some aspects of the latter will be treated in the fourth chapter.

The second chapter lists some recent stochastic volatility models and illustrates a common method for examining them.
The third chapter introduces the multiple stochastic volatility LIBOR market model. Its characteristics are analyzed in detail and a calibration routine supplied.
In the final fourth chapter the multiple stochastic volatility LIBOR market model is extended to a multicurrency setting. The results of the single currency case are successfully applied to this scenario.

Notation For ease of notation the time argument of processes is often suppressed. In this work L, v, W, $\widetilde{W}$, $\overline{W}$, Z are always time-dependent processes, γ usually a deterministic time-dependent function. Others, like ρ and σ, are considered time-dependent in the majority of cases, but not in general. The context should clarify when. The lower index k in expressions like v_k, γ_{jk} and dW_k indicates vector components.

1.1 Principles of Derivative Pricing

An adequate development of the mathematical theory of derivatives pricing and the tools needed, is a challenge that would require a treatise using noticeably more space then we are endowed with in this work. Good books have been written that offer deep insight into the subtleties of this elegant theory. Among these are Björk (1998), Duffie (1996), Hunt and Kennedy (2000) and Glasserman (2003). Of particular importance are three principles that we introduce shortly.

1. Replicating strategies. If a security can be replicated, or hedged, through trading in the market assets, then the contracts price is the cost of the replicating strategy.

2. Under probability measures associated with the choice of a discount factor process, or numeraire, deflated asset prices are martingales. Prices are expected values of discounted payoffs under such martingale measures.

3. If there is only one such martingale measure, the market is called complete and any measurable claim can be perfectly replicated by a trading strategy.

Whereas the first principle supplies an intuitive approach of how derivative prices are determined, it does not provide tools or advice about how to actually find those replicating, later also self-financing, strategies and how to evaluate their cost of implementation. The second principle, however, does the job. It gives us a recipe to represent prices as expectations which can, given dynamics and corresponding measure, be evaluated. The subtlety of this approach underlies the fact that we must choose the dynamics of asset prices not as we observe them in the "real-world" measure, but as they evolve under a risk-adjusted probability measure.
The third point above insinuates that in an incomplete market there seem to be derivative contracts that cannot be perfectly hedged. This is literally so by definition. But what is a perfect hedge? When speaking of a hedge, or synonymously replicating strategy, one associates usually that in the replicating process one captures the unexpected movement of the underlying process. Other *unknown* stochastic variables are naturally ignored. The *classical* partial differential equations developed in that regard, some of which will find mention later, never contain partial derivatives with respect to unknown variables other then the underlying stock, for example. In other words, the hedge has to be evaluated with respect to the model employed. In Black & Scholes' formula from 1973 the option price is a function of more

then one variable. But the partial derivative with respect to a volatility σ is rarely analyzed. A typical statement about a "perfect hedge" insinuates usually nothing but the ordinary delta hedge technique and conceals other risks. Therefore, one cannot sufficiently emphasize the careful usage that should be employed with expressions like "perfectly replicated“ or "hedged“ by market assets. In some markets the assets may be of very limited number or restricted to a small subclass of all possible assets. The above phrases are not wrong as such, because they are made and meant with respect to a given model. But it is still misleading to believe that no further risks are existent. The whole discussion reduces thus to the question: What is the *correct* model? Humility would not be amiss when trying to answer this question. Here is a simple example:

Consider a market consisting of one stock and a bond in a binomial tree. The dynamics can usually be chosen arbitrage free and since the corresponding measure is unique, the market is complete. Consequently any measurable claim can be replicated as long as the dynamics do not change by unspecified sources of uncertainty. But what if the dynamics are influenced by other factors? This could happen for example when the market participants change their view about the future volatility. No trader will circumvent a possible loss through sticking to its replicating strategy recommended by his complete market model. As this example indicates, substantial problems arise when assuming a stochastic volatility, partly because the market turns incomplete. In a final remark, let us allude to a source of misunderstanding and thus frequently asked question. Working with "stochastic volatility", what we will intend to do in the rest of this thesis, does not mean that the factor in front of the Brownian differential suddenly turns into a random variable. In a general model it rather has already been a random quantity, as

$$\sigma_i(t) = \sigma_i(t, B(t, \omega))$$

indicates and in which sense the quantities below are always to be interpreted. A volatility coefficient $\sigma_i(t)$ defined in this form, however, is adapted to the filtration generated only by the Brownian motions perturbing the assets. As such, introducing stochastic volatility to a model insinuates the novelty of utilizing larger filtrations. Those will then be generated not just by the processes driving the assets, but also by additional sources of uncertainty.

1.1.1 Arbitrage Concept for Zerobond Processes

In a mathematical finance context, interest rate LIBOR forward contracts are not seen as *tradable* securities. This convention is confusing – admittedly. Are not billions of notional in swap contracts traded daily between banks or other institutional investors? As a traded security, however, only those contracts are considered feasible that can be utilized as numeraire assets. A sensible arbitrage condition for the LIBOR model can thus only be formulated by a detour over a proper arbitrage theory for zerobonds. The LIBOR forwards are derived from them later in a unique way.

For the concepts in this section we will be working in a probability space $(\Omega, \mathcal{F}, P)$ with filtration $\mathcal{F}_\infty = (\mathcal{F}_t)_{0 \le t \le T_\infty < \infty}$ generated by an $\mathbb{R}^m$-valued Brownian motion W which satisfies the *usual conditions*. On this space consider an n-dimensional process $B = (B(t))_{0 \le t \le T}$ of tradable securities which is a solution of the SDE

$$\begin{aligned} \frac{dB_i(t)}{B_i(t)} &= \mu_i(t)dt + \sigma_i(t) \cdot dW(t), \qquad B_i(0) > 0, \quad i = 1, ..., n, \\ &= \mu_i(t)dt + \sum_{j=1}^{m} \sigma_{ij}(t) dW_j(t). \end{aligned} \tag{1}$$

The coefficient processes μ and σ are assumed to be $\mathcal{F}_\infty$-predictable and to satisfy Lipschitz and growth conditions necessary to guarantee existence and uniqueness of solutions. See for example Kloeden & Platen (1992) or Glasserman (2003).
We interpret B as a price vector of n risky bonds, our assets, whose evolution is described by (1). We further enrich this market by a security representing the money market account B_0 with the help of which we define the extended market $B^0 := \{B_i, i = 0, ..., n\}$. It can be shown, see Reiss, Schoenmakers, Schweizer (2001), that in a complete market environment of system (1) the existence of a money market account can be concluded instead of required. Jamshidian (1997) showed that the existence of such a savings account is not even necessary for pricing and hedging interest rate derivative products based on LIBOR forwards. However, to require its existence from the outset will avoid technicalities and reduce the complexity of some of our proofs.
Let the money market account evolve according to the SDE

$$\frac{dB_0(t)}{B_0(t)} = r(t)dt\,, \tag{2}$$

for some predictable scalar process r.
Since the seminal work of Harrison & Kreps (1979) and Harrison & Pliska

(1981), it is known that a market is arbitrage free, if and only if there exists a state price deflator in that market. This equivalence leads to the following definition.

Definition 1 *The market B^0 in (1) is said to be arbitrage-free, if there exists an adapted process ξ on $(\Omega, \mathcal{F}, P)$ with $\xi > 0$ and $\xi_0 = 1$, such that ξB_i are martingales for all $0 \le i \le n$. The process ξ is called a state price deflator.*

Lemma 2 *ξ is a solution of the Itô SDE*

$$\frac{d\xi(t)}{\xi(t)} = -r(t)dt - \sum_{j=1}^{m} \lambda_j(t)\, dW_j(t), \quad \xi(0) = 1\,, \tag{3}$$

for a predictable vector process $\lambda \in \mathbb{R}^m$ and the predictable scalar process r. We will refer to λ as the market price of risk process.

Proof. The process ξ is adapted to $\mathcal{F}_\infty$ and ξB_0 is a martingale. We thus obtain by the martingale representation theorem

$$\xi(t)B_0(t) = \xi(0)B_0(0) + \int_0^t \widetilde{\Lambda}(u) \cdot dW(u),$$

for some adapted process $\widetilde{\Lambda} \in \mathbb{R}^m$.
Write $B_0(t) = \exp\left(\int_0^t r(u)du\right)$ and observe that from

$$\xi(t) = B_0^{-1}(t) \left(\xi(0)B_0(0) + \int_0^t \widetilde{\Lambda}(u) \cdot dW(u) \right)$$

we can deduce that

$$\begin{aligned} d\xi(t) = &-r(t)B_0^{-1}(t) \left(\xi(0)B_0(0) + \int_0^t \widetilde{\Lambda}(u) \cdot dW(u) \right) dt \\ &+ B_0^{-1}(t)\, \widetilde{\Lambda}(t) \cdot dW(t). \end{aligned}$$

Division by ξ concludes that

$$\begin{aligned} \frac{d\xi(t)}{\xi(t)} &= -r(t)dt - \Lambda(t) \cdot dW(t) \\ &= -r(t)dt - \sum_{j=1}^{m} \lambda_i(t)\, dW_j(t), \end{aligned}$$

where we define

$$\Lambda(t) := -\widetilde{\Lambda}(t) / \left(\xi(0) B_0(0) + \int_0^t \widetilde{\Lambda}(u) \cdot dW(u) \right).$$

The choice $\xi(0) = 1$ can be made without loss of generality. The representation theorem guarantees the adaptedness of $\widetilde{\Lambda}$. Except for evanescence we can assume that it is predictable. Adapted processes whose paths are left continuous and have right limits a.s. (caglad) are in particular predictable with a.s. bounded paths. ■

The bond prices and the price deflator, as solutions of SDEs (1) and (3), can be given in the following integrated forms

$$B_i(t) = B_i(0) \exp \left[\int_0^t \left(\mu_i - \frac{1}{2} |\sigma_i|^2 \right) ds + \int_0^t \sigma_i \cdot dW(s) \right],$$

$$\xi(t) = \exp \left[\int_0^t \left(-r - \frac{1}{2} |\lambda|^2 \right) ds - \int_0^t \lambda \cdot dW(s) \right],$$

as an application of Itô‘s Lemma demonstrates. We immediately see that

$$\xi B_i = B_i(0) \exp \left[\int_0^t \left(\mu_i - \frac{1}{2} |\sigma_i|^2 - r - \frac{1}{2} |\lambda|^2 \right) ds + \int_0^t (\sigma_i - \lambda) \cdot dW(s) \right]. \tag{4}$$

Since ξB_i are martingales for $i = 1, ..., n$, this implies the equality

$$\mu_i - r - \frac{1}{2} |\sigma_i|^2 - \frac{1}{2} |\lambda|^2 + \frac{1}{2} |\sigma_i - \lambda|^2 = 0,$$

or equivalently

$$\mu_i = r + \sigma_i \cdot \lambda \ , \tag{5}$$

to hold for $i = 1, ..., n$. To summarize, if the price system in (1) and (2) is arbitrage-free, the dynamics of a price deflator ξ can be written in the form (3), where r and λ satisfy (5). Conversely, if such predictable processes r and λ exist so that (5) holds, then the system is arbitrage-free. The last conclusion requires an integrability condition for ξB_i, such as the one introduced by Novikov.

We have avoided as yet a rigorous definition for the notion of replicability of a claim. At this stage, let us just note that a market is called complete if every measurable claim can be replicated. In advanced mathematical finance literature it is shown that there is an equivalence between completeness of a market and uniqueness of its price deflator. We will use this result to directly define completeness.

Definition 3 *Let the price system B^0 from (1) and (2) be arbitrage-free. The system is then said to be complete, if the price deflator $\xi > 0$ is unique.*

Let us have a closer look at (5) to analyze under which conditions markets are arbitrage-free and complete. To this end we first write (5) as a vector equation with matrix $\sigma \in \mathbb{R}^{n \times m}$ and column vector $\mathbf{1} := (1, ..., 1)^\mathsf{T} \in \mathbb{R}^n$.

$$\mu = \mathbf{1}r + \sigma \cdot \lambda$$

where $\mu \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^m$ and $r \in \mathbb{R}^1$.

Proposition 4 *Suppose that for all t and P-almost all ω with $(t, \omega) \in [0, T] \times \Omega$, the $n \times m$-matrix $\sigma(t, \omega)$ has full rank, $\min(m, n)$. Then with $\mathbf{1} := (1, ..., 1)^\mathsf{T} \in \mathbb{R}^n$ we have*

(i) *For $m \geq n$ the market is arbitrage-free but incomplete.*

(ii) *In the case when $m = n - 1$:*

   (a) *if $\mathbf{1} \notin$ span(σ), the market is arbitrage-free and complete.*

   (b) *if $\mathbf{1} \in$ span(σ) and $\mu \in$ span(σ), the market is arbitrage-free but incomplete.*

   (c) *if $\mathbf{1} \in$ span(σ) and $\mu \notin$ span(σ), the market is not even arbitrage-free.*

(iii) *In the case when $m < n - 1$:*

   (a) *if $\mu \in$ span($\mathbf{1}, \sigma$) and $\mathbf{1} \notin$ span(σ), the market is arbitrage-free and complete.*

   (b) *if $\mu \in$ span($\mathbf{1}, \sigma$) and $\mathbf{1} \in$ span(σ), the market is arbitrage-free but incomplete.*

   (c) *if $\mu \notin$ in span($\mathbf{1}, \sigma$), the market is not even arbitrage-free.*

Proof. Follows directly from definitions (1) and (3) and conditions for the solvability of linear systems and uniqueness of their solutions. ■

The condition "$\mathbf{1} \notin \text{span}(\sigma)$" when $m \leq n-1$ is necessary for completeness in the above proposition. To see this define the extended matrix

$$\widehat{\sigma} := (\mathbf{1}, \sigma) = \begin{pmatrix} 1 & \sigma_{11} & \dots & \sigma_{1m} \\ 1 & \sigma_{21} & \dots & \sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \sigma_{n1} & \dots & \sigma_{nm} \end{pmatrix} \in \mathbb{R}^{n \times (m+1)},$$

and summarize the two postulates, full rank σ and "$\mathbf{1} \notin \text{span}(\sigma)$", both into one requirement by imposing a full rank assumption on $\widehat{\sigma}$. An according proposition can be found in Duffie (1996).

Proposition 5 *If rank($\widehat{\sigma}$)$= m+1$ almost everywhere, then there is at most one state price deflator.*

Note that $m+1 \leq n$.

1.1.2 Arbitrage Condition for the Standard Libor Model

In this section we demonstrate, how the above arbitrage condition for a zerobond process $B \in \mathbb{R}^n$ translates into a LIBOR forward rate process $L \in \mathbb{R}^{n-1}$. In other words, we derive dynamics of LIBOR forward rates from the respective dynamics of an arbitrage-free bond system $B \in \mathbb{R}^n$ solving (1) and satisfying (5).
In full generality we consider a sequence of tenor dates, $0 < T_1 < T_2 < ... < T_n = T < \infty$, corresponding to the maturity dates of the zerobonds, together with the sequence of day-count fractions $\delta_i := T_{i+1} - T_i, i = 1, ..., n-1$. With respect to this tenor structure, let B be such an arbitrage-free system of bonds as in (1) that satisfies (5). Each B_i, for $i = 1, ..., n$, is a random process that *lives* on $[0, T_i]$ and converges towards its face value $B_i(T_i) = 1$ at maturity T_i. This condition indeed implies a restriction on the processes σ_{ij} and μ_i. This does, however, not effect the result of Propostion 4.

A LIBOR forward rate system is defined by

$$L_i(t) := \frac{1}{\delta_i}\left(\frac{B_i(t)}{B_{i+1}(t)} - 1\right), \quad 0 \le t \le T_i, \quad 1 \le i \le n-1. \tag{6}$$

Applying the vector version of Itô's Lemma to (6) we obtain

$$\begin{aligned} dL_i &= \delta_i^{-1}(1+\delta_i L_i)\left[(\mu_i - \mu_{i+1} - \sigma_{i+1}\cdot(\sigma_i - \sigma_{i+1}))dt + (\sigma_i - \sigma_{i+1})\cdot dW\right] \\ &= \delta_i^{-1}(1+\delta_i L_i)(\sigma_i - \sigma_{i+1})\cdot(dW + (\lambda - \sigma_{i+1})dt), \quad 1 \le i < n, \end{aligned}$$

where we used (5) in the last equality. For a convenient notation we introduce LIBOR volatility processes $\gamma_i \in \mathbb{R}^m$ defined by

$$L_i\gamma_i := \delta_i^{-1}(1+\delta_i L_i)(\sigma_i - \sigma_{i+1}), \tag{7}$$

and for $1 < j \le n$ the shifted Brownian motions by

$$dW^{(j)} := dW + (\lambda - \sigma_j)dt. \tag{8}$$

With these definitions we may write for $1 \le i < n$,

$$\begin{aligned} dL_i &= L_i\gamma_i \cdot dW^{(i+1)} \qquad (9) \\ &= -\sum_{j=i+1}^{n-1}\frac{\delta_j L_j L_i}{1+\delta_j L_j}\gamma_i\cdot\gamma_j\,dt + L_i\gamma_i\cdot dW^{(n)}, \quad 0 \le t \le T_i, \end{aligned}$$

where we chose $j = i + 1$ for the shifted Brownian motion in (8).
The last equality follows from

$$\begin{aligned} dW^{(i+1)} &= dW + (\lambda - \sigma_{i+1})dt \\ &= dW^{(n)} - (\lambda - \sigma_n)dt + (\lambda - \sigma_{i+1})dt \\ &= dW^{(n)} + (\sigma_n - \sigma_{i+1})dt \\ &= dW^{(n)} + \sum_{j=i+1}^{n-1} (\sigma_{j+1} - \sigma_j)dt \\ &\overset{(7)}{=} dW^{(n)} - \sum_{j=i+1}^{n-1} \frac{\delta_j L_j \gamma_j}{1 + \delta_j L_j} dt. \end{aligned}$$

A model for forward LIBOR rates whose dynamics are specified by (9) is called a LIBOR model. Note that $\gamma = (\gamma_1, ..., \gamma_{n-1})$ is in general a stochastic process.

A closer look at (9) suggests three questions:

1. Is there a measure under which the above shifted Brownian motion is a *standard* Brownian motion, so that (9) indicates that the forward LIBOR process L_i is a martingales under that measure?

2. Since we will be working in an incomplete market, will this measure be unique?

3. If not, will there arise problems for our intention to price interest rate derivatives?

These questions will be addressed in the next section. Finally, we define what we understand under a LIBOR *market* model.

Definition 6 *A* LIBOR *model where the volatility process* $\gamma(t) = (\gamma_1(t), ..., \gamma_{n-1}(t))^\mathsf{T}$ *is a deterministic function of time is called a* LIBOR *market model.*

1.1.3 Change of Numeraire

In statistics underlying "true" or "real world" measures are estimated from data at first hand. One thereby aims at identifying a measure that suits the data best, according to some criteria. Having found it, further analysis starts from here.

In financial mathematics it can be advantageous to consider the process under a different, yet unknown, measure. It pays off, when desired expressions are easier to evaluate under that measure. A change in measure may be induced by a measurable positive stochastic process, a numeraire process. The fact that bond prices B are of that class is a particular benefit. The following definition is more general in considering an arbitrary numeraire process N.

Definition 7 *Let the system (1) be arbitrage-free and ξ a price deflator such that ξB_i are P-martingales. Let further N be a positive $\mathcal{F}_T$-adapted process such that ξN is also a P-martingale.*

1. *The N-numeraire measure P_N is defined via the Radon-Nikodym derivative*

$$\frac{dP_N}{dP} := \frac{\xi(T)N(T)}{N(0)}, \quad i.e.$$

$$P_N(A) := \int_A \frac{dP_N}{dP} dP,$$

where the random variable $\frac{dP_N}{dP} \in L^1(P)$.

2. *The density process of P_N is defined by*

$$Z_t = E\left(\frac{dP_N}{dP} | \mathcal{F}_t\right) = E_P^{\mathcal{F}_t}\left(\frac{dP_N}{dP}\right) = \frac{\xi(t)N(t)}{N(0)}.$$

Some facts can be concluded.

Remark 8 *Since P_N is equivalent to P, $P_N \backsim P$, we have $\frac{dP_N}{dP} > 0$, P-as. Consequently Z is a positive P-martingale with $Z_0 = 1$ and $Z_T = \frac{dP_N}{dP}$. Note also the correctness of the implication*

$$P(Z_T > 0) = 1 \quad \Rightarrow \quad P(Z_t > 0) = 1, \quad \textit{for all} \quad t \in [0, T],$$

which guarantees that the "inverse" density exists. Let $\frac{dP}{dP_N}$ denote the Radon-Nikodym derivative of P with respect to P_N, then for $\frac{dP}{dP_N} \in L^1(P_N)$ we have

$$\frac{dP}{dP_N} = \left(\frac{dP_N}{dP}\right)^{-1}.$$

The central lemma of this section follows. It will allow to present derivative prices as expectations with respect to different measures.

Lemma 9 *Let $P_N \backsim P$ and $E_P^{\mathcal{F}_t}\left(\frac{dP_N}{dP}\right)$ the corresponding density process. An adapted process X/N is a P_N-martingale, if and only if ξX is a P-martingale.*

Proof. We have for all $0 \leq s \leq t \leq T$.

$$\begin{aligned}
X/N \quad \text{is} \quad P_N\text{-martingale} \quad &\Leftrightarrow \quad E_{P_N}^{\mathcal{F}_s}\left(\frac{X}{N}(t)\right) = \frac{X}{N}(s) \\
&\Leftrightarrow \quad E_{P_N}^{\mathcal{F}_s}\left(\frac{dP}{dP_N}\frac{dP_N}{dP}\frac{X}{N}(t)\right) = \frac{X}{N}(s) \\
&\Leftrightarrow \quad E_P^{\mathcal{F}_s}\left(\frac{dP_N}{dP}\frac{X}{N}(t)\right) = \frac{X}{N}(s) \\
&\Leftrightarrow \quad E_P^{\mathcal{F}_s}\left(\frac{\xi N(t)}{\xi N(s)}\frac{X}{N}(t)\right) = \frac{X}{N}(s) \\
&\Leftrightarrow \quad E_P^{\mathcal{F}_s}\left(\xi N(t)\frac{X}{N}(t)\right) = \xi N(s)\frac{X}{N}(s) \\
&\Leftrightarrow \quad E_P^{\mathcal{F}_s}\left(\xi X(t)\right) = \xi X(s) \\
&\Leftrightarrow \quad \xi X \quad \text{is} \quad P\text{-martingale.}
\end{aligned}$$

Here s took over the role of zero and $\xi(s) \neq 1$ in general. The argument that the conditional density given time s information $\frac{dP_N}{dP} = \frac{\xi N(t)}{\xi N(s)}$ is immediate. Alternatively we may use the definition of a version of the conditional expectation and exchange the column on the right hand side by

$$E_{P_N}\left(\mathbf{1}_A\left(\frac{X}{N}(t)-\frac{X}{N}(s)\right)\right)=0, \quad \text{for all} \quad s\le t,\ A\in\mathcal{F}_s$$

$$E_P\left(\mathbf{1}_A\left(\frac{X}{N}(t)-\frac{X}{N}(s)\right)Z_T\right)=0, \quad \text{for all} \quad s\le t,\ A\in\mathcal{F}_s$$

$$E_P\left(\mathbf{1}_A\left(\frac{XZ}{N}(t)-\frac{XZ}{N}(s)\right)\right)=0, \quad \text{for all} \quad s\le t,\ A\in\mathcal{F}_s$$

$$E_P\left(\mathbf{1}_A\left(X\xi(t)-X\xi(s)\right)\right)=0, \quad \text{for all} \quad s\le t,\ A\in\mathcal{F}_s,$$

to again conclude that ξX is a P-martingale. ■

Of particular interest in the above derivation are the cases where $N=B_i$ for some i. By assumption all processes B_i are positive, adapted and ξB_i are martingales. The constructed, thus existing, B_0-numeraire measure is known as the *risk-neutral-measure* and denoted by $P_0:=P_{B_0}$. Another measure deserving an own name is the *terminal measure* generated by $N=B_n$ and abbreviated $P_n:=P_{B_n}$. It will be utilized frequently in this work.
The question whether there exists a measure such that the shifted Brownian motions in (8) are standard Brownian motions, therefore martingales, can now be answered affirmatively. Pick $P_j:=P_{B_j}$ to see that

$$\begin{aligned}
P_j(W_T^{(j)}\le x) &= \int_A dP_j\,, \quad \text{with} \quad A=\{\omega: W_T^{(j)}(\omega)\le x\}\\
&= \int_A \frac{\xi(T)B_j(T)}{B_j(0)}dP\\
&\overset{(4)}{=} \int_A \exp\left[-\frac{1}{2}\int_0^T|\sigma_j-\lambda|^2ds+\int_0^T(\sigma_j-\lambda)\cdot dW(s)\right]dP\\
&= P(W_T\le x),
\end{aligned}$$

where the last equality follows by a version of Girsanov's Theorem, see Glassermann (2004, appendix B). We obtain

$$\frac{dP_j}{dP}=\exp\left[-\frac{1}{2}\int_0^T|\sigma_j-\lambda|^2ds+\int_0^T(\sigma_j-\lambda)\cdot dW(s)\right].$$

Consequently with $j = i + 1$ in (9), the forward LIBOR process L_i is indeed a martingal under P_{i+1}.

A major application of the above concept is the following. Suppose the payoff function C multiplied by our state price deflator is a martingale and let its arbitrage-free price be given by

$$C_t = \frac{1}{\xi(t)} E_P^{\mathcal{F}_t}(\xi(T)C_T).$$

Then an alternative way to compute its value is, to change measure and evaluate

$$C_t = P_j(t) E_{P_j}^{\mathcal{F}_t}\left(\frac{C_T}{P_j(T)}\right).$$

By lemma (9), with $X = C$ and $N = P_j$, the values must coincide.

1.1.4 Replicability and Derivative Pricing

We have seen that, as long as ξC is a martingale, the price of a derivative can be determined conveniently by evaluation of an expectation

$$C_t = \frac{1}{\xi(t)} E_P^{\mathcal{F}_t}\left(\xi(T)C_T\right).$$

Unfortunately, this martingale property can not always be assumed to hold. It can be concluded, however, for so called replicable claims. To demonstrate this result is the task of this section.

Before proving an invariance property, we first give a definition of replicability.

Definition 10 *Self-Financing Trading Strategy & Replicable Claims*

1. *A trading strategy in our bond market B is an $\mathbf{R}^n$-valued, $\mathcal{F}$-predictable and B-integrable process φ. If φ with corresponding value process*

$$V^\varphi := \varphi \cdot B = \sum_{i=1}^{n} \varphi_i B_i$$

satisfies the self-financing condition

$$V^\varphi(t) = V^\varphi(0) + \int_0^t \varphi(u) \cdot dB(u), \quad 0 \le t \le T, \tag{10}$$

we speak of a Self-Financing Trading Strategy (SFTS) in the market B.

2. *A measurable claim C_T is called replicable, if there exists an SFTS φ such that*

$$C_T = \varphi(T) \cdot B(T) = \varphi(0) \cdot B(0) + \int_0^T \varphi(u) \cdot dB(u).$$

Note that replicable claims form a linear subspace in all measurable claims. The following lemma illustrates that self-financing trading strategies satisfy an invariance condition with respect to certain semimartingales ξ.

Lemma 11 *Let φ be a self-financing trading strategy in the system B, i.e. (10) holds. For any positive continuous adapted semimartingale ξ with respect to which φ is ξ-integrable, we have*

$$\varphi(T) \cdot \xi(T)B(T) = \varphi(0) \cdot \xi(0)B(0) + \int_0^T \varphi(u) \cdot d(\xi(u)B(u))\,.$$

Consequently φ is also a self-financing trading strategy in ξB.

Proof. The SFTS condition (10) implies that the value process $\varphi \cdot B$ is a continuous semimartingale, because $d(\varphi \cdot B) = \varphi \cdot dB$. By this same property, Itô's product rule $d(XY) = YdX + XdY + dXdY$, and recalling that $\xi \in \mathbb{R}^1$ we obtain

$$\begin{aligned}
d(\varphi \cdot (\xi B)) &= d(\xi(\varphi \cdot B)) \\
&= d\xi(\varphi \cdot B) + \xi d(\varphi \cdot B) + d\xi d(\varphi \cdot B) \\
&= \varphi \cdot (d\xi B) + \xi \varphi \cdot dB + d\xi \varphi \cdot dB \\
&= \varphi \cdot (d\xi B) + \varphi \cdot (\xi dB) + \varphi \cdot (d\xi dB) \\
&= \varphi \cdot (d\xi B + \xi dB + d\xi dB) \\
&= \varphi \cdot d(\xi B).
\end{aligned}$$

■

We now show the forementioned result that the deflated process ξC is indeed a martingale, if C is a replicable claim.

Proposition 12 *If C_T is a replicable claim in an arbitrage-free system (1) then ξC is a martingale.*

Proof. Let φ be a self-financing trading strategy which replicates C_T. Consequently

$$C_T = \varphi(T) \cdot B(T) = \varphi(0) \cdot B(0) + \int_0^T \varphi(u) \cdot dB(u).$$

Recall that in an arbitrage-free market there exists a price deflator ξ such that ξB_i are martingales for all $i = 1, .., n$. By lemma (11) we conclude

$$\begin{aligned}\xi(T)C_T &= \xi(T)\varphi(T) \cdot B(T) = \varphi(T) \cdot (\xi(T)B(T)) \\ &= \varphi(0) \cdot (\xi(0)B(0)) + \int_0^T \varphi(u) \cdot d(\xi(u)B(u))\,.\end{aligned}$$

The last equality holds, because φ is also a self-financing trading strategy in ξB. Taking conditional expectations over left and right hand side yields

$$\begin{aligned}E^{\mathcal{F}_t}(\xi(T)C_T) &= \varphi(0) \cdot (\xi(0)B(0)) + \int_0^t \varphi(u) \cdot d(\xi(u)B(u)) \\ &= \varphi(t) \cdot (\xi(t)B(t)) \\ &= \xi(t)\varphi(t) \cdot B(t) \\ &= \xi(t)C_t\,.\end{aligned}$$

The first equality follows from the martingale property of the process

$$\int_0^t \varphi \cdot d(\xi B).$$

As it is an Itô integral with respect to the martingale ξB. The second is due to Lemma (11) and the last follows, since C_t is replicated by ξ. ■

We obtained the designated martingale property for ξC which permits to evaluate derivative prices by

$$C_t = \varphi(t) \cdot B(t) = \frac{1}{\xi(t)} E_P^{\mathcal{F}_t}(\xi(T)C_T). \tag{11}$$

From (11) and the *Law-of-One-Price*, which is required to hold for replicable claims in arbitrage-free markets, two remarkable facts can be deduced:

1. If $\widetilde{\varphi}$ is another SFTS replicating C_T, we have $\varphi(t) \cdot B(t) = \widetilde{\varphi}(t) \cdot B(t)$ for all $t \in [0, T]$. Thus both strategies generate the same price process.

2. In an incomplete market the price deflator ξ is not unique. Nevertheless, the right hand side of (11) must not depend on ξ – because the left side does not. We therefore must have

$$\frac{1}{\xi_1(t)} E_{P_1}^{\mathcal{F}_t}(\xi_1(T)C_T) = \frac{1}{\xi_2(t)} E_{P_2}^{\mathcal{F}_t}(\xi_2(T)C_T).$$

As a conclusion to this section, we may state that the two open questions from Section 1.1.2 are now affirmatively answered. Although the respective measure is not unique in an incomplete market, it will not pose a problem as long as we restrict ourselves to replicable claims. The price process remains then unique. For non-replicable claims only an interval can be specified in which the arbitrage-free price has to lie.

It is a well-known fact that in a complete market any measurable claim C_T can be replicated. Moreover, replicability of all claims can be shown to be sufficient for completeness of the market. In many text books replicability is therefore used as the defining property of complete markets. However, in this work we will introduce continuous stochastic volatility models in markets that are profoundly *incomplete*. It will thus be interesting to know, which class of claims will remain replicable and will therefore allow the price evaluation in the form illustrated above.

1.1.5 Existence, Uniqueness and Replicability

Now that all open questions from Section 1.1.2 are answered, let us turn to one that evolved implicitly and which is of significant theoretical importance. We usually start with modeling the LIBOR process directly, that is we specify its volatility structure $\gamma_i(t, L)$, for $1 \leq i < n$. A natural question arising is, whether a corresponding arbitrage-free bond process B exists at all. A fundamental result in this direction was proved by Jamshidian (1997). His work provides a justification for the approach taken by practitioners. The following proposition gives an affirmative answer, which is sufficient for our needs.

Proposition 13 ***Existence of the* LIBOR *forward process***
Given a measurable, bounded and locally Lipschitz volatility structure $\gamma_i(t, L)$, $1 \leq i < n$, there exists an arbitrage-free system B of bond prices (1) satisfying $B_i(T_i) = 1$ and a price deflator ξ, such that (9), with $\gamma_i = \gamma_i(t, L)$, has a unique solution for which (6) holds.

Proof. See Jamshidian (1997), Theorems 5.3 and 7.1. ■

Another fundamental question of interest is regarding the replicability of a derivative security. If it is, traders in a bank call it hedgeable, meanwhile applied mathematicians and business researchers refer to it as replicable. Pure mathematicians and probabilists, however, associate measurability conditions with it. In any case, it is good to speak the language of all three of them. From a probabilistic perspective, finding a sufficient condition comes essentially down to clarifying with respect to which σ-field measurability is meant. In the following theorem it is crucial to observe that the filtration generated by $W^{(n)}$ – and no further sources – is considered.

Theorem 14 ***Hedging*** LIBOR ***derivative Claims***
Consider an arbitrage-free bond system (1) and the corresponding LIBOR *forward process (6) with dynamics given by (9). Suppose that $W^{(n)}$ is an m-dimensional Brownian motion with $m < n$ and the volatility process γ has constant rank m and is adapted to $\mathcal{F}^n$, the filtration generated by $W^{(n)}$. Then any claim C_T such that $C_T/B_n(T)$ is measurable with respect to $\mathcal{F}^n_T$ can be priced and hedged by a self-financing strategy φ in the bond system B. The arbitrage-free price of the claim is given by*

$$C_t := \varphi(t) \cdot B(t) = B_n(t) E_n \left[\frac{C_T}{B_n(T)} | \mathcal{F}^n_t \right] = B_n(t) E_n^{\mathcal{F}^n_t} \left[\frac{C_T}{B_n(T)} \right]. \qquad (12)$$

Since only the distribution of the LIBOR process is relevant, we may assume $m < n$ and that $\gamma_i(t, L)$ has full rank m. Indeed, for $\widehat{m} \geq n$ and arbitrary $\widehat{\gamma}$, it is possible to construct a LIBOR process with the same distribution given by (9) with $m < n$ and some different full m-rank γ. As such, the assumptions above do not constitute a restriction. Of course, an incomplete market will remain incomplete and by no means be suddenly turned complete.
In fact it is remarkable that completeness of the bond system B is not an issue in theorem (14). As we will see, its generality will make it applicable to almost all LIBOR derivatives in practice.

The main achievement of the theorem consists in providing us with a condition under which a claim is replicable. In incomplete markets it is in general difficult to characterize the replicable claims which are perceived to consist of a rather "meager" set among all possible measurable ones. Only the affine-linear combinations of the market securities, here the zerobonds, are easy to identify as such. All the more surprising appears to be the fact that this condition is merely a measurability condition, a property usually associated as weak or easy to satisfy. In fact this is not the case here. It is a very characteristic and appropriate measurability that enables us to find an SFTS. We elaborate further on this qualities in the next example.

Proof. (of Theorem 14, see Jamshidian 1997, Theorem 5.2.) Since the forward LIBOR process L is $\mathcal{F}^n$-adapted, all B_i/B_n are $\mathcal{F}^n$-adapted martingales under the terminal measure $W^{(n)}$. The martingale representation theorem thus implies the existence of an $\mathcal{F}^n$-adapted (matrix) process $\nu \in \mathbb{R}^{(n-1)\times m}$ such that

$$d(\frac{B}{B_n}) = \nu dW^{(n)}$$

where $(B/B_n) := (B_i/B_n)_{1\leq i<n} \in \mathbb{R}^{n-1}$. The matrix (process) ν must be of full rank m also, for otherwise at least one component in $W^{(n)}$ could be exchanged by the others which would imply that γ was not of full rank m. The pseudo-inverse of ν is now well-defined by $\psi = \nu(\nu^\top\nu)^{-1} \in \mathbb{R}^{(n-1)\times m}$ which is again $\mathcal{F}^n$-adapted.

Observe now that by the measurability assumption

$$c_t = E_n\left(\frac{C_T}{B_n(T)}|\mathcal{F}^n_t\right)$$

is an $(\mathcal{F}^n,\ P_n)$-martingale. The martingale property can be easily shown by the tower law for conditional expectations. Again by the representation theorem we obtain an adapted process $\beta \in \mathbb{R}^m$ such that

$$c_t = c_0 + \int_0^t \beta \cdot dW^{(n)}. \tag{13}$$

Set $\theta = \psi\beta\ \in \mathbb{R}^{n-1}$ and observe

$$\begin{aligned}
\theta^\top d\left(\frac{B}{B_n}\right) &= (\psi\beta)^\top d\left(\frac{B}{B_n}\right)\\
&= \beta^\top\psi^\top d\left(\frac{B}{B_n}\right)\\
&= \beta^\top(\nu(\nu^\top\nu)^{-1})^\top\nu dW^{(n)}\\
&= \beta^\top(\nu^\top\nu)^{-1}\nu^\top\nu\, dW^{(n)}\\
&= \beta^\top dW^{(n)}\\
&= dc.
\end{aligned}$$

From this and the definition of $\varphi \in \mathbb{R}^n$ as an extension of θ by defining

$$\varphi_i := \theta_i \quad \text{for} \quad 1 \leq i < n$$

and

$$\varphi_n := c - \theta^\top\left(\frac{B}{B_n}\right),$$

we obtain with $(B^*/B_n) := (B_i/B_n)_{1\le i\le n} \in \mathbb{R}^n$ that

$$\begin{aligned} d\left(\frac{\varphi^\mathsf{T} B}{B_n}\right) = d\left(\varphi^\mathsf{T}\left(\frac{B^*}{B_n}\right)\right) &= d\left(\theta^\mathsf{T}\left(\frac{B}{B_n}\right) + \varphi_n\right) \\ &= dc \\ &= \theta^\mathsf{T} d\left(\frac{B}{B_n}\right) \\ &= \sum_{i=1}^{n-1} \varphi_i d\left(\frac{B_i}{B_n}\right) + \varphi_n d(1) \\ &= \sum_{i=1}^{n} \varphi_i d\left(\frac{B_i}{B_n}\right) \\ &= \varphi^\mathsf{T} d\left(\frac{B^*}{B_n}\right), \end{aligned}$$

where we recall that $B \in \mathbb{R}^n$. Hence the pair φ is an SFTS for $B_i/B_n, 1 \le i \le n$. In lemma (11) choose $\xi = B_n$ and conclude that (φ, B) is also a self-financing portfolio. Since $\varphi B = B_n c$ we have in particular that $\varphi_T B_T = B_n(T)c_T = C_T$. Recall that $c = (\varphi B)/B_n$ is an $(\mathcal{F}^n,\, P_n)$-martingale to finish the proof.
Note finally that in our setting of continuous functions on a compact set, $[0, T]$, and the previously mentioned restriction on pure martingales rather than local martingales, we immediately get the finiteness of all integrals and thus do not have to show it explicitly. ■

Example 15 *As indicated in the proof, we have*

$$\sigma(B/B_n) = \sigma((B_i/B_n)_{1\le i<n}) \subset \mathcal{F}^n = \sigma(W^{(n)}) \subset \sigma(W) = \sigma(B_1, ..., B_n).$$

Consequently B_n is $\sigma(W)$-measurable, but not necessarily $\mathcal{F}^n$-measurable. If it would be $\mathcal{F}^n$-measurable, the theorem would imply the existence of an SFTS φ that would have to satisfy

$$B_n^2 = \varphi \cdot B.$$

In an incomplete market a solution is $\varphi_n(t) = B_n(t)$ and $\varphi_j(t) = 0$, for $j = 1, ..., n-1$, which is not self-financing.
In a complete market there are always self-financing solutions, but in that case B_n could also be $\mathcal{F}^n$-measurable.

On the other hand, by the theorem, for $C_T = f(L_{n-1}(T))B_n(T)$ *with any continuous* f *there must exist an SFTS replicating the claim* C_T*, since* $f(L_{n-1}(T))$ *is* $\mathcal{F}^n$*-measurable.*

The position of $\mathcal{F}^n$ between $\sigma(B/B_n)$ and $\sigma(B_1, ..., B_n)$ gives an indication how "complete" the market is. If $\mathcal{F}^n$ coincides with $\sigma(B_1, ..., B_n)$, i.e. $\mathcal{F}^n = \sigma(B_1, ..., B_n)$, no restriction is imposed on C_T and the market must in fact be complete, since now every claim is replicable. On the other hand, if $\mathcal{F}^n$ coincides with $\sigma((B_i/B_n)_{1\leq i<n}) = \sigma((B_i/B_j)_{1\leq i<j<n}) = \sigma((L_i)_{1\leq i<n})$, then the measurability condition on C_T essentially amounts to $C_T/B_n(T)$ being measurable with respect to the σ-algebra generated by the $(L_i)_{1\leq i<n}$. In an incomplete market this restriction can be substantial.

1.2 The Standard Libor Model

1.2.1 An Alternative Definition for the Standard Model

The standard LIBOR market model is usually defined in terms of a vector-valued standard Brownian motion. Before we analyze its parametrization in detail, we show that an alternative definition with correlated one-dimensional Brownian components can be obtained. Both prove to be equivalent. On some occasions the latter may be useful.

To remain consistent with the notation in the literature, we will not introduce other greek characters. We rather mark the familiar ones by an upper case s to distinguish them from the earlier. Thus σ^s, ρ^s and μ^s denote the parameters of the standard model given in the form

$$\frac{dL_i(t)}{L_i(t)} = \mu_i^s(t)dt + \sigma_i^s(t)dZ_i(t), \qquad i = 1, ..., n-1, \tag{14}$$

where $Z_i : (t, \omega) \to \mathbb{R}^1$ are *dependent* Brownian motions with

$$dZ_i(t)dZ_j(t) = \rho_{ij}^s(t)dt. \tag{15}$$

Here $\sigma_i^s : [0, T] \to \mathbb{R}^1$ and $\rho_{ij}^s : [0, T] \to \mathbb{R}^1$ denote *scalar* deterministic functions of time, whereas the drift μ_i^s is an $\mathcal{F}_t$-adapted stochastic process. The dependence on the argument t will be suppressed for notational convenience. In the setting (14) every LIBOR is perturbed by exactly one source of uncertainty. Nevertheless it is equivalent to the one given before

$$\frac{dL_i}{L_i} = -\sum_{j=i+1}^{n-1} \frac{\delta_j L_j \gamma_i \cdot \gamma_j}{(1 + \delta_j L_j)} dt + \gamma_i \cdot dW, \tag{16}$$

where $\gamma_i \in \mathbb{R}^m$ and $W \in \mathbb{R}^m$. Note that in this version the components of W are *independent*.

Lemma 16 *The two stochastic differential equations (14) and (16) have the same solutions.*

Proof. To see the equivalence consider the (generalized) Cholesky decomposition of the correlation matrix $R = (\rho_{ij}^s)_{i,j=1,...,n-1}$ of rank m

$$\rho_{ij}^s = \sum_{k=1}^{m} f_{i,k} f_{j,k} = f_i \cdot f_j,$$

where $f_i \in \mathbb{R}^m$ for $i = 1, ..., n-1$. As a result, there exists an $(n-1) \times m$ -matrix $F = (f_{i,j})$ such that

$$dZ_i = \sum_{k=1}^{m} f_{i,k} dW_k = f_i \cdot dW$$

or simply

$$dZ = F\, dW, \tag{17}$$

with $Z := (Z_1, ...Z_{n-1})^\mathsf{T}$ and $W = (W_1, ..., W_m)^\mathsf{T}$.

The equivalence is established by setting

$$\gamma_i := \sigma_i^s f_i\,.$$

Since R is a correlation matrix the rows f_i , $i = 1, ..., n-1$, are in fact unit vectors. Therefore

$$|\gamma_i| = \sigma_i^s\,.$$

Moreover, as required, we observe

$$\begin{aligned}
& \rho_{ij}^s dt = dZ_i\, dZ_j \\
\Leftrightarrow \qquad & \sigma_i^s \sigma_j^s\, \rho_{ij}^s\, dt = \sigma_i^s dZ_i\, \sigma_j^s dZ_j \\
& = \sigma_i^s f_i \cdot dW\ \ \sigma_j^s f_j \cdot dW \\
& = \gamma_i \cdot dW\, \gamma_j \cdot dW \\
& = \left(\sum_{k=1}^{m} \gamma_{ik} dW_k \right) \left(\sum_{k=1}^{m} \gamma_{jk} dW_k \right) \\
& = \sum_{k=1}^{m} \gamma_{ik} \gamma_{jk} \left(dW_k \right)^2 + \sum_{l \neq k}^{m} \gamma_{il} \gamma_{jk} dW_l dW_k \\
& = \gamma_i \cdot \gamma_j\, dt\,.
\end{aligned}$$

In the last equality the second sum vanishes due to independence of the Brownian components.

These arguments justify an alternative definition for the instantaneous correlation as

$$\rho_{ij}^s = \frac{\gamma_i \cdot \gamma_j}{|\gamma_i||\gamma_j|}. \tag{18}$$

For the drift μ_i^s, directly set

$$\mu_i^s = -\sum_{j=i+1}^{n-1} \frac{\delta_j L_j \gamma_i \cdot \gamma_j}{(1+\delta_j L_j)}.$$

■

Remark 17 *In the pertinent literature the instantaneous correlation process ρ_{ij}^s is sometimes defined by a covariance analogon. The following other notations exist*

$$cov\left(\frac{dL_i}{L_i}, \frac{dL_j}{L_j}\right) = \gamma_i \cdot \gamma_j dt$$

or in a logarithmic exposition

$$cov(d\ln L_i, d\ln L_j) = \gamma_i \cdot \gamma_j dt.$$

The last two equations are understood to hold in linear approximation, as all other terms are of higher order.

1.2.2 Parametrization of Scalar Volatility Function and Correlation Matrix

General Aspects The parametrization and calibration of the factor loadings γ_i is a main issue in the work with LIBOR market models. These loadings can be arbitrary processes in *general* LIBOR models. In LIBOR *market* models, however, they are, by definition, required to be deterministic. The results of the last section explain the self-evident definition for the scalar volatility process σ^s of the forward LIBOR rates L_i through

$$\sigma_i^s(t) := |\gamma_i(t)| = \sqrt{\sum_{k=1}^{m} \gamma_{ik}^2(t)}, \qquad 0 \leq t \leq T,\, 1 \leq i < n. \tag{19}$$

A correlation structure ρ^s is defined by

$$\rho_{ij}^s(t) := \frac{\gamma_i(t) \cdot \gamma_j(t)}{|\gamma_i(t)||\gamma_j(t)|} \qquad \text{for} \quad 0 \leq t \leq \min(T_i, T_j),\, 1 \leq i, j < n. \tag{20}$$

Both functions are refered to as the *instantaneous* volatility and correlation, as opposed to *terminal* quantities. Some properties of these concepts are given immediately after the following proposition, which illustrates an invariance feature that is important for the subsequent analysis. It can be deduced directly from the above definition.

If γ_i's are transformed by an orthogonal matrix Q,

$$\widetilde{\gamma}_{ik} := \sum_{l=1}^{m} \gamma_{il} Q_{lk} \,, \tag{21}$$

neither the scalar volatility nor the correlation structure will be affected. Consequently the volatility structure γ is not uniquely determined by the structures σ^s and ρ^s alone. It is not difficult to show, see Proposition 2.1 in Schoenmakers (2005), that the following proposition holds.

Proposition 18 *If two volatility structures γ and $\widetilde{\gamma}$ are related by (21), their corresponding scalar volatility and correlation structures coincide and their corresponding* LIBOR *processes are identical in distribution. In this sense, γ and $\widetilde{\gamma}$ can be regarded as equivalent volatility structures.*

Since prices of derivative contracts are expectations, only the distributions of the involved stochastic variables are relevant. Proposition 18 therefore implies that we have a certain freedom to choose the γ structure according to our needs. Furthermore, we conclude that σ^s and ρ^s are the basic economic objects that fully specify the LIBOR market model. Once these objects are determined, we can work with any convenient deterministic volatility structure γ that satisfies (19) and (20). We will see later that triangular matrices are a possible choice for (γ_{ij}) that is useful for our purposes.

The parametrization of the above functions is necessary, because at any given time the two product groups, caps and swaptions, provide only a finite number of prices of contracts terminating at the tenor dates $T_i, i = 1, ..., n-1$. The implied parameters are averages over these parametrized functions up to these times.

Implied volatilities are usually determined through Black's formula, the reason we denote them with a superscript B. In the caplet market we have for σ_i^s

$$(\sigma_i^B)^2 = \frac{1}{T_i - t_0} \int_{t_0}^{T_i} (\sigma_i^s(s))^2 ds = \frac{1}{T_i - t_0} \int_{t_0}^{T_i} |\gamma_i|^2(s) ds. \tag{22}$$

This relation is obtained by a simple Itô argument on the geometric Brownian SDE with a following integration.

We are not that fortunate with respect to the swaptions. All we can hope to receive in their case are *approximations* to the averages over ρ^s at the grid points. Fortunately, on the other hand, these turn out to work quite well. Denoting the tenor subscript of swap start and maturity with p and q, respectively, we will consider

$$\begin{aligned}(\sigma^B_{p,q})^2 &\approx \sum_{i,j=p}^{q-1} \frac{\mathcal{K}(i,j)}{T_p - t_0} \int_{t_0}^{T_p} (\gamma_i \cdot \gamma_j)(s) ds \\ &= \sum_{i,j=p}^{q-1} \frac{\mathcal{K}(i,j)}{T_p - t_0} \int_{t_0}^{T_p} \rho_{ij}(s) |\gamma_i| |\gamma_j| (s) ds\end{aligned} \tag{23}$$

where the constant $\mathcal{K}(i,j)$ is given

$$\mathcal{K}(i,j) = \frac{w_i^{p,q} w_j^{p,q} L_i L_j}{\sum_{k,l=p}^{q-1} w_k^{p,q} w_l^{p,q} L_k L_l}.$$

The weight factors $w_i^{p,q} = \delta_i B_{i+1} / \sum_{k=p}^{q-1} \delta_k B_{k+1}$ behave in practice rather smooth compared to the L_i processes. In a good approximation, it is therefore justified to neglect their covariation process, as shown in Schoenmakers & Coffey (1999, equation 54). Small increments of $d \ln w_i^{p,q}$ are assumed to be negligible compared to increments $d \ln L_i$.

While (22) is an equation that can be derived easily, the last relation (23) requires a more detailed explanation. We will provide this in the next section. We mention it already here, because we wish to make some comments on the often utilized *constant factor loading* assumption, which assumes γ_i to be constant vectors. Due to the ease of implementation associated with it, in practice models regularly adopt this assumption. However there are disadvantages to be kept in mind.

Constant Factor Loadings In principle a market model with *constant* loadings contains just enough degrees of freedom to be calibrated to a complete system of caplet and swaption prices or volatilities on the given tenor structure. The reason for this is that the inner product $\gamma_i \cdot \gamma_j$ fully determines the model, as can be seen from (19) and (20). Assuming time independence, we can immediately deduce from (22) and (23), where integration is now obsolete, that the number of given parameters – all possible inner products – coincides with the number of given volatilities. A simple counting argument shows that we have $n(n-1)/2$ of both. Apart from orthogonal transformations of γ, we therefore have a "*unique*" solution under this assumption.

An interesting perspective to look at it is, to observe that the norms $|\gamma_i|^2 = \gamma_i \cdot \gamma_i$ are determined by the market caplet volatilities of which we have exactly

$n-1$. However, the individual components of the γ_i, the $\gamma_{ik}, k = 1, ..., m$, which contain the information about the correlation of the different LIBOR forwards, cannot be recovered from the caplet prices alone. The swaption prices, on the other hand, do significantly depend on the correlation structure and thus constitute plausible candidates for the calibration. Naturally, we come to the same conclusion that we have just as much as we need.

Unfortunately, there are various reasons why a time-independent γ assumption should be avoided.

1. The identification of the parameters may cause stability problems.
2. Most importantly, time-independent σ_i^s are considered unrealistic as empirical LIBOR volatilities decrease with time-to-maturity.
3. Calibration of a full factor model is unpopular among practitioners, who prefer a faster evaluation of their portfolio, especially in regard of the fact that 90% of yield curve movements seem to be explainable with only three factors. Tempting as it is to restrict to less factors, one has then to deal with the problem of a lower rank matrix. It is not immediate to know which prices to ignore.

In a first step to a more general parametrization, suppose that the scalar volatilities σ_i^s and correlations ρ_{ij}^s given by (19) and (20) are piecewise constant functions of time. The scalar volatilities σ_i^s will provide already $n(n-1)/2$ unknown variables, not to mention how much the correlations will contribute. In this case the model (16) is therefore over-determined.

The Compromise In the last paragraph we listed well-founded arguments why the constant factor loading assumption should be avoided. A natural remedy is to consider time-dependent instantaneous volatilities and correlations. Apart from the necessary relations (19) and (20), we would like the resulting functions to have some further features. These are necessary to guarantee consistency with empirically observed economical realities. Some facts:

1. Low factor market models calibrated to prices tend to imply unrealistic instantaneous correlations between different LIBOR forwards. See Schoenmakers (2005) for examples.
2. Calibration of a full factor model, however, tends to be unstable due to the large parameter space dimension.

3. The shape of scalar volatility functions inhibits a hump in short maturities.

4. Towards maturity the functions behave similar for all forwards, thus time homogeneity is an issue.

5. The further LIBORS are apart, the smaller is their absolute correlation, i.e.: $\rho_{ij} \geq \rho_{ik}$, for $|k-i| > |j-i|$.

6. Correlation of adjacent LIBORS increases with maturity, i.e.: $\rho_{ij} \geq \rho_{kl}$ if $|j-i| = |k-l|$ but $k < i$ or $l < j$.

The first two observations concern the dimension of γ_i and involve a modeling problem. Point 2 can be accounted for by regularizing restrictions.

The last four economical facts comprise requirements on the functional form of the instantaneous scalar volatility and matrix correlation. Let us foreclose that unfortunately we will not succeed to meet all points. But if we can not satisfy all these characteristics, the problem remains to identify an *acceptable compromise* between possible solutions to these findings.

Here is what we choose to do:

a.) We work in a full factor model and find cure in regularization. That covers point 1 and 2 above.

b.) Good news regarding the hump-shapedness in point 3. The function g, stated below, will ensure it.

c.) Time homogeneity will not be obtained in its purest form, but we will be close to it, so "quasi" time homogeneous. With the help of constants c_i we will achieve point 4. approximately.

d.) Observations 5 and 6 can be accomplished by a parametrization of the correlation matrix proposed by Schoenmakers (2005). We will supply it below.

In a first approach to find time homogeneous functions consider

$$\begin{aligned} \sigma_i^s(t) &= \qquad c_i g(T_i - t), \quad 0 \leq t \leq T_i\,, \quad 1 \leq i < n\,, \quad c_i > 0\,, \\ \rho_{ij}^s(t) &= \varrho(T_i - t, T_j - t), \quad 0 \leq t \leq \min(T_i, T_j)\,, \quad 1 \leq i, j < n\,, \end{aligned} \tag{24}$$

for a non-negative function $g : [0, \infty) \to \mathbb{R}_+$ and a function $\varrho : [0, \infty) \times [0, \infty) \to [-1, 1]$, which satisfies the usual conditions of a correlation function.

Such ρ^s_{ij} are immediately seen to be time homogeneous. The c_i in (24), however, destroy the pure time homogeneity feature, unless they are all the same, that is $c_i \equiv c$. Their diversity is responsible for the add-on "quasi" in our structure. They will play a crucial role in the calibration of instantaneous volatility to individual caplets in (22). If g is choosen

$$g(s) = g_{a,b,g_\infty}(s) := g_\infty + (1 - g_\infty + as)e^{-bs} \ , \quad a, b, g_\infty > 0$$

a hump-shaped form is obtained.

In a second step, let ρ be a fixed correlation matrix of rank $n-1$ and consider an arbitrary set of decomposing unit vectors $\{\mathbf{e}_i \in \mathbb{R}^{n-1}\, , \, 1 \le i < n\}$,

$$\rho_{kl} = \mathbf{e}_k \cdot \mathbf{e}_l \, .$$

Lemma 19 *Define* $m(t) := \min\{m \in \mathbb{N}_0 : T_{m+1} \ge t\}$ *and set for* $0 \le t \le T_i$ *and* $1 \le i < n$,

$$\gamma_i(t) = c_i g(T_i - t)\mathbf{e}_{i-m(t)} \, .$$

With this setting we obtain functions σ^s_i *and* ρ^s_{ij} *of the form required in (24).*

Proof. Since $\mathbf{e}_{i-m(t)}$ are unit vectors, we have

$$\sigma^s_i(t) = |\gamma_i(t)| = c_i g(T_i - t)|\mathbf{e}_{i-m(t)}| = c_i g(T_i - t)$$

and also

$$\begin{aligned} \rho^s_{ij}(t) &= \frac{\gamma_i(t) \cdot \gamma_j(t)}{|\gamma_i(t)||\gamma_j(t)|} = \frac{c_i g(T_i - t)\mathbf{e}_{i-m(t)} \cdot c_j g(T_j - t)\mathbf{e}_{j-m(t)}}{|c_i g(T_i - t)\mathbf{e}_{i-m(t)}||c_j g(T_j - t)\mathbf{e}_{j-m(t)}|} \\ &= \mathbf{e}_{i-m(t)} \cdot \mathbf{e}_{j-m(t)} \end{aligned}$$

A function of the form $\mathbf{e}_{i-m(t)}$ is easily seen to be time homogeneous for $t \le T_i$. ■

Remark 20 *Schoenmakers (2005) shows that correlation matrices of the form*

$$\rho_{kl} = \frac{\min(b_k\, , \, b_l)}{\max(b_k\, , \, b_l)}$$

with $1 \le k, l < n$*, for a strictly increasing sequence* $b = (b_1, ..., b_{n-1})$ *which in addition offers that* b_k/b_{k+1} *is strictly increasing, are finitely decomposable. Such correlation matrices are of particular interest for a parametrization, since they offer features 5 and 6, as can be shown easily.*

The calibration steps are:

1. In the context of an LS-optimization, find three parameters a, b, g_∞ such that the market caplet volatilities in (22) are best possibly explained. Choose c_i to make the fit exact. Caplets are then priced precisely.

2. Given a parametrized form for ρ, see Schoenmakers (2005, pp 43), optimize over parameters to obtain minimal root mean square (RMS) compared to swaption data given by (23). Take the resulting ρ and Clolesky decompose it, as indicated in the paragraph before Lemma 19.

The first (*Cap*-) part of the calibration is straightforward, because in a LIBOR market model caplets are assumed lognormally distributed and thus the Black & Scholes prices are recovered. Not so in the second (*Swaption*-) part. Here we need an approximating equality of the kind given in (23). The following section is devoted to this problem.

1.3 Approximation of Swaptions in a Libor Market Model

It is a well known fact that caps and swaptions cannot both be priced simultaneously in a consistent way in a LIBOR market model. For the latter a *Swap* market model would be more appropriate, if only swaptions needed to be considered in a specific analysis, as for example in the pricing of Bermudan Swaptions. Whenever individual LIBOR forwards are of concern, however, the LIBOR market model is the better choice. A possible compromise is to price both product groups in a LIBOR market model, while introducing approximating formulas for the swaptions involved. A formula frequently used by practitioners when calibrating the model to swaption volatilities is

$$S_{p,q}^2|\sigma_{p,q}|^2 \approx \sum_{i,j=p}^{q-1} w_i w_j L_i L_j |\gamma_i||\gamma_j|\rho_{ij}^s. \tag{25}$$

As before the correlation matrix ρ^s is defined by $\rho_{ij}^s := \gamma_i \cdot \gamma_j / |\gamma_i||\gamma_j|$. The notation, particularly regarding the definition of the one-dimensional process $|\sigma_{p,q}|^2$, will be explained in the following two sketches of different proofs. While the first introduces rather heuristical arguments and bracket calculus, the second is more rigorous and thus, not surprisingly, more technical. Both approaches are hoped to bring some insight as to why the formula is reasonable.
The derivations will be considering only *standard* swaps, where the number of LIBOR fixings and settlement dates on the variable side coincide with those on the fixed side. In practice these will differ in every country. In the Euro area annual fixed rates against semi-annual variable LIBORS are usual, meanwhile in US, UK and Yen markets semi-annual fixed against quarterly variable rates are common.

A Heuristic Argument We consider $[T_p, T_q]$-swaps, starting at T_p and maturing at T_q for $1 \leq p < q \leq n$, with respect to a usual tenor structure T_j ; $j = 1, ..., n$ with equidistant intervals δ. The swap rate $S_{p,q}$ at time $t < T_p$ is then given by

$$S_{p,q}(t) = \frac{B_p(t) - B_q(t)}{\sum_{k=p+1}^{q} \delta B_k(t)} = \frac{\sum_{i=p+1}^{q} B_i(t) L_{i-1}(t)}{\sum_{k=p+1}^{q} B_k(t)}, \tag{26}$$

which can be abbreviated to

$$S_{p,q}(t) = \sum_{i=p}^{q-1} w_i^{p,q}(t) L_i(t),$$

by introducing random variables w_i defined by

$$w_i = w_i^{p,q}(t) := \frac{B_{i+1}(t)}{\sum_{k=p+1}^{q} B_k(t)}, \quad \text{for} \quad i = p, ..., q-1.$$

For ease of notation we omit superscript and argument as indicated, but should remember that these weights actually depend on swap dates p and q and time. The coefficients w_i can be regarded as (stochastic) weight factors, since they add to one

$$\sum_{i=p}^{q-1} w_i = 1.$$

In differential form, though only the integrated version is well defined by Itô integrals, we informally obtain

$$dS_{p,q} := \sum_{i=p}^{q-1} w_i dL_i + \sum_{i=p}^{q-1} L_i dw_i + \sum_{i=p}^{q-1} d\langle w_i \, , L_i \rangle \, .$$

We use angle brackets to denote a *previsible* covariation process. It is equal to a regular covariation process for continuous semimartingales. Continuing in a bracket calculus notation, we receive

$$\begin{aligned} d\langle S_{p,q}\rangle &:= d\langle S_{p,q}\, , S_{p,q}\rangle = dS_{p,q}dS_{p,q} \qquad (27) \\ &= \sum_{i,j=p}^{q-1} w_i w_j L_i L_j \left(d\langle \ln L_i, \ln L_j\rangle + 2d\langle \ln w_i, \ln L_j\rangle + d\langle \ln w_i, \ln w_j\rangle \right), \end{aligned}$$

where we used that $d\langle L_i, L_j\rangle = L_i L_j d\langle \ln L_i, \ln L_j\rangle$ and a similar equality for w_i. The question, whether

$$d\langle X, Y\rangle = XY d\langle \ln X, \ln Y\rangle \, ,$$

which holds for general semimartingales, applies to w_i , $i = p, ..., q-1$, is not self-evident. We need w_i , $i = p, ..., q-1$, to be semimartingales. But since the denominator in their definition remains positive with probability one, it is ensured that w_i , $i = p, ..., q-1$, in fact are semimartingales.

Time series data in practice reveal that the w_i behave far less erratic than the L_i. Moreover, they seem to move in a rather smooth manner, although they are random processes. The reason lies in the nature of their definition. The bond in the numerator appears in the denominator. This has a flattening effect on the quantities. As a reasonable approximation we assume their quadratic variation processes to be identically zero, which renders the differentials in (27) involving w_i negligible. In effect, we could consider w_i with finite variation.

$$\begin{aligned} d\langle S_{p,q}\rangle &= S_{p,q}^2 d\,\langle \ln S_{p,q}, \ln S_{p,q}\,\rangle \\ &\approx \sum_{i,j=p}^{q-1} w_i w_j L_i L_j \, d\,\langle \ln L_i, \ln L_j\,\rangle \\ &= \sum_{i,j=p}^{q-1} w_i w_j L_i L_j \gamma_i \cdot \gamma_j \, dt. \end{aligned}$$

If we now introduce a scalar volatility process $|\sigma_{p,q}|$ for the swap rate by defining

$$|\sigma_{p,q}|^2 dt \;:=\; d\,\langle \ln S_{p,q}\rangle\,,$$

we obtain (25)

$$\begin{aligned} S_{p,q}^2 |\sigma_{p,q}|^2 \;&\approx\; \sum_{i,j=p}^{q-1} w_i w_j L_i L_j \gamma_i \cdot \gamma_j \\ &= \sum_{i,j=p}^{q-1} w_i w_j L_i L_j |\gamma_i||\gamma_j| \rho_{ij}^s\,. \end{aligned}$$

Remark 21 .

1. *The definition of $|\sigma_{p,q}|$ makes sense, since $S_{p,q}$ is an Itô process.*
2. *The market prices swaptions in assuming that $S_{p,q}$ follows a geometric Brownian motion. In this case $|\sigma_{p,q}|$ may be chosen as a deterministic function. We then say to price in a Swap market model, as opposed to pricing in a* LIBOR *market model, where γ are deterministic. It can not be accomplished to have both processes deterministic at the same time.*

3. *In a* LIBOR *market model the relative instantaneous volatility process* $\sigma_{p,q}$ *is stochastic in general. This inherits a further approximation, when in following calibration procedures this process is compared with Black volatility processes stemming from a Swap market model.*

We carefully defined the scalar process $|\sigma_{p,q}|$ as a norm of a vector process $\sigma_{p,q}$. We now explicitly determine it in a more rigorous second justification for (25).

A Technical Argument With the same definitions for $S_{p,q}$ and w_i's as above and

$$B_{p,q}(t) := \sum_{k=p+1}^{q} \delta B_k(t) = \sum_{k=p}^{q-1} \delta B_{k+1}(t),$$

which gives

$$w_i(t) = \frac{\delta B_{i+1}(t)}{B_{p,q}(t)}, \quad \text{for} \quad i = p, ..., q-1\,,$$

we obtain by an application of Itô's Lemma, see (1.27) in Schoenmakers (2005), for $p \leq r \leq q$

$$\frac{d(B_r/B_{p,q})}{(B_r/B_{p,q})} = \sum_{j=p}^{q-1} w_j(\sigma_r - \sigma_{j+1}) \cdot \left(\lambda dt - \sum_{k=p}^{q-1} w_k \sigma_{k+1} dt + dW\right)$$

$$=: \sum_{j=p}^{q-1} w_j(\sigma_r - \sigma_{j+1}) \cdot dW^{p,q}.$$

$W^{p,q}$ denotes a standard Brownian motion in the measure induced by the annuity numeraire $B_{p,q}$. The market price of risk process λ and the bond volatility process σ are the same as defined in Section 1.1.1 on arbitrage pricing. Let successively $r = p$ then $r = q$ and use (26) to obtain for the difference

$$dS_{p,q} = \left(\frac{B_p}{B_{p,q}} \sum_{j=p}^{q-1} w_j(\sigma_p - \sigma_{j+1}) - \frac{B_q}{B_{p,q}} \sum_{j=p}^{q-1} w_j(\sigma_q - \sigma_{j+1})\right) \cdot dW^{p,q}$$

$$= S_{p,q}\left(\sum_{j=p}^{q-1} w_j(\sigma_p - \sigma_{j+1}) + \frac{B_q}{B_p - B_q}(\sigma_p - \sigma_q)\right) \cdot dW^{p,q}$$

$$=: \; S_{p,q}\sigma_{p,q} \cdot dW^{p,q},$$

As before $\sigma_{p,q}$ can be interpreted as a relative volatility process which is now clearly seen to be of higher dimension m. It is also obvious that $\sigma_{p,q}$ is in general a random variable in a LIBOR market model. Only in a Swap market model, where it is required to be deterministic, $\ln(S_{p,q})$ can have a gaussian law.

We also may express $\sigma_{p,q}$ in terms of the LIBOR volatility structure $\gamma_i \in \mathbb{R}^m$ by using its definition

$$L_i \gamma_i := \delta^{-1}(1+\delta L_i)(\sigma_i - \sigma_{i+1}).$$

Note that

$$\sigma_p - \sigma_{j+1} = \sum_{k=p}^{j} \frac{\delta L_k}{1+\delta L_k} \gamma_k$$

and

$$\sigma_p - \sigma_q = \sum_{k=p}^{q-1} \frac{\delta L_k}{1+\delta L_k} \gamma_k \,.$$

We receive

$$\begin{aligned} \sigma_{p,q} &= \sum_{j=p}^{q-1} w_j \sum_{k=p}^{j} \frac{\delta L_k}{1+\delta L_k} \gamma_k + \frac{B_q}{B_p - B_q} \sum_{k=p}^{q-1} \frac{\delta L_k}{1+\delta L_k} \gamma_k \\ &= \sum_{k=p}^{q-1} \frac{\delta L_k}{1+\delta L_k} \gamma_k \left(\sum_{j=k}^{q-1} w_j + \frac{B_q}{B_p - B_q} \right), \end{aligned}$$

where the last equality is obtained by rearranging terms in the first sum above. Naturally the swap volatility $\sigma_{p,q} \in \mathbb{R}^m$, since $\gamma_i \in \mathbb{R}^m$ for all $i = p, ..., q-1$. Squaring the Euclidean norm on the left amounts to taking a product of sums on the right hand side. We obtain the exact form of the total volatility process $|\sigma_{p,q}|$ in a LIBOR market model

$$|\sigma_{p,q}|^2 = \frac{1}{S_{p,q}^2} \sum_{k=p}^{q-1} \sum_{l=p}^{q-1} L_k L_l v_k^{p,q} v_l^{p,q} \big(\gamma_k \cdot \gamma_l\big) \,, \tag{28}$$

with definition

$$v_k^{p,q} := \frac{\delta S_{p,q}}{1+\delta L_k} \left(\sum_{j=k}^{q-1} w_j + \frac{B_q}{B_p - B_q} \right) = \frac{\delta}{1+\delta L_k} \left(\sum_{j=k}^{q-1} w_j \frac{B_p - B_q}{B_{p,q}} + \frac{B_q}{B_{p,q}} \right).$$

Equation (28) will be the starting point for approximations such as (25). It is the difference between w_i in (25) and $v_i^{p,q}$ from (28) that entails the major part of inexactness in the approximating formula. Another part is due to *freezing*, as will be explained below. We abbreviate $v_i = v_i^{p,q}$ for a shorter notation, just as we did for w_i.

1.3.1 A General Swaption Approximation

The connection between the instanteneous swap volatility process $|\sigma_{p,q}|$ and the actual observed implied volatility, measured only at finitely many time points on the grid, is obtained by integration of (28) from the present calender date $t_0 = 0$ to $T_p \geq 0$:

$$\int_0^{T_p} |\sigma_{p,q}|^2(t)dt = \sum_{k,l=p}^{q-1} \int_0^{T_p} \frac{v_k(t)v_l(t)L_k(t)L_l(t)}{S_{p,q}^2(t)} \big(\gamma_k(t) \cdot \gamma_l(t)\big)dt\,. \tag{29}$$

If we now assume the instantaneous volatility process $|\sigma_{p,q}|$ to be non-stochastic, we may interpret its integrated average as a realization of an implied Black volatility of an approximated swaption price in the LIBOR market model:

$$(\sigma_{p,q}^B)^2 = \frac{1}{T_p} \int_0^{T_p} |\sigma_{p,q}|^2(t)dt.$$

This involves a first major approximation to the hitherto exact derivation. To be meaningful, the required deterministic behavior of the left hand side in (29), needs to be matched to the right hand side. This can be obtain by *freezing* the stochastic fractions in the integrand at their value at time $t_0 = 0$. Legitimately, Schoenmaker calls the result a general approximation formula:

$$(\sigma_{p,q}^B)^2 \approx \frac{1}{T_p} \sum_{k,l=p}^{q-1} \frac{v_k v_l L_k L_l}{S_{p,q}^2}(0) \int_0^{T_p} \gamma_k(t) \cdot \gamma_l(t)dt. \tag{30}$$

Remark 22 *Does the freezing technique ensure an acceptable result?*

Fortunately the answer to this second question is affirmative. A first observation in this direction is demonstrated in the following lemma. The freezed fractions add up to one, when the yield curve is assumed to be flat. In the

general case, they will at least approximate unity. The fractions can therefore be considered as weights.
The second remarkable feature is that they vary relatively slow in practice.

Lemma 23 *If the swap curve is flat, i.e. if*

$$S_{p,q} = \frac{B_p - B_q}{B_{p,q}} = \frac{B_k - B_q}{B_{k,q}} = S_{k,q}\,, \quad \textit{for all} \quad p \leq k \leq q,$$

then

$$\sum_{k,l=p}^{q-1} \frac{v_k v_l L_k L_l}{S_{p,q}^2} = 1.$$

Proof. Recall the definitions of $B_{p,q}$, w_j, v_j and L_j:

$$B_{p,q} = \sum_{j=p}^{q-1} \delta B_{j+1}, \quad w_j = \frac{\delta B_{j+1}}{B_{p,q}}, \quad L_j = \frac{1}{\delta}\left(\frac{B_j}{B_{j+1}} - 1\right).$$

With these we obtain

$$\begin{aligned}\frac{v_k v_l L_k L_l}{S_{p,q}^2} &= \frac{\delta L_k}{1+\delta L_k}\Big(\sum_{j=k}^{q-1} w_j + \frac{B_q}{B_p - B_q}\Big)\frac{\delta L_l}{1+\delta L_l}\Big(\sum_{j=l}^{q-1} w_j + \frac{B_q}{B_p - B_q}\Big) \\ &= \big(1 - \frac{B_{k+1}}{B_k}\big)\Big(\sum_{j=k}^{q-1} w_j + \frac{B_q}{B_p - B_q}\Big)\big(1 - \frac{B_{l+1}}{B_l}\big)\Big(\sum_{j=l}^{q-1} w_j + \frac{B_q}{B_p - B_q}\Big).\end{aligned}$$

Since the k and l-terms factorize, the double sum over these can be written as a product of two. It therefore suffices to show that every individual sum is equal to one. Take for example the sum over k:

$$
\begin{aligned}
\sum_{k=p}^{q-1} \frac{v_k L_k}{S_{p,q}} &= \sum_{k=p}^{q-1} (1 - \frac{B_{k+1}}{B_k})(\sum_{j=k}^{q-1} w_j + \frac{B_q}{B_p - B_q}) \\
&= \sum_{k=p}^{q-1} (1 - \frac{B_{k+1}}{B_k})(\sum_{j=k}^{q-1} \frac{\delta B_{j+1}}{B_{p,q}} + \frac{B_q}{B_p - B_q}) \\
&= \sum_{k=p}^{q-1} (1 - \frac{B_{k+1}}{B_k})(\sum_{j=k}^{q-1} \frac{\delta B_{j+1}(B_k - B_q)}{B_{k,q}(B_p - B_q)} + \frac{B_q}{B_p - B_q}) \\
&= \sum_{k=p}^{q-1} (1 - \frac{B_{k+1}}{B_k})(\frac{\sum_{j=k}^{q-1} \delta B_{j+1}}{B_{k,q}} \frac{(B_k - B_q)}{(B_p - B_q)} + \frac{B_q}{B_p - B_q}) \\
&= \sum_{k=p}^{q-1} (1 - \frac{B_{k+1}}{B_k})(\frac{B_k - B_q}{B_p - B_q} + \frac{B_q}{B_p - B_q}) \\
&= \sum_{k=p}^{q-1} (1 - \frac{B_{k+1}}{B_k})(\frac{B_k}{B_p - B_q}) \\
&= (\frac{1}{B_p - B_q}) \sum_{k=p}^{q-1} (B_k - B_{k+1}) \\
&= 1.
\end{aligned}
$$

This shows the result. ■

The assumption of a flat curve is strong and may appear not even realistic. Its case should, however, serve as an anchor around which perturbances will naturally occur. The perturbances force the sum of weights to then differ from one. In any case the idea of an approximate weighting of the volatility components can be supported. Furthermore, empirically the fractions do not vary in an erratic way. These facts seem to be enough evidence for practitioners to accept the inexactness due to freezing.

1.3.2 A Further Approximation

The flat yield curve assumption from above will play a significant role also in the next step towards (25). The last obvious step is to exchange the v_i by w_i and to answer the question whether this cut will not be too distorting. The fluctuations are difficult to track, so a course of action is to again approach the issue by considering special cases. If all forward swaps coincide for all maturities (flat coupon-curve), it can be shown that

$$v_k = w_k.$$

Lemma 24 *If*

$$S_{p,q} = \frac{B_p - B_q}{B_{p,q}} = \frac{B_k - B_q}{B_{k,q}} = S_{k,q}\,, \quad \textit{for all} \quad p \leq k < q,$$

then

$$v_k = w_k\,, \quad \textit{for all} \quad p \leq k < q.$$

Proof.

$$\begin{aligned} v_k - w_k &= \frac{\delta}{1+\delta L_k}\Big(\sum_{j=k}^{q-1} w_j \frac{B_p - B_q}{B_{p,q}} + \frac{B_q}{B_{p,q}}\Big) - \frac{\delta B_{k+1}}{B_{p,q}} \\ &= \frac{\delta}{B_{p,q}(1+\delta L_k)}\Big(\sum_{j=k}^{q-1} \frac{\delta B_{j+1}}{B_{p,q}}(B_p - B_q) + B_q - B_{k+1}(1+\delta L_k)\Big) \\ &= \frac{\delta}{B_{p,q}(1+\delta L_k)}\Big(\frac{B_{k,q}}{B_{k,q}}\frac{B_k - B_q}{B_p - B_q}(B_p - B_q) + B_q - B_{k+1}\frac{B_k}{B_{k+1}}\Big) \\ &= \frac{\delta}{B_{p,q}(1+\delta L_k)}\big(B_k - B_q + B_q - B_k\big) \\ &= 0 \end{aligned}$$

for all $p \leq k < q$. ∎

Exchanging w_k for v_k finally yields (25), a popular formula among practitioners, see Jäckel and Rebonato (2003).

The reason for our in-depth analysis of (25), lies in the fact that it will play a major role in the upcoming calibration procedures as a valuable tool for stabilizing the optimization algorithm. We will thus employ the formula extensively for regularization purposes.

1.4 Calibration

1.4.1 General Aspects

Now that we have meaningful parametrizations and adequate approximation tools for swaption volatilities, we can discuss various optimization procedures. In particular there are three different approaches to the calibration problem:

1. First, calibrate the function g to caplet prices. Determine the coefficients c_i from (32). In a second step, calibrate a suitably parametrized correlation function to swaption prices.
2. Joint calibration to the set of caplets and swaptions, i.e. joint identification of g and ϱ.
3. Regularization of least squares routine by an additional restriction.

The first method is fairly stable and ensures that cap volatilities are chosen as time-homogeneous as possible. It suffers, however, from an unsatisfactory fit to the given data. The second method is instable. The last method is the most effective that we refer to in all numerical tests.

1.4.2 Calibration to Caplets

Given (24), minimize first over the parameter $a, b, g_\infty > 0$ of g_{a,b,g_∞} through

$$\min \sum_{p=1}^{n-1} \left((\sigma_p^B)^2 T_p - c_p^2 \int_0^{T_p} g_{a,b,g_\infty}(T_p - s) ds \right)^2 .$$

Determine c_p from (32) for an exact fit. Then optimize over the remaining three parameters η_1, η_2 and ρ_∞ from the correlation parametrized form proposed by Schoenmakers (2005) to swaption volatilities according to the procedure outlined below.

1.4.3 Calibration to Swaptions

Although (25) will be the underlying estimation formula in the calibration procedure, let us start here by considering the general swaption approximation (30). The quantities $(\sigma_{p,q}^B)^2$ are observed in the market directly, as solutions of inverse problems in a swap market model.

$$
\begin{aligned}
(\sigma^B_{p,q})^2 &\approx \frac{1}{T_p} \sum_{k,\,l=p}^{q-1} \frac{v_k v_l L_k L_l}{S^2_{p,q}}(0) \int_0^{T_p} \gamma_k(t)\cdot\gamma_l(t)dt \\
&= \frac{1}{T_p} \sum_{k,\,l=p}^{q-1} \frac{v_k v_l L_k L_l}{S^2_{p,q}}(0) \int_0^{T_p} |\gamma_k(t)||\gamma_l(t)|\rho^s_{kl}dt \\
&= \frac{1}{T_p} \sum_{k,\,l=p}^{q-1} \frac{v_k v_l L_k L_l}{S^2_{p,q}}(0) \int_0^{T_p} \sigma^s_k(t)\sigma^s_l(t)\varrho(T_k-t,T_l-t)dt \\
&= \frac{c_k c_l}{T_p} \sum_{k,\,l=p}^{q-1} \frac{v_k v_l L_k L_l}{S^2_{p,q}}(0) \int_0^{T_p} g(T_k-t)g(T_l-t)\varrho(T_k-t,T_l-t)dt\,,
\end{aligned}
$$

for $1 \leq p < q \leq n$, shows the relation to the parametrized forms of volatility and correlation function introduced in Section 1.2.2. The corresponding estimate for squared market (Black-) volatilities to a simpler approximation involving the w_k is

$$
(\sigma^B_{p,q})^2 \approx \frac{c_k c_l}{T_p} \sum_{k,\,l=p}^{q-1} \frac{w_k w_l L_k L_l}{S^2_{p,q}}(0) \int_0^{T_p} g(T_k-t)g(T_l-t)\varrho(T_k-t,T_l-t)dt\,, \tag{31}
$$

for $1 \leq p < q \leq n$. Note that even though we use the same relation sign, we actually have different values on the right side.
Since one-period swaptions are caplets, denoting $\sigma^B_p := \sigma^B_{p,p+1}$, we will in this case obtain from (31):

$$
(\sigma^B_p)^2 \approx \frac{c_p^2}{T_p} \int_0^{T_p} g^2(T_p-t)dt\,, \quad \text{for} \quad 1 \leq p < n. \tag{32}
$$

This is naturally consistent with (22), as it should. Once the parameters in the function g are determined, the c_k's can immediately be evaluated from this equation.

Regarding the parameters in the correlation function ϱ, define the quantity:

$$\alpha_{k,l,p}^{g,\varrho} := \frac{\sqrt{T_k}\sqrt{T_l}}{T_p} \frac{\int_0^{T_p} g(T_k - t)g(T_l - t)\varrho(T_k - t, T_l - t)dt}{\sqrt{\int_0^{T_k} g^2(T_k - t)dt}\sqrt{\int_0^{T_l} g^2(T_l - t)dt}},$$

for $p \leq \min(k, l)$. An alternative and simpler expression for (31) is then:

$$(\sigma_{p,q}^B)^2 \approx \sum_{k,l=p}^{q-1} \frac{w_k w_l L_k L_l}{S_{p,q}^2}(0)\ \sigma_k^B\ \sigma_l^B\ \alpha_{k,l,p}^{g,\varrho},$$

for $1 \leq p < q \leq n$. Note that $\sigma_{p,q}^B$ and σ_p^B are real numbers extracted from the market. In order to define an objective function, let

$$\sigma_{p,q}^B(g, \varrho) := \sqrt{\sum_{k,l=p}^{q-1} \frac{w_k w_l L_k L_l}{S_{p,q}^2}(0)\ \sigma_k^B\ \sigma_l^B\ \alpha_{k,l,p}^{g,\varrho}},$$

and consider the root mean square minimization:

$$RMS(g, \varrho) := \sqrt{\frac{2}{(n-1)(n-2)} \sum_{1 \leq p \leq q-2, q \leq n} \left(\frac{\sigma_{p,q}^B - \sigma_{p,q}^B(g, \varrho)}{\sigma_{p,q}^B} \right)^2},$$

which is to be solved for six parameters in the problem. Unfortunately, there are many examples that demonstrate the instability of the above non-linear optimization problem.

1.4.4 Calibration by Regularization

An additional restriction is obtained by an equation that reflects a relationship between caplet and swap volatilities. Among market practitioners it is accepted as a reasonable rule-of-thumb. If we exchange the *variance*-type quantities

$$\frac{\int_0^{T_k} g^2(T_k - t)dt}{T_k}$$

in $\alpha^{g,\varrho}_{k,l,p}$ above, by somewhat similar expressions

$$\frac{\int_0^{T_p} g^2(T_k - t)dt}{T_p},$$

where integration and average are considered over a shorter interval $[0, T_p]$, since $p \leq \min(k, l)$, we find an approximation

$$\alpha^{g,\varrho}_{k,l,p} \approx \rho^{global;g,\varrho}_{kl,p} := \frac{\int_0^{T_p} g(T_k - t)g(T_l - t)\varrho(T_k - t, T_l - t)dt}{\sqrt{\int_0^{T_p} g^2(T_k - t)dt}\sqrt{\int_0^{T_p} g^2(T_l - t)dt}}.$$

This quantity remarkably reminds one of the global correlation between $L_k(T_p)$ and $L_l(T_p)$. Indeed it is easily comprehensible that

$$\begin{aligned} Cor(L_k(T_p), L_l(T_p)) &\approx Cor(\ln L_k(T_p), \ln L_l(T_p)) \\ &\approx \frac{\int_0^{T_p} g(T_k - t)g(T_l - t)\varrho(T_k - t, T_l - t)dt}{\sqrt{\int_0^{T_p} g^2(T_k - t)dt}\sqrt{\int_0^{T_p} g^2(T_l - t)dt}} \\ &= \rho^{global;g,\varrho}_{k,l,p}. \end{aligned}$$

Hence if we define

$$\begin{aligned} (\sigma^{MSF}_{p,q})^2 &:= \sum_{k,l=p}^{q-1} \frac{w_k w_l L_k L_l}{S^2_{p,q}}(0)\ \sigma^B_k\ \sigma^B_l\ \rho^{global;g,\varrho}_{k,l,p} \\ &= \sum_{k,l=p}^{q-1} \frac{w_k w_l L_k L_l}{S^2_{p,q}}(0)\ \sigma^B_k\ \sigma^B_l\, Cor(L_k(T_p), L_l(T_p)), \end{aligned}$$

we obtain a further representation for the Black volatilities in swaption prices. This representation in terms of global model correlations of the LIBOR process may be implemented as a regularizing equation. Incorporating it into the least squares minimization routine helps to identify correlations less ambiguously. It thus serves as a remedy against the intrinsic instability in the joint calibration method.

1.5 Heston and the CIR Square-root Process

Let us briefly review the stochastic volatility model proposed by Heston (1993), as the model we analyze in this work has some similar components. Heston's approach to explain deviations between model and market prices was at first developed for stocks, as were many other attempts like displaced diffusions (DD), constant elasticity of variance (CEV) or (pure-) jump models, to name a few. The Black-Scholes model tends to underestimate prices for out-of-the-money call options and to overestimate prices for out-of-the-money put options due to an imperfect match between lognormal and true density. The same holds true for interest rate derivatives. Stochastic volatility models attempt to achieve more degrees of freedom by introducing additional parameters. The following Feller square root process was proposed by Cox, Ingersoll, and Ross (1985), referred to as CIR process, who justified the evolution of short interest rates on economic grounds. They developed a general equilibrium framework based on the assumption that if change in production opportunity is required to follow a square-root diffusion

$$dv_t = \kappa(\theta - v_t)\, dt + \sigma\sqrt{v_t}\, dW_t\,, \tag{33}$$

then so should short term interest rates.

1.5.1 Features and Distributional Properties

All auspicious characteristics for interest rates are, however, also interesting for a volatility environment. We consider here the case where κ and θ are positive. If $v_0 > 0$, v_t will almost surely never be negative. In case that $2\kappa\theta \geq \sigma^2$, v_t remains even strictly positive, almost surely, for all t. As in a model previously proposed by Vasicek (1977), the drift term in (33) suggests that v_t is pulled back towards θ at a speed determined by κ. Both are therefore referred to as *mean-reversion models*. In contrast to Vasicek's model, the diffusion term in the CIR offers the feature of decreasing to zero as volatility v_t approaches the origin. This prevents process v_t from taking negative values, a quality that makes (33) attractive for modeling volatility processes.

It is noteworthy that the Feller diffusion process (33) was originally introduced to model interest rates, but in the follow was used to explain volatility evolutions. Meanwhile interest rates are modeled by more general Itô processes, as in our case the LIBORS L_i.

Here is a summary of positive features a CIR process offers when used to explain volatility behavior.

(i) Negative values are precluded.

(ii) If zero is reached, it can again become positive (non-absorbing).

(iii) There is a steady state distribution for the volatility.

(iv) Analytical tractability due to known conditional distributions.

The last item is of particular importance. The probability density of a volatility at time t, conditioned on its value at current time s, is given by:

$$f(v_t, t; v_s, s) = ce^{-u-w}\left(\frac{w}{u}\right)^{q/2} I_q\left(2\,(uw)^{1/2}\right), \tag{34}$$

where

$$\begin{aligned} c &= \frac{2\kappa}{\sigma^2(1 - e^{-\kappa(t-s)})}, \\ u &= cv_s e^{-\kappa(t-s)}, \\ w &= cv_t, \\ q &= \frac{2\kappa\theta}{\sigma^2} - 1. \end{aligned}$$

$I_q(\cdot)$ denotes the modified Bessel function of first kind and order q. The distribution function is noncentral chi-square, $\chi^2(2cv_t; 2q+2, 2u)$, with $2q+2$ degrees of freedom and noncentrality parameter $2u$.
Straightforward calculations show that expected value and variance of v_t are:

$$\begin{aligned} E(v_t|v_s) &= v_s e^{-\kappa(t-s)} + \theta(1 - e^{-\kappa(t-s)}), \\ Var(v_t|v_s) &= v_s\left(\frac{\sigma^2}{\kappa}\right)(e^{-\kappa(t-s)} - e^{-2\kappa(t-s)}) + \theta\left(\frac{\sigma^2}{2\kappa}\right)(1 - e^{-\kappa(t-s)})^2. \end{aligned}$$

The case $s = 0$ is of significance to our model. We explicitly use the expected value as an input. Consequently the quantity we need is

$$E(v_t) = v_0 e^{-\kappa t} + \theta(1 - e^{-\kappa t}).$$

Remark 25 *CIR processes offer a particular useful feature. When part of a Heston model, we can determine the underlying's characteristic function explicitly. In an attempt to solve an inverse problem, optimization over parameter sets is then possible.*

1.5.2 Existence of Weak and Strong Solutions

Strong Solution for v. In (33) the classical Lipschitz-condition is not satisfied for all diffusion coefficients. They do, however, satisfy the Yamada & Watanabe assumption, see Richard Durrett (1996) or earlier Ikeda & Watanabe (1992). Pathwise uniqueness and uniform L^2-convergence to a strong solution is therefore ensured.

Weak Solution for X. More involved is the case of the Heston model as a vector valued SDE, as in (35) below. The existence of a weak solution for X can be concluded from general versions of Strook-Varadhan's existence results, for example Theorem (24.1) in Rogers/Williams (2000). Under the given assumptions the martingale problem is well-posed, which is equivalent to the existence of a unique weak solution for X. Moreover, the strong Markov property holds, i.e. X is in fact a Markov process.

Strong Solution for X. Regarding the existence of a strong solution for X, Theorem 3.1.1. in Prevot/Röckner (2007) supplies a suitable result. In their theorem it is crucial that the coefficient functions of X are progressively measurable for fixed x. Given the strong solution for v, as argued above, this is the case in (35). It thus exists a unique solution up to P-indistinguishability for the SDE. This strong solution is, furthermore, P-as continuous and adapted.

In this work the concept of a local martingale, the introduction of stopping times and the utilization of the localisation machinery can be abstained from. Uniform L^2-integrability suffices to conclude that all local martingales are martingales.

Furthermore, we consider only continuous processes, both as integrands and integrators as part of Itô integrals. In more general settings integrands, interpreted as hedging strategies, are chosen left continuous with right limits (caglad). Integrators in form of a jump-process are naturally considered right continuous with left limits (cadlag). In our framework, however, all integrators are continuous and integrands at least adapted. Itô integrals are well-defined under these conditions.

1.5.3 Moment Explosion in a Heston Model

In a recent monograph Andersen/Piterbarg (2007) address a problem that occurs in many stochastic volatility models. Moments of order higher one can become infinite in finite time. With respect to arbitrage free pricing of products this undesirable property could produce infinite prices. In LIBOR-in-arrears and constant maturity swap structures, for example, expectations of functions with super-linear growth are to be evaluated. We will, however, demonstrate that this feature will become austere not till a maturity region which is of no significance to the market.

In their treatise Andersen/Piterbarg analyze general stochastic volatility models of the form

$$\begin{aligned} dX(t) &= \gamma(t) f(X(t)) \sqrt{v(t)} \cdot dW_X(t), \\ dv(t) &= \kappa(\theta - v(t))dt + \sigma v^p(t) dW_v(t). \end{aligned}$$

For brevity, we consider the case $p = 1/2$, $f = id$ and γ constant. The last postulate is a simplification of our model, where γ is left time-dependent. As long as this time dependence ensures bounded and deterministic γ, the results presented in the sequel hold as stated. Note further that in this section we ignore a possible drift of X without loss of generality. The process represents a forward LIBOR L_i in the appropriate martingale measure. We thus examine a model of the form

$$\begin{aligned} dX(t) &= \gamma X(t) \sqrt{v(t)} \cdot dW_X(t), \\ dv(t) &= \kappa(\theta - v(t))dt + \sigma \sqrt{v(t)} dW_v(t), \end{aligned} \tag{35}$$

where the parameters κ, θ, σ and γ are strictly positive. W_X and W_v are correlated Brownian motions on a probability space with given measure P, satisfying $dW_X(t)dW_v(t) = \rho dt$.

It is well known that a solution for v in SDE (35) can reach zero if $2\kappa\theta < \sigma$. In our case, however, the Yamada condition, see Karatzas/Shreve (1991) proposition (2.13) from Yamada & Watanabe, shows that v has a unique strong solution on $[0, \infty)$. Given this, the origin is strongly reflecting, in the sense that the length of time spent at $v = 0$ is of Lebesque measure zero, as shown in Revuz/Yor (1994, Chap XI). No specific boundary condition is therefore needed. In the sequel we recap the main results, whose proofs are reduced to sources and short comments.

Proposition 26 *The stationary distribution density of v in (35) is given by*

$$\varphi(y) = c\, y^{(2\kappa\theta\sigma^{-2}-1)} e^{-2\kappa\sigma^{-2}\, y},$$

with

$$c = \int_0^\infty u^{(2\kappa\theta\sigma^{-2}-1)} e^{-2\kappa\sigma^{-2}\, u} du\,,$$

thus ensuring a probability measure.

Proof. By direct computation from (34) and the non-central χ^2 density. See also Andersen/Piterbarg (2007). ■

The next lemma shows how moments of X can be expressed as exponential moments of the process v.

Lemma 27 *For the m^{th} moment of X we have*

$$EX^m(T) = X^m(0)E^{Q_m}(e^{k\int_0^T v(u)du}), \quad k = \gamma^2 m(m-1)/2 .$$

With respect to the measure Q_m, the dynamics of v are given by

$$dv(t) = [\kappa\theta + (\rho\sigma\gamma m - \kappa)v(t)]dt + \sigma\sqrt{v(t)}dW_v^m(t),$$

with W_v^m a Q_m-Brownian motion.

Proof. Can be deduced from Sin (1998) by an extension of a limit argument (Lemma 4.2). Note here that in case of a square root diffusion, $p = 1/2$, the explosion time for v,

$$\tau_\infty = \lim_{n\to\infty} \tau_n , \quad \tau_n = \inf\{t \in \mathbb{R}^+ : v(t) \geq n\}.$$

is not finite, therefore $1_{\{\tau_\infty > T\}} = 1 , as.$ ■

The following proposition is the main result of this section. It derives sharp conditions for finiteness of moments of X.

Proposition 28 *Consider the process (35). Fix $k = \gamma^2 m(m-1)/2 > 0$ and define*

$$b = 2k/\sigma^2 > 0 , \quad a = 2(\rho\sigma\gamma m - \kappa)/\sigma^2 , \quad D = a^2 - 4b.$$

$EX^m(T)$ will be finite for $T < T^*$ and infinite for $T \geq T^*$, where T^* is determined in various cases as follows:

1. *For $D \geq 0$ and $a < 0$:*
$$T^* = \infty ,$$
2. *for $D \geq 0$ and $a > 0$:*
$$T^* = c^{-1}\sigma^{-2}\ln\left(\frac{a/2 + c}{a/2 - c}\right), \quad c = \sqrt{D}/2 ,$$

3. for $D < 0$:

$$T^* = 2d^{-1}\sigma^{-2}\left(\pi 1_{\{a<0\}} + \arctan(2d/a)\right) , \quad d = \sqrt{-D}/2 .$$

Proof. Again see Andersen/Piterbarg (2007). ■

Note first that empirical data suggest to consider only $\rho \leq 0$. We will therefore always have $a < 0$, since $\kappa > 0$. If additionally $D \geq 0$ there will be no restriction at all, $T^* = \infty$, and any moment will exist. The only restrictive case will thus be when $D < 0$. The figure below depicts a realistic case for $m = 1.81$. One sees that for zero or negative correlation the critical T^*-zone begins far from maturity regions which are traded in the market.

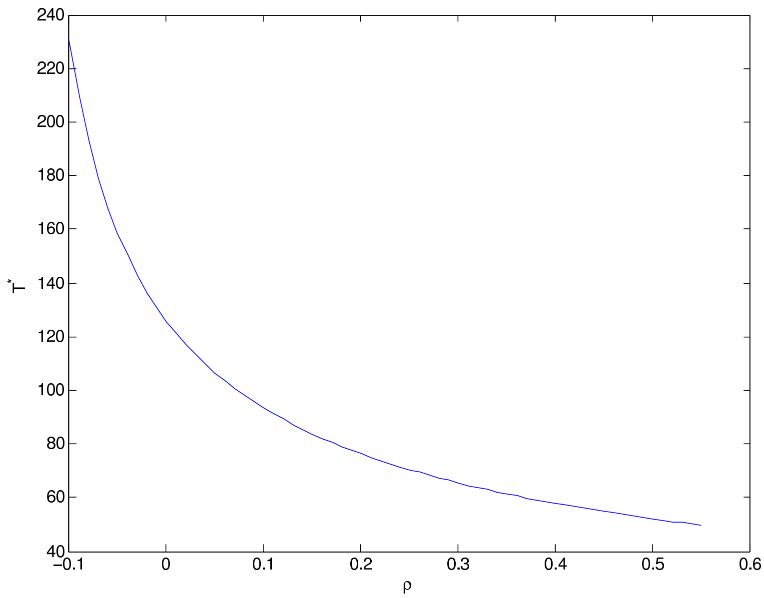

Figure 1: Critical time T^* vs ρ for $m = 1.81$.
The parameters are $\gamma = 30\%$, $\sigma = 30\%$, $\kappa = 10\%$.

1.5.4 Moment Matched Log-Euler Method

Numerical integration of a coupled stochastic volatility system (35) is a twofold exercise. For one, we have to find an appropriate method for the approximation of the stochastic volatility process, while we secondly have to incorporate it into the dynamics of the underlying process X. In this section we deal with the first problem. Good news in that regard is that we can treat the volatility process (33) separately, since it does not explicitly depend on the underlying X. Bad news, on the other hand, is that a discrete numerical integration scheme might not enjoy the theoretically so required structural characteristic of staying positive, a feature the continuous process inhibits, at least for a sufficient parameter region, as seen in the last subsection. Already the simplest approach, namely the explicit Euler-Maruyama scheme

$$v_{n+1} = v_n + \kappa(\theta - v_n)\Delta t_n + \sigma\sqrt{v_n}\Delta W_n,$$

fails to preserve positivity. The same holds true for standard Milstein and Milstein+ schemes.

Various alternative schemes have been proposed in order to avoid this deficiency. Among these are the Balanced Implicit Method (BIM) and the Balanced Milstein Method (BMM). Both methods require the introduction of two control or weight functions, respectively, the choice of which depends strongly on the SDE structure. This arbitrariness and the fact that the BIM method has worse convergence properties than the explicit Euler method, according to Kahl/Jaeckel (2006), are certainly disadvantageous attributes. On the other hand, in fairness we have to admit, that the BMM method seems to perform better than the moment matched Log-Euler method we want to employ here, again see Kahl/Jaeckel (2006). However, the latter, introduced in the work of Andersen/Brotherton-Ratcliffe (2001), is better suited for applications to processes in the financial fields.

While the explicit Euler, BIM and BMM methods leave the stochastic differential equation in its orignal form when discretizing it, Andersen/Brotherton-Ratcliffe start from a log-transformed version. Using Itô's Lemma they obtain

$$d\ln v_t = \frac{2\kappa(\theta - v_t) - \sigma^2}{2v_t}dt + \frac{\sigma}{\sqrt{v_t}}dW_t\,.$$

The Euler scheme applied to this SDE preserves positivity, but is likely to turn instable as both drift and diffusion coefficient explode near zero. For that reason Andersen/Brotherton-Ratcliffe (2001) propose an integration scheme, we will refer to as the moment matched Log-Euler method:

$$\begin{aligned} \widehat{v}_{n+1} &= \left(\theta + (\widehat{v}_n - \theta)e^{-\kappa\Delta t_n}\right) e^{-\frac{1}{2}D_n^2 + D_n\zeta}\,, \\ D_n^2 &= \ln\left(1 + \frac{\sigma^2}{2\kappa\theta}\left(1 - \frac{e^{-2\kappa\Delta t_n}\widehat{v}_n^2}{(\theta + (\widehat{v}_n - \theta)e^{-\kappa\Delta t_n})^2}\right)\right)\,, \end{aligned} \tag{36}$$

where $\zeta \sim \mathcal{N}(0,1)$ is chosen independent of v. Its derivation is given in their appendix. As the analysis of Kahl/Jaeckel shows, the moment matched Log-Euler method has hardly improved convergence properties compared to the straightforward explicit Euler method. Depending on the regions of σ, it is occasionally faster, but never significantly slower. It should thus be even considered a fortunate case, to have a scheme at hand delivering beneficial features like positivity and not having to pay a price .

We seek to find an easier representation for the numerical scheme (36) and succeed by normalizing it appropriately. In fact one could normalize the stochastic process rather than the scheme and still reach the same conclusion. But first note that an easy calculation gives

$$E(\widehat{v}_{n+1}|\widehat{v}_n) = \theta + (\widehat{v}_n - \theta)e^{-\kappa \Delta t_n} \,,$$

whence we can write equation (36) as

$$D_n^2 = \ln\left(1 + \frac{\sigma^2}{2\kappa\theta}\left(1 - \frac{e^{-2\kappa\Delta t_n}\widehat{v}_n^2}{E^2(\widehat{v}_{n+1}|\widehat{v}_n)}\right)\right) .$$

Again the appendix of Andersen/Brotherton-Ratcliffe (2001) and another simple calculation show that

$$\begin{aligned} E(\widehat{v}_{n+1}|\widehat{v}_n) &= E(v_{n+1}|\widehat{v}_n), \qquad (37)\\ E(\widehat{v}_{n+1}^2|\widehat{v}_n) &= E(v_{n+1}^2|\widehat{v}_n). \end{aligned}$$

Note that the first equation in (37) implies that $\widehat{v}$ is a consistent numerical scheme for v. Concerning the above mentioned normalization, define

$$\overline{v}_{n+1} := \frac{\widehat{v}_{n+1}}{E(\widehat{v}_{n+1}|\widehat{v}_n)} \,,$$

It immediately follows that

$$\overline{v}_{n+1} = e^{-\frac{1}{2}D_n^2 + D_n\zeta} \,.$$

We thus have

$$E(\overline{v}_{n+1}|\overline{v}_n) = 1 \,,$$

which holds, because the generated fields in the conditional expectation are the same, that is $\sigma(\overline{v}_n) = \sigma(\widehat{v}_n)$. More importantly, the last equality signifies that $\overline{v}$ is also a consistent numerical scheme, but now for $v_{t_{n+1}}/E(v_{t_{n+1}}|v_{t_n})$. This is true, because apparently

$$E\left(\frac{v_{t_{n+1}}}{E(v_{t_{n+1}}|v_{t_n})}|v_{t_n}=\overline{v}_n\right)=1\,.$$

As a last step, the explicit determination of the second conditional moment will help to simplify (36).

$$\begin{aligned}
E(\overline{v}_{n+1}^2|\overline{v}_n) &= E\left(\frac{\widehat{v}_{n+1}^2}{E^2(\widehat{v}_{n+1}|\widehat{v}_n)}|\frac{\widehat{v}_n}{E(\widehat{v}_n|\widehat{v}_{n-1})}\right)\\
&= \frac{1}{E^2(\widehat{v}_{n+1}|\widehat{v}_n)}E\left(\widehat{v}_{n+1}^2|\widehat{v}_n\right)\\
&= \frac{1}{E^2(\widehat{v}_{n+1}|\widehat{v}_n)}E\left(v_{n+1}^2|\widehat{v}_n\right)\\
&= \frac{1}{E^2(\widehat{v}_{n+1}|\widehat{v}_n)}\left[\left(1+\frac{\sigma^2}{2\kappa\theta}\right)E^2\left(v_{n+1}|\widehat{v}_n\right)-\frac{\sigma^2}{2\kappa\theta}e^{-2\kappa\Delta t_n}\widehat{v}_n^2\right]\\
&= \frac{1}{E^2(\widehat{v}_{n+1}|\widehat{v}_n)}\left[\left(1+\frac{\sigma^2}{2\kappa\theta}\right)E^2\left(\widehat{v}_{n+1}|\widehat{v}_n\right)-\frac{\sigma^2}{2\kappa\theta}e^{-2\kappa\Delta t_n}\widehat{v}_n^2\right]\\
&= \left(1+\frac{\sigma^2}{2\kappa\theta}\left(1-\frac{e^{-2\kappa\Delta t_n}\widehat{v}_n^2}{E^2(\widehat{v}_{n+1}|\widehat{v}_n)}\right)\right),
\end{aligned}$$

where the second equality follows from the same σ-field argument as above, the third and the fifth follow from (37) and the fourth equality from the appendix. We thus obtain a scheme that can be written in a simpler way:

$$\begin{aligned}
\overline{v}_{n+1} &= e^{-\frac{1}{2}D_n^2+D_n\zeta}\,,\\
D_n^2 &= \ln\left(E(\overline{v}_{n+1}^2|\overline{v}_n)\right)\\
&=: \ln\left(S_n(\overline{v}_n)\right)\,.
\end{aligned}$$

Remark 29 *Here is an idea for future work: If one approximates*

$$E(\widehat{v}_{n+1}|\widehat{v}_n)\approx E(\widehat{v}_n|\widehat{v}_{n-1})\,,$$

then D_n^2 *can be written*

$$\begin{aligned} D_n^2 &= \ln\left(1 + \frac{\sigma^2}{2\kappa\theta}\left(1 - \frac{e^{-2\kappa\Delta t_n}\widehat{v}_n^2}{E^2(\widehat{v}_{n+1}|\widehat{v}_n)}\right)\right) \\ &\approx \ln\left(1 + \frac{\sigma^2}{2\kappa\theta}\left(1 - \frac{e^{-2\kappa\Delta t_n}\widehat{v}_n^2}{E^2(\widehat{v}_n|\widehat{v}_{n-1})}\right)\right) \\ &= \ln\left(1 + \frac{\sigma^2}{2\kappa\theta}\left(1 - e^{-2\kappa\Delta t_n}\overline{v}_n^2\right)\right). \end{aligned}$$

We obtain a nice and short scheme that can be evaluated fairly easy.

2 Stochastic Volatility Libor Models and a Common Method

2.1 General Aspects

A Definition: Stochastic volatility models are used in the field of quantitative finance to evaluate derivative securities, such as options. As opposed to earlier versions, these models treat the underlying security's volatility as a random process, governed by state variables such as price level of underlying, tendency of volatility to revert to some long-run mean value, and variance of the volatility process itself, among others.

A famous earlier version is the Black-Scholes model. Stochastic volatility models are an approach to resolve some of its shortcomings. In particular, Black-Scholes models assume that the underlying volatility is constant over lifetime of the derivative, thus unaffected by changes in price level of the underlying. Under these assumptions, they cannot explain long-observed features of implied volatility surfaces, such as its smile and skew. Both indicate that implied volatility does tend to vary with respect to strike price and expiration. By assuming that the volatility of the underlying price is a stochastic process rather than a constant, it is possible to model derivatives more accurately.

A Justification for Existence: The definition above is kept general in the sense that it is applicable to any underlying. A working paper by Jarrow/Li/Zhao (2003) under the meaningful title "Interest Rate Caps "Smile" Too! But Can the LIBOR Market Models Capture It?" is a source, where specifically the drawback of LIBOR market models is elaborated.
Here are their objectives and results in a short summary. They

- study the ability of generalized LIBOR market models to capture smiles.
- identify constant elasticity of variance (CEV) models with uncorrelated stochastic volatility as the best among the ones considered.
- discover, however, that they have a bias for short- and medium-term caps.
- conclude that the existing LIBOR market models do not capture fully the volatility smile.

The current caps and swaption pricing literature has focused on two issues over the last years. The first is referred to as the "unspanned stochastic volatility" puzzle documented by Collin-Dufresne and Goldstein (2002) and Heidari and Wu (2001). It comprises that there appear to be additional risk factors driving caps and swaption markets, different from those factors explaining LIBOR or swap rates. In other words, factors that are not in the space spanned by the latter. The second issue concerns the relative pricing between caps and swaptions, which has been addressed in papers as those by Longstaff, Santa-Clara, Schwartz (2001) and Jagannathan, Kaplin, Sun (2001). Both identify a significant and systematic mispricing between caps and swaptions using various multi-factor term structure models, see also Collin-Dufresne and Goldstein (2001).
Note that both, forward and swap rates, cannot be required to simultaneously follow a lognormal distribution in the same model. It is only natural that one of them has to be approximated in one way or the other. An error in either product group is therefore foreseeable. The remaining question is, how it can be minimized.

Jarrow, Li and Zhaos work is one of very few studies involving data of caps and swaptions with different strikes. It was their hope that studying these "will provide new insights about existing term structure models that are not available from ATM options". As already pointed out by Dai and Singleton (2002) a year earlier, there is an "enormous potential for new insights from using derivatives data in (dynamic term structure) model estimations". We commit ourselves to this impulse and believe their line of reasoning. Our ambition is devoted to identifying and extracting more information from different strikes in our interest rate derivatives, because capturing the volatility smile in caps and swaptions offers not only an interesting challenge to existing term structure models, but also provides an alternative perspective for examining their performance.

In analyzing the LIBOR market model, the authors draw the further conclusions:

- The standard LIBOR market model has large pricing errors and performs especially poor after September 11, 2001, a period with a more pronounced volatility smile.
- The constant elasticity variance model combined with uncorrelated stochastic volatility [...] performs best, except for short and medium-term caps.[...] Although an improvement over the standard LIBOR market model, they show that these generalized LIBOR models are incapable of capturing the entire smile.

- With regard to the forementioned unspanned stochastic volatility puzzle, their findings affirm that the three term structure factors explain only 50% to 70% of the variations in implied cap volatilities, even when combined with the two additional factors moneyness and time-to-maturity.

A Common Method: Already in the introduction we noticed that only a few models allow for correlation between underlying and stochastic volatility process. Among these the model by Wu/Zhang (2006). Assuming independence, the multivariate distribution of LIBORS, as in (9), can be determined as a simple product of marginal distributions, even in a stochastic volatility model. Not so in case of dependence. Consequently, other methods have to be employed. In particular, if stochastic volatility is modeled by CIR processes, the distribution of an underlying may be hard to find. However, the characteristic function of its distribution is known. Wu/Zhang were the first to apply this result to a one-dimensional volatility process. Remarkable is that they use the moment generating function (MGF) in their derivation, meanwhile Carr/Madan show their results only for characteristic functions. For both cases an inverse transformation exists, given in form of an integral-representation.

The following section is dedicated to the Fast Fourier Transform (FFT) method by which the option price can be determined in an efficient way. The pioneering work for this methodology was done by Carr/Madan in 1998. In order to apply it successfully, we need either the *characteristic* or the *moment-generating* function of a given variable. The seminal result of Feynman-Kac helps to identify it as the solution of a (deterministic) partial differential equation. If this PDE can be solved, and we consider such a fortunate case, we can proceed to look for a stable calibration procedure. The following two sections elaborate on these concepts. The last three then present some alternative stochastic volatility models.

2.2 Option valuation by Fast Fourier Transform

Wu/Zhang pay full credit to Carr/Madan: "...[they] discovered that the Fourier transform of an option price can be expressed in terms of the moment generating function of the underlying state variable". We would like to make two comments regarding this quotation in first posing two questions that may arise very naturally expressing some confusion about what we know from probability theory.

(a) The Fourier transform is, in general, a complex function. The moment generating function, on the other hand, was introduced for its simplicity of mapping into the reals. How can they occur, in full generality, on different sides of an equation?

(b) Without discrediting their outstanding achievement, did Carr/Madan indeed *discover* the forthcoming results, or didn't they rather *apply* known results from probability theory to the financial problem of determining option prices?

We give the resolution shortly and illustrate how the statement of Wu/Zhang is to be interpreted in comparison to the original paper of Carr/Madan. The following definition illustrates the difference between two concepts.

Definition 30 *The characteristic funtion of a probability measure μ on the line is defined for real t by*

$$\begin{aligned}\phi(t) &= \int_{-\infty}^{\infty} e^{\mathrm{i}tx}\mu(dx) \\ &= \int_{-\infty}^{\infty} \cos(tx)\mu(dx) + \mathrm{i}\int_{-\infty}^{\infty} \sin(tx)\mu(dx).\end{aligned}$$

A random variable X with distribution μ has characteristic function

$$\phi(t) = \mathbb{E}[e^{\mathrm{i}tX}] = \int_{-\infty}^{\infty} e^{\mathrm{i}tx}\mu(dx).$$

Whereas its moment generating function is defined by

$$M(s) = \mathbb{E}[e^{sX}] = \int_{-\infty}^{\infty} e^{sx}\mu(dx).$$

The characteristic function can thus be defined in terms of a moment generating function, with $\mathrm{i}t$ replacing a real argument s. As opposed to the moment genarating function, it always exists, because $\exp(\mathrm{i}\,tx)$ is bounded. The

characteristic function of a random variable is, in non-probabilistic terms, nothing but the Fourier transform of its density. Reverting it by the inversion formula gives back the original density.
The question under which circumstances an option price can be determined using characteristic functions, reduces to the question of existence of an explicit form for it. The cardinal assumption in the paper of Carr/Madan is thus to have ϕ in analytic form. Before we discuss their results we elaborate on some fundamental properties of characteristic functions. Although only the second is of relevance later, we list three:

(i) If μ_1 and μ_2 have respective characteristic functions $\phi_1(t)$ and $\phi_2(t)$, then $\mu_1 * \mu_2$ has characteristic function $\phi_1(t) \cdot \phi_2(t)$. Naturally it is simpler to study products of characteristic functions, instead of convolutions of measures.

(ii) Characteristic functions uniquely determine distributions. Therefore in studying them, no information is lost.

(iii) Pointwise convergence of characteristic functions implies weak convergence of corresponding distributions.

The starting point of the idea of Carr/Madan is the unique correspondence between densities and their Fourier transforms.

2.2.1 Carr and Madan

A security maturing at time T is denoted by S_T . Recall that the characteristic function of $s_T = \ln S_T$ is given by

$$\phi_T(u) = \mathbb{E}[\exp \mathrm{i}\, u s_T].$$

Denoting the risk-neutral density of this ln price s_T by $q_T(s)$, we obtain

$$\phi_T(u) = \int_{-\infty}^{\infty} e^{\mathrm{i}\, us} q_T(s) ds \tag{38}$$

as a first expression for the characteristic function.
In terms of this density, the price of a T-maturity call with strike K can be computed by

$$C_T(k) = \mathbb{E}\,(S_T - K)^+ = \int_k^{\infty} e^{-rT}(e^s - e^k) q_T(s) ds,$$

where $k = \ln K$ denotes the natural logarithm of the strike price. Since $C_T(k)$ converges to S_0 as $k \to -\infty$, this call price is not square integrable as a function of k. Square integrability, as an assumption of the inversion formula for Hilbert spaces, facilitates results considerably. A simple scaling transformation avoids this dilemma and ensures that the resulting function is square integrable over the entire real line. Consider the following damped version of the price

$$c_T(k) = e^{\alpha k} C_T(k),$$

with $\alpha > 0$. The Fourier transform of $c_T(k)$ is:

$$\psi_T(v) = \int_{-\infty}^{\infty} e^{\mathrm{i}vk} c_T(k) dk.$$

By inversion formula and rescaling we receive back the original option price

$$C_T(k) = \frac{\exp(-\alpha k)}{2\pi} \int_{-\infty}^{\infty} e^{-\mathrm{i}vk} \psi_T(v) dv = \frac{\exp(-\alpha k)}{\pi} \int_{0}^{\infty} e^{-\mathrm{i}vk} \psi_T(v) dv. \quad (39)$$

The second equality holds, because $C_T(k)$ is real, which implies an odd imaginary and an even real part for $\psi_T(v)$.
The ingenuity of the approach is more obvious, when this latter function $\psi_T(v)$ is expressed in terms of the characteristic function ϕ_T.

$$\begin{aligned}
\psi_T(v) &= \int_{-\infty}^{\infty} e^{\mathrm{i}vk} \int_{k}^{\infty} e^{\alpha k} e^{-rT} (e^s - e^k)\, q_T(s) ds\, dk \\
&= \int_{-\infty}^{\infty} e^{-rT} q_T(s) \int_{-\infty}^{s} (e^{s+\alpha k} - e^{(1+\alpha)k}) e^{\mathrm{i}vk} dk\, ds \\
&= \int_{-\infty}^{\infty} e^{-rT} q_T(s) \left[\frac{e^{(\alpha+1+\mathrm{i}v)s}}{\alpha + \mathrm{i}v} - \frac{e^{(\alpha+1+\mathrm{i}v)s}}{\alpha + 1 + \mathrm{i}v} \right] ds \\
&= \frac{e^{-rT} \phi_T(v - (\alpha+1)\mathrm{i})}{\alpha^2 + \alpha - v^2 + \mathrm{i}(2\alpha+1)v}. \quad (40)
\end{aligned}$$

Once (40) is substituted into (39), all that remains to be done is a numerical quadrature of an integral. This can most efficiently be handled by the Fast Fourier method, abbreviated usually by FFT.

Before returning to it, a short comment on the way Wu/Zhang interpreted the results is appropriate. In their paper the last equality of (40) reads as follows

$$\psi_T(v) = \frac{\widetilde{\phi}_T(1 + \alpha + \mathrm{i}v)}{(\alpha + \mathrm{i}v)(1 + \alpha + \mathrm{i}v)}. \quad (41)$$

Ignoring the discount factor e^{-rT} for a while – their underlying is a forward interest rate – equality between these two expressions can only be reached if

$$\phi_T(v - (\alpha+1)\mathrm{i}) = \widetilde{\phi}_T(1 + \alpha + \mathrm{i}v)$$

which is the case, if $\widetilde{\phi}$ is defined as

$$\widetilde{\phi}(v) = \mathbb{E}[\exp vs_T] = \int_{-\infty}^{\infty} e^{vs} q_T(s) ds.$$

Apparently $\widetilde{\phi}$ is a moment generating rather than a characteristic function. It therefore turns out, that both derivations are correct as such, only different definitions have been taken into account.
Regarding the initial quotation of Wu/Zhang at the beginning of the paragraph, we can now adopt the actual correctness of their statement. On the left hand side of (41) with $\psi_T(v)$ we have in fact a "Fourier transform of an option price", meanwhile the right hand side can be interpreted as an expression given "in terms of the moment generating function of the underlying state variable". This resolves the first part (a) of the question. Concerning the second part (b), it should be clear by now that only an application of the already existing apparatus from probability theory has been used.

When considering only the intrinsic value, rather then the total price, it is even possible to avoid the introduction of a damping constant a. By a smart application of the put-call parity one again arrives at an expression involving an integral as in (39). Whether damping is employed or not, after inversion of the Fourier transform we are required to solve an integral of the form

$$\int_0^{\infty} e^{-\mathrm{i}vk} \psi_T(v) dv$$

numerically. Since the Fourier transform maps $L^2(\mathbb{R})$ onto $L^2(\mathbb{R})$ itself, ψ_T, as the Fourier-image of c_T, will be square integrable as well. An upper bound for the tail can be easily found from (40) in the case with damping.

$$\left| \int_A^{\infty} e^{-\mathrm{i}vk} \psi_T(v) dv \right| \leq \int_A^{\infty} \frac{1}{v^2} dv < \frac{1}{A}.$$

When the intrinsic value is analyzed, the equivalent of ψ_T can be shown to be a function η_T given by

$$\eta_T(v) = \frac{1 - \phi_T(1 + \mathrm{i}\, v)}{v^2 - v}.$$

It will not be derived here. Again with $\phi_T \leq 1$ we can analogously find that

$$\left| \int_A^{\infty} e^{-\mathrm{i}vk} \eta_T(v) dv \right| \leq \int_A^{\infty} \frac{2}{v^2} dv < \frac{2}{A}.$$

In any case, we can find a truncation constant A, such that a given error bound holds. With a composite trapezoidal rule, for example, we obtain

$$\int_0^\infty e^{-\mathrm{i}vk}\psi_T(v)dv \cong \left(\frac{\psi_T(0)}{2} + \sum_{m=1}^{N-1} e^{-\mathrm{i}v_m k}\psi_T(u_m) + \frac{e^{\mathrm{i}u_N k}\psi_T(u_N)}{2}\right)\Delta u, \quad (42)$$

where $u_m = m\Delta u$ and $\Delta u = A/N$.

In order to analyze smile and skew behavior around the strike K, we will have to compute sums as in (42) for various different values of $k = \ln K$. The "Fast Fourier Transform", abbreviated FFT, is a method or better an algorithm for evaluating such sums in an efficient way, even for more accurate numerical quadrature schemes.

2.3 Wu/Zhang Model

2.3.1 Motivation

Stochastic volatility extensions of LIBOR models intend to accommodate market smiles and skews in predicting a more realistic evolution of future implied volatility. A general form of such a model could be, as Brigo and Mercurio put it,

$$\begin{aligned} dL_i(t) &= a_i(t)\psi(L_i(t))[v(t)]^\gamma dW_i(t), \\ dv(t) &= a_v(t)dt + b_v(t)d\widehat{W}(t)\,, \end{aligned}$$

where time arguments are maintained to indicate stochastic processes. We will be more reserved when possible.

If W_i and $\widehat{W}$ are chosen uncorrelated, it can be shown for specific cases that implied volatility functions have local minima at ATM values. So happened in Renault/Touzi (1996) in case ψ is equal to identity. In this case, $\psi(x) = x$, not even a skew can be expected for ATM options. In case $\psi(x) = x^\alpha$, for $0 < \alpha < 1$, downward-sloping volatility skews are indeed possible, but the resulting smiles do not have all desired properties, namely to be non-monotonic and U-shaped. More precisely, a hockey-stick type of curve is desired. On the other hand, if these latter features were postulated, functions ψ would have to be chosen which would imply quite non-stationary behavior of the volatility smile as a function of forward rate levels. Empirical evidence, however, indicates that both properties for caplet-volatility curves are significant. The local minimum should be at higher than at-the-money rates, namely in deep-in the money regions and stationarity should be ensured.
Overall, the caplet volatility curve, as a function of strike, is postulated to incorporate the following four features:

1. The smile, exhibiting convexity.
2. The skew, having nonzero slope at ATM levels.
3. U-shapedness or hockey-stick form.
4. Stationarity with regard to different forward yield levels.

Several suggestions in order to obtain these properties have been made:

(a) Displaced-diffusion dynamics, that is setting $\psi(x) = x+\alpha$ with constant α,

(b) Choosing a non-linear function ψ, for example, $\psi(x) = x^\alpha$,

(c) Introducing a non-zero (instantaneous) correlation between rates and volatility.

Approach (c) is considered the most promising and therefore is subject of this work. Following it requires, however, solutions to problems we would not have in other cases. The most original work in this direction was probably done by Heston (1993) for an arbitrary asset S. Heston proposes a CIR square-root process for the volatility.

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{v_t} S_t dW_t, \\ dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t} d\widehat{W_t} \end{aligned}$$

where

$$dW_t d\widehat{W_t} = \rho dt$$

indicates the non-zero correlation.

His work involved the observation that the known characteristic function of the underlying variables will provide analytic expressions for option prices once Fourier Transforms are inversed. The characteristic function itself can be derived by solving an associated PDE, which is found by an application of the Feynman-Kac formula.
This approach is more likely to be successful in cases where the characteristic function can be determined easily. Even very simple volatility dynamics, however, can lead to tremendous difficulties for deriving the joint probability density of the variables when these are correlated.
A first application of the non-zero correlation assumption to an interest rate framework was established by Wu/Zhang (2006).

2.3.2 Setup

Other stochastic volatility approaches to LIBOR rates, prior to Wu/Zhang, have been by Andersen/Brotherton-Ratcliffe (2001) and Piterbarg (2005). Opposed to their models, the non-zero correlation assumption in the model by Wu/Zhang generates additional difficulties at measure changes. In an original setup a model asks for the specification of a measure. A change to a different measure will, in general, affect the volatility dynamics, namely shift the drift. This is not the case, when a zero correlation is assumed between W and $\widehat{W}$. In the sequel we will explain, how Wu/Zhang managed the problem. This short summary is hoped to be instructive and illustrating, because later we have to run through a more detailed but similar procedure for our model.

Wu/Zhang specify the dynamics of LIBORS under the risk-neutral measure by

$$
\begin{aligned}
dL_i &= L_i \sum_{k=1}^{i} \frac{\delta L_k}{1+\delta L_k} v\, \gamma_i \cdot \gamma_k dt + L_i \sqrt{v}\, \gamma_i \cdot dW \\
&= L_i \sqrt{v} \gamma_i \cdot [dW - \sqrt{v}\, \sigma_{i+1} dt], \\
dv &= \kappa(\theta - v)dt + \sigma \sqrt{v}\, d\widehat{W},
\end{aligned}
$$

with

$$
\sigma_{i+1} = -\sum_{k=1}^{i} \frac{\delta L_k}{1+\delta L_k} \gamma_k.
$$

Note that they confine to a one-dimensional Brownian motion as volatility process. Furthermore, they allow for correlation in setting

$$
E\left[\left(\frac{\gamma_i(t)}{|\gamma_i(t)|} \cdot dW_t\right) \cdot d\widehat{W}_t\right] = \rho_i(t)dt.
$$

2.3.3 Where Feynman-Kac Comes In

A change to *individual* forward measures is necessary for caplet pricing. Through this transformation the process $v(t)$ gains an extra drift term as indicated by the following proposition.

Proposition 31 *Let W and $\widehat{W}$ be Brownian motions under the risk-neutral measure P. Define $W^{(i+1)}$ and $\widehat{W}^{(i+1)}$ through*

$$
\begin{aligned}
dW^{(i+1)} &= dW - \sqrt{v}\, \sigma_{i+1}\, dt\,, \\
d\widehat{W}^{(i+1)} &= d\widehat{W} + \xi_i(t) \sqrt{v}\, dt\,,
\end{aligned}
$$

where

$$
\xi_i(t) = \sum_{k=1}^{i} \frac{\delta L_k\, \rho_k(t)\, |\gamma_k|}{1+\delta L_k},
$$

then $W^{(i+1)}$ and $\widehat{W}^{(i+1)}$ are Brownian motions under P_{i+1}.

Proof. See Wu/Zhang (2006) ■

In terms of $W^{(i+1)}$ and $\widehat{W}^{(i+1)}$, Wu/Zhang's model alters to

$$
\begin{aligned}
dL_i &= L_i \sqrt{v}\, \gamma_i \cdot dW^{(i+1)}, \\
dv &= [\kappa\theta - (\kappa + \sigma \xi_i(t) v)]\, dt + \sigma \sqrt{v}\, d\widehat{W}^{(i+1)},
\end{aligned}
$$

where $\sigma\xi_i(t)$ is the forementioned, time-dependent, additional driftterm. The benefit of having L_i represented as a martingale with respect to the new measure is partly offset by the disadvantage that the new driftterm of v depends on random variables through the definition of $\xi_i(t)$, namely on $L_k, k \leq i$. When generating paths for a Monte Carlo pricing, some sort of "freezed" approximation is to be considered.
A first simplification is obtained in defining

$$\widetilde{\xi}_i(t) = 1 + \frac{\sigma}{\kappa}\xi_i(t)\,,$$

which gives a nicer equation for v, namely

$$dv = \kappa[\theta - \widetilde{\xi}_i(t)v]\,dt + \sigma\sqrt{v}\,d\widehat{W}^{(i+1)}\,.$$

Since the forward price of a caplet $C_i(t)$ is a martingale under forward measure P_{i+1}, we receive following expression for today's value

$$\begin{aligned} C_i(0) &= P(0,T_{i+1})\,\delta\,E_{i+1}[(L_i(T_i)-K)^+] \\ &= P(0,T_{i+1})\,\delta\,L_i(0)\,G(0,L_i(0),v(0),K)\,. \end{aligned} \qquad (43)$$

We denote

$$G(0,L_i(0),v(0),K) := E_{i+1}\left[\left(\frac{L_i(T_i)}{L_i(0)} - \frac{K}{L_i(0)}\right)^+\right]$$

$$= E_{i+1}[e^{\ln(L_i(T_i)/L_i(0))}\mathbf{1}_{L_i(T_i)>K}] - \frac{K}{L_i(0)}E_{i+1}[\mathbf{1}_{L_i(T_i)>K}].$$

In this form the two expected values can be evaluated in terms of moment generating functions (MGF). Define

$$\phi(X(t),v(t),t;z) := E[e^{zX(T_i)}|\mathcal{F}_t],$$

where $X(t) = \ln(L_i(t)/L_i(0))$ and z is allowed to be complex, $z \in \mathbb{C}$. This convention by Wu/Zhang differs from the usual way to define moment generating functions for reals. Existence of the MGF has to be ensured, in general. It does not pose a problem in our case, because all distributions involved have finite variances. In defining a moment generating function in this way, the characteristic function emerges as a special case. Take z to be purely complex, $z = \mathbf{i}\,u$. Recall that the method of Carr/Madan was derived for characteristic functions.

For simplicity define $\phi_T(z) := \phi(0, V(0), 0; z)$. It has been shown that the expected values above can be ascertained from

$$E_{i+1}[\mathbf{1}_{L_i(T_i)>K}] = \frac{\phi_T(0)}{2} + \frac{1}{\pi}\int_0^\infty \frac{\Im(e^{-\mathrm{i}u\ln(K/L_i(0))}\phi_T(\mathrm{i}\,u))}{u}du,$$

and

$$E_{i+1}[e^{\ln(L_i(T_i)/L_i(0))}\mathbf{1}_{L_i(T_i)>K}] = \frac{\phi_T(1)}{2}+\frac{1}{\pi}\int_0^\infty \frac{\Im(e^{-\mathrm{i}u\ln(K/L_i(0))}\phi_T(1+\mathrm{i}\,u))}{u}du.$$

See Kendall (1994) or more recently Duffie, Pan, Singleton (2000) for the respective results. If ϕ_T is known, the price can be determined from a numerical integration. It turns out that $\phi(x, V, t; z)$ satisfies the Kolmogorov backward equation corresponding to the joint process of forward rate and stochastic factor:

$$\frac{\partial\phi}{\partial t}+\kappa(\theta-\widetilde{\xi}_i v)\frac{\partial\phi}{\partial v}-\frac{1}{2}|\gamma_i|^2 v\frac{\partial\phi}{\partial x}+\frac{1}{2}\sigma^2 v\frac{\partial^2\phi}{\partial v^2}+\sigma\rho_i v|\gamma_i|\frac{\partial^2\phi}{\partial v\partial x}+\frac{1}{2}|\gamma_i|^2 v\frac{\partial^2\phi}{\partial x^2}=0\,,$$

subject to the boundary condition

$$\phi(x, v, T_i\,; z) = e^{zx}.$$

There are two ways to prove this:

1. An application of the Feynman-Kac formula.
2. Apply Itô's Lemma to $\phi(X, v, t; z)$, which is by definition a martingale. Set the resulting drift equal to zero.

Application of the Feynman-Kac formula is demonstrated later in detail, therefore we abstain from doing it here. The partial differential equation can always be solved numerically. In finding an explicit solution, however, we have to require piecewise constant coefficients.

2.3.4 Solution of PDE

If coefficients $\gamma_i(t)$ and $\rho_i(t)$ are time-independent between tenors of a grid, $0 = T_0 < T_1 < \cdots < T_n = T$, we speak of piecewise constant coefficients. For a first simplification, fix t in an arbitrary interval and define

$$\lambda := |\gamma_i(t)| \quad \text{and} \quad \rho := \rho_i(t)$$

on its length. The characteristic function is determined individually for every caplet. Naturally, λ and ρ do not dependent on i. Since $\widetilde{\xi}_i(t)$ depends on stochastic LIBOR rates, some additional argument is needed to make them constant between grid points. Wu/Zhang propose the usual freezing technique. Details are given in their paper. At this stage they assume constant

$$\xi = \widetilde{\xi}_i(t).$$

The PDE is then given by

$$\frac{\partial \phi}{\partial t} + \kappa(\theta - \xi v)\frac{\partial \phi}{\partial v} - \frac{1}{2}\lambda^2 v \frac{\partial \phi}{\partial x} + \frac{1}{2}\sigma^2 v \frac{\partial^2 \phi}{\partial v^2} + \sigma\rho\lambda v \frac{\partial^2 \phi}{\partial v \partial x} + \frac{1}{2}\lambda^2 v \frac{\partial^2 \phi}{\partial x^2} = 0,$$

with terminal condition

$$\phi(x, v, T; z) = e^{zx}.$$

Heston determined a solution for the characteristic function ϕ by the ansatz

$$\widetilde{\phi}(x, v, \tau; z) = e^{A(\tau,z)+B(\tau,z)v+zx},$$

where the substitution $\tau = T - t$ represents time to maturity. Inserted into the PDE, a subsequent comparison of coefficients of the resulting polynomial in v, will reduce the problem. Namely, to the objective to find solutions for two ordinary differential equations

$$\begin{aligned} \frac{dA}{d\tau} &= \kappa\theta B\,, \\ \frac{dB}{d\tau} &= \frac{1}{2}\sigma^2 B^2 + (\rho\sigma\lambda z - \kappa\xi)B + \frac{1}{2}\lambda^2(z^2 - z)\,, \end{aligned} \qquad (44)$$

subject to the initial conditions

$$A(0, z) = 0, \qquad B(0, z) = 0.$$

The second ODE for B in (44) is a Riccati equation, which is known to have an analytical solution for constant coefficients. We summarize the result in a proposition. Naturally $\tau_i = T_i - t$.

Proposition 32 *For piecewise constant coefficients and for $\sigma \neq 0$, equations (44) have for $\tau_i \leq \tau < \tau_{i+1}$, $i = 0, 1, ..., n-1$, a solution of the form*

$$A(\tau, z) = A(\tau_i, z) + \frac{\kappa\theta}{\sigma^2}\left((a + d)(\tau - \tau_i) - 2\ln\left(\frac{1 - g_i e^{d(\tau-\tau_i)}}{1 - g_i}\right)\right),$$

$$B(\tau, z) = B(\tau_i, z) + \frac{(a + d - \sigma^2 B(\tau_i, z))(1 - e^{d(\tau-\tau_i)})}{\sigma^2(1 - g_i e^{d(\tau-\tau_i)})},$$

where

$$a = \kappa\xi - \rho\sigma\lambda z, \qquad d = \sqrt{a^2 - \lambda^2\sigma^2(z^2 - z)}, \qquad g_i = \frac{a + d - \sigma^2 B(\tau_i, z)}{a - d - \sigma^2 B(\tau_i, z)}.$$

An explicit solution of the above PDE allows direct computation of caplet prices by (43). On the other hand, given prices, an explicit form for ϕ enables calibration of parameters.

Remark 33 *The crucial assumption in the above solution ansatz is that the parameter functions are piecewise constant. The reasonability of this assumption may be questioned. Wu/Zhang consider it to be "by no means a restriction". Some practioners, however, including the author, believe it to be very restrictive and even unrealistic. What is more, it causes untolerable instabilities in the calibration. This topic has already been discussed in Section 1.2.2. Therefore, depending on the user's stance, assumptions about the form of the instantaneous parameter functions should be carefully scrutinized. In any case, if the parameters of the resulting PDE can be considered constant, the Fourier inversion method is an acceptable and valuable tool. If they are not, alternative methods have to be looked for.*

2.4 Zhu Model

In his working paper Zhu (2007) immediately considers the dynamics of forward LIBORS $L_i(t)$ under the forward measure P_{i+1}

$$dL_i(t) = v_i(t)dW_i(t), \quad t \le T_i, \quad 1 \le i \le n-1.$$

He proposes two possible dynamics for a volatility process v, respectively variance process V, but elaborates only on the first. The alternatives are between a mean-reverting Ornstein-Uhlenbeck process and the square-root process by Cox, Ingersoll and Ross.

1. $dv_i = \kappa_i(\theta_i - v_i)dt + \sigma_i dW,$
2. $dV_i = \kappa_i(\theta_i - V_i)dt + \sigma_i v_i dW,$

with $V = v^2$.

Note that $n-1$ stochastic volatilities v_i, respectively variances V_i, are introduced, which are all driven by the same one-dimensional Wiener process. As such, this idea generalizes the model by Wu/Zhang, where a one-dimensional CIR-process perturbs all LIBORS. This is a first fundamental difference to our model. We propose to consider as many CIR-processes as we have LIBORS, where each is perturbed by its own source.

Zhu's model can not be reduced to a special case from the model we introduce later. Not surprisingly, this is mainly due to the fact that all square-root processes in his model are driven by a one-dimensional W only. At first sight, it appears to be simpler. But he obtains a fully restructured LIBOR model setup. In particular, all information from the standard model are lost, namely the γ_i. Furthermore, it remains unclear, how a stable calibration can be obtained.
(For later reference: In the terminology of our model, let $r = 1$, $\widehat{W} = W$ and set the dimension of the extended U-space to $d = m = n-1$. This is as close as we can possibly come to Zhu's model.)

Two empirical observations:

1. Smile structures of different caplets display similar patterns. It is therefore not very restrictive to have just one driving force for the volatility processes. This is a comprehensible argument. Nevertheless, our model offers additional variety.

2. Due to its gaussianity the Ornstein-Uhlenbeck process allows negative yields. Although such negative rates have occasionally occurred in Switzerland, the process has been avoided for interest rate modeling. For the modeling of volatilities it seems even more inappropriate to use a gaussian process.

To keep the number of parameters small, Zhu introduces deterministic "nesting functions" for each of the parameter sets: reversion-speed, -level, initial volatility, volatility of variance and correlation.
These are in particular:

$$\begin{aligned}
\kappa_i &= \kappa \exp aT_i, \\
\theta_i &= \theta + (b_1 T_i + b_2) \exp b_3 T_i, \\
v_i(0) &= v + (c_1 T_i + c_2) \exp c_3 T_i, \\
\sigma_i &= \sigma \exp dT_i, \\
\rho_i &= e_1(1 - \exp e_2 T_i).
\end{aligned}$$

These nesting functions reduce the total number of parameters from $5(n-1)$ to just 14 and supposedly still capture a humped cap volatility structure with adequate choices of θ_i and $v_i(0)$ (Brigo and Mercurio, 2006). With nesting functions defined and designed in this fashion, not only a parsimonious model is accomplished, but another desired feature is exhibited: time homogeneity. That much to the positive aspects. But what happens if one believes in nested structures of a different form?

When analyzing Zhu's model, the probably most striking criticism one could bring in is the existence of a fundamental calibration error.
Zhu's "closed-form pricing" solution, involving the characteristic function, is only available under individual forward measures. Specifying the whole model under only one, for example the terminal measure, will destroy this pleasant property. Zhu's "joint calibration" is performed simultaneously over all parameters, therefore their estimates are inconsistent with the one that would result if calibrated under the terminal measure.

Following this approach he, in effect, ignores the drift term required by the arbitrage-free condition. An interesting thought, admittedly. The deviations are expected to be small, so why not disregard this last corset, after "no transaction-cost" and "complete market" assumptions have already been dropt. In the end, departments in business schools have implemented research activity in the latter two subjects only a few years ago, meanwhile arbitrage desks have had their entitlement for existence in banks from the beginning of trading.

2.5 Piterbarg Model

Another stochastic volatility model manages to create not only a smile in the implied volatility structure, but also the observed slope in it. Piterbarg (2005) sticks to the assumption of zero correlation between the two Brownian motions W and $\widehat{W}$, which ensures that the dynamics of the v_i remain unchanged under the different measures. His model postulates that forward rates evolve, under a generic measure P, according to

$$\begin{aligned} dL_i &= \sqrt{v}\,[b_i(t)L_i + (1-b_i(t))L_i(0)]\,\underline{\sigma}_i(t)\cdot\left(\sqrt{v}\,\underline{\mu}(t)dt + dW\right), \\ dv &= \kappa(v(0)-v)dt + \eta\sqrt{v}\,d\widehat{W}, \end{aligned}$$

where W is a vector of independent P-Brownian motions, σ_i's are deterministic vector functions, κ and η are positive constants, b_i's are deterministic scalar functions, $\widehat{W}$ is another P-Brownian motion uncorrelated with W, $dWd\widehat{W} = 0$, and μ is a suitable adapted vector process, i.e. a numeraire-specific drift that ensures lack of arbitrage within the model.
Piterbarg proposes this model as a generalization of the well-known displaced-diffusion stochastic volatility model, see Andersen/Andreasen (2002), which is arrived at, if σ_i and b_i are chosen constant. The latter model already generates both, smile and skew, in the cap market; Piterbarg, however, wants to improve the fit with respect to the swaption grid. The time-homogeneity of parameters in Andersen/Andreasen (2002) reduces the models ability to reproduce the empirically observed term structure of volatility. With the additional flexibility introduced by the functions b_i Piterbarg's model allows a better fit.

In Wu/Zhang (2004) slope and skew are controlled by correlations between volatility and rates. In Piterbarg's model versatility is accomplished by the functions b_i. We here recap the main results. Thereby we use a more *general* model in the sense that it is applicable to various markets not just interest rates. It will be more *restricted*, however, since we consider a one-dimensional Brownian for the underlying. In a first step the above mentioned versatility is immediately restrained by a notion of "parameter averaging", how Piterbarg terms it. He seeks for some "effective" constant skew and volatility parameters b and λ so that the (terminal) distribution of a process of the form

$$dS = \sqrt{v}\,[\beta(t)S + (1-\beta(t))S(0)]\,\sigma(t)dW, \tag{45}$$

where we set $\mu(t) = 0$ without loss of generality, can be approximated by the distribution of a process

$$d\bar{S} = \sqrt{v}\,\left[b\bar{S} + (1-b)\bar{S}(0)\right]\lambda dW. \tag{46}$$

The underlying idea is thus to match a time-dependent local volatility function to a time-homogeneous one. Aware of the fact that term stochastic volatility parameters b and λ are already determined from vanilla options for different maturities, this idea bears considerable advantages for the calibration. Since $\sigma(t)$ and $\beta(t)$ will be directly inferred from b and λ without making a detour over the actual european option prices, a higher numerical efficiency is to be expected.
Piterbarg shows that the 'effective' skew over a time horizon $[0, T]$ – the skew thus actually depends on T – for the system

$$\begin{aligned} dS &= \sqrt{v}\, [\beta(t)S + (1-\beta(t))S(0)]\, \sigma(t) \cdot dW, \\ dv &= \kappa(v(0) - v)dt + \eta\sqrt{v}\, d\widehat{W}, \end{aligned}$$

is given by

$$b = \int_0^T \beta(t) w(t) dt, \tag{47}$$

where the weights $w(\cdot)$ are computed from

$$w(t) = \frac{\nu^2(t)\sigma^2(t)}{\int_0^T \nu^2(u)\sigma^2(u)du},$$

with

$$\nu^2(t) = v^2(0)\int_0^t \sigma^2(s)ds + v(0)\eta^2 e^{-\kappa t}\int_0^t \sigma^2(s)\frac{e^{\kappa s} - e^{-\kappa s}}{2\kappa}ds.$$

Remark 34 *Because ν depends on parameters κ and η, which affect the dynamics of the square root process but have not yet be determined, we have to retreat to special cases in finding $\beta(t)$.*

The deterministic volatility function $\sigma(t)$ needs to be derived from knowledge of the 'effective' volatility λ *and* the 'effective' skew b. It has to be chosen such that

$$\varphi_0\left(-\frac{g''(\zeta)}{g'(\zeta)}\lambda^2\right) = \varphi\left(-\frac{g''(\zeta)}{g'(\zeta)}\right)$$

is satisfied. φ_0 and φ are the Laplace transforms of the random variables

$$\int_0^T v(t)dt \quad \text{and} \quad V(T) = \int_0^T \sigma^2(t)v(t)dt,$$

respectively. We have

$$\zeta = E(V(T)) = \int_0^T \sigma^2(t)E(v(t))dt = v_0\int_0^T \sigma^2(t)dt.$$

The function g is of the form

$$g(x) = \frac{S_0}{b}\left(2\mathcal{N}(b\sqrt{x}/2) - 1\right),$$

with $\mathcal{N}$ denoting the standard normal distribution. Piterbarg considers an approximation from the second-order Taylor expansion

$$g(x) \approx a + be^{-cx}.$$

The *given* effective volatility λ and the effective skew b stem from a first calibration to market options. The practical idea of Piterbarg is thus to take advantage of fitting a simpler model (46). Since that is a known exercise, the optimization is numerically quite efficient. The calibration to the more demanding model (45) should be performed along the following steps:

1. Due to the interdependence of λ and b, in a first step fix b at a reasonable value, for example take the average over all values determined in the calibration of (46).

2. Obtain $\sigma(t)$ from fitting to λ over all times T.

3. Determine $\beta(t)$ by (47) from fitting to b over all times.

4. If necessary, recalibrate.

Details are given in Piterbarg (2005).

2.6 Andersen, Brotherton-Ratcliffe Model

A predecessor of Piterbarg's stochastic volatility model is proposed by Andersen and Brotherton-Ratcliffe (2001). Their primary interest is to obtain efficient approximations to caplet prices, which is a more demanding exercise in a more general model. The dynamics of forward LIBORS and a volatility process under a measure P are given by

$$\begin{aligned} dL_i &= \sqrt{v}\,\varphi(L_i)\sigma_i(t)\left(\sqrt{v}\,\mu_i(t)dt + dZ_i\right), \\ dv &= \kappa(\theta - v)dt + \epsilon\psi(v)\,d\widehat{W}, \end{aligned} \tag{48}$$

where in their formulation σ_i are deterministic *scalar* functions and Z_i are dependent *scalar* P-Brownian motions, independent of another P-Brownian motion $\widehat{W}$.

$$dZ_i dZ_k = \rho_{ik}\,dt, \quad i,k = 1,...,n-1,$$
$$dZ_i d\widehat{W} = 0, \qquad i = 1,...,n-1.$$

With this setup the authors choose to represent a LIBOR market model with dependent components Z_i. Recall Section 1.2.1 for details. Again μ_i are numeraire-specific, adapted drift-processes that ensure lack of arbitrage within the market. Note that in this representation μ_i are scalar. The constants κ, θ and ϵ are positive.

In Section 1.2.1 we have seen that with the help of a Cholesky decomposition, a representation of the form

$$dZ_i = \sum_{k=1}^{n-1} f_{i,k}\, dW_k = f_i \cdot dW$$

can be given, with now independent W_k's and unit vectors $f_i \in \mathbb{R}^{n-1}$. Defining a new vector $\underline{\sigma}_i(t) := \sigma_i(t) f_i \in \mathbb{R}^{n-1}$ and stating the dynamics under measure P_{i+1}, we obtain a striking similarity to the Piterbarg model.

Andersen/Brotherton-Ratcliffe:

$$\begin{aligned} dL_i &= \sqrt{v}\, \varphi(L_i)\underline{\sigma}_i(t) \cdot dW^{(i+1)}, \\ dv &= \kappa(\theta - v)dt + \epsilon\psi(v)\, d\widehat{W}^{(i+1)}, \end{aligned}$$

where $\underline{\sigma}_i(t) \in \mathbb{R}^{n-1}$.

Piterbarg:

$$\begin{aligned} dL_i &= \sqrt{v}\, [b_i(t)L_i + (1 - b_i(t))L_i(0)]\, \underline{\sigma}_i(t) \cdot dW^{(i+1)}, \\ dv &= \kappa(v(0) - v)dt + \eta\sqrt{v}\, d\widehat{W}^{(i+1)}, \end{aligned}$$

with again $\underline{\sigma}_i(t) \in \mathbb{R}^{n-1}$.

None of the models can be interpreted as a pure generalization of the other. The skew function φ can not produce the corresponding skew factor in the Piterbarg model, which needs an additional unknown function b_i. On the other hand, the arbitrariness of the function ψ demonstrates that in the Andersen/Brotherton-Ratcliffe model a more general volatility process than in Piterbarg's model can be generated. But the intention persued by Andersen/Brotherton-Ratcliffe is a different anyway.

The technical contribution of Piterbarg's paper is a "Markov-projection" onto a simpler process from which an efficient calibration could be initiated. From there accurate european option prices under general time-dependent parameters, and not only piecewise constant as in Wu/Zhang (2002), could

be obtained.
The motivation of Andersen/Brotherton-Ratcliffe, on the other hand, is to explore *asymptotic* expansion techniques to provide closed-form pricing formulas for caps and swaptions. Therefore they introduce rather arbitrary functions φ and ψ for which exact formulas can not be given.

The approach of Andersen/Brotherton-Ratcliffe stipulates to observe that the caplet price at time t,

$$C_j(t;K) = \delta_j B_{j+1}(t) E_{t,j+1}(L_j(T_j) - K)^+ = \delta_j B_{j+1}(t) G(t; L_j(t), v(t)),$$

can be interpreted as a function $G(t; L_j, v)$ that, by Feynman-Kac, must satisfy the PDE

$$\frac{\partial G}{\partial t} + \kappa(\theta - v)\frac{\partial G}{\partial v} + \frac{1}{2}\epsilon^2\psi^2(v)\frac{\partial^2 G}{\partial v^2} + \frac{1}{2}\varphi^2(L_j)v|\underline{\sigma}_j(t)|^2\frac{\partial^2 G}{\partial L_j^2} = 0.$$

Dropping the subscript j and defining a scalar $\sigma(t) := |\underline{\sigma}_j(t)|$ we thus have to solve

$$\frac{\partial G}{\partial t} + \kappa(\theta - v)\frac{\partial G}{\partial v} + \frac{1}{2}\epsilon^2\psi^2(v)\frac{\partial^2 G}{\partial v^2} + \frac{1}{2}\varphi^2(L)v\sigma^2(t)\frac{\partial^2 G}{\partial F^2} = 0, \tag{49}$$

subject to $G(t; L, v) = (L - K)^+$.

Andersen/Brotherton-Ratcliffe tackle the problem in first considering v to be constant. By scaling, without loss of generality, it can then even be assumed $v = 1$. The PDE then reduces to

$$\frac{\partial G}{\partial t} + \frac{1}{2}\varphi^2(F)\sigma^2(t)\frac{\partial^2 G}{\partial L^2} = 0\,, \tag{50}$$

subject to $G(t; L, 1) = (L - K)^+$.
Note that the additional assumption $\varphi \equiv id$ delivers the well-known Black price, as can also immediately be seen from (48). It is therefore a tempting idea to take this solution and replace the constant implied volatility in it by some asymptotic expansion of the implied volatility in powers of $\tau = T - t$. For that reason introduce a new variable

$$x = \frac{1}{\tau}\int_t^T \sigma^2(s)ds$$

and suppose the solution has the form

$$G(t;L,1) = g\left(t;L,\frac{1}{\tau}\int_t^T \sigma^2(s)ds\right)$$

$$g\,(t;L,x) = L\mathcal{N}(d_+) - K\mathcal{N}(d_-)$$

$$d_\pm = \frac{\ln\left(L/K\right) \pm \frac{1}{2}\Omega^2(t;L,x)}{\Omega(t;L,x)}.$$

The merit of Andersen/Brotherton-Ratcliffe is to have shown that such an expansion is given by

$$\Omega(t;L,x) = \sum_{i\geq 0}(x\tau)^{i+1/2}\Omega_i(L),$$

and that the first two terms of the expansion

$$\Omega(t;L,x) = \Omega_0(L)(x\tau)^{1/2} + \Omega_1(L)(x\tau)^{3/2} + O(\tau^{5/2})$$

serve as a good enough approximation in practical applications. The functions Ω_i, which are independent of τ, are arrived at by substituting the assumed form of $g\,(t;L,x)$ into the PDE (50) and comparing coefficients of the same order. The first two are then found by imposing a finite limit for $L \to K$ and solving ordinary Bernoulli type equations:

$$\Omega_0(L) = \frac{\ln\left(L/K\right)}{\int_K^L \varphi^{-1}(u)du}$$

$$\Omega_1(L) = -\frac{\Omega_0(L)}{\left(\int_K^L \varphi^{-1}(u)du\right)^2}\ln\left(\Omega_0(L)\left(\frac{LK}{\varphi(L)\varphi(K)}\right)^{1/2}\right).$$

In the next stage Andersen/Brotherton-Ratcliffe introduce stochastic volatility as proposed in the dynamics of the SDE (48). Referring back to the Hull and White (1987) decomposition result, the function G solving (49) can be written as

$$G(t;L,v) = E_t[g(t;L,\tau^{-1}U(T)],$$

where

$$U(T) := \int_t^T \sigma^2(s)v(s)ds.$$

The result of Hull and White (1987) is based on the independence of underlying and volatility process. It is the application of the decomposition result that requires the independence $dZ_i d\widehat{W} = 0$. The following treatise of expanding the above expectation and the subsequent calculation of the resulting higher moments of $U(T)$ by means of analytical approximations for the Laplace tranform of the density of U, are rather involved and will not be repeated here.

Though with different approaches, for the sake of analytical tractability Andersen/Brotherton-Ratcliffe, as well as Piterbarg before, assume that the Brownian motions are uncorrelated.

In summary:

Model	Contributor	Skew	Smile
CEV, $\varphi(x) = x^\gamma$	Andersen/Andreasen (2000)	yes	no
DD , $\varphi(x) = x + \beta$	Rebonato/Joshi (2001)	yes	no
DD & SV	Andersen/Andreasen (2002)	yes	yes
SV & CEV & DD, $\rho = 0$	Andersen/Broth.-Ratcl.(2001)	yes	yes
SV, $\rho \neq 0$	Wu/Zhang (2006)	yes	yes
SV, $\varphi(b(t), L_j(t))$, $\rho = 0$	Piterbarg (2005)	yes	yes

Table 2:

3 A Multiple Stochastic Volatility Libor Market Model

3.1 Some Features of the Model

In a recent working paper Schoenmakers/Belomestny (2006) propose a jump-diffusion LIBOR model which extends the classical LIBOR market model in two aspects. They not only introduce a jump-diffusion term, but, secondly, allow its proportion to either be specified directly by the user or determined by an optimization. Through this additional term in its setup the model can be regarded as a perturbation of a standard LIBOR market model.
Two new concepts are introduced by their proceedings. The first idea is to "take away a little" off a standard or classical model and exchange that little by a new component. The second innovation consists in the requirement that the local covariance structures of standard and perturbed model shall be the same. This latter assumption helps to considerably reduce computation complexity in the first place, and, secondly, to determine the newly implemented parameters in a more stable way.
These two key ideas are reintegrated in the model construction we explore in this work, only this time we apply it to a stochastic volatility model. Working "close" to a classical model thereby not only endows us with confidence that we are not too far from "the real thing", but also with an initial and field-tested estimate of instantaneous volatility and correlation structures.

Specifically, we consider an additive component involving a CIR process for the volatility and hence the total construction resembles a Heston type model. The CIR process, as proposed by Cox, Ingersoll, Ross (1985), was originally introduced to explain short term interest rates in an equilibrium framework. We do not further elaborate on its underlying idea, but rather intend to profit from the outstanding performance of Heston models in practical applications to stock and currency markets, and see whether it helps to improve models in the interest rate environment.

The idea of utilizing a Heston type process has already been formulated in a paper by Wu/Zhang (2006) and in a working paper by Zhu. Here are the differences between their approaches and ours:

1. We propose a multi-dimensional partial-Gaussian and partial-Heston type model, where each forward LIBOR is driven by linear combinations of independent CIR volatility processes. As such, we introduce a *vector* volatility process, as opposed to the single dimensional by Wu/Zhang,

with independent sources of uncertainty, as opposed to just one source as in Zhu's model.

2. In both other models it remains unclear whether the optimization is robust. This calibration is the most interesting problem from a practitioners point of view. More importantly, the calibration should be "robust", that is the parameters should be estimable in a unique way, in a certain sense. Robustness is a major problem that is often underestimated. Does it make sense at all to work with a model which does not allow a robust determination of its parameters? Imagine the standard Black/Scholes model would not deliver a unique implied volatility for the stock option. Already in this simple case, we have an inverse problem with a hopefully unique solution. The fact that it is unique contributes a significant part to the Black/Scholes model's success. Newton methods or any other numerical scheme will find a unique minimum.

3. We introduce a multitude of new parameters. At first sight it may appear questionable whether this bunch can be handled at all. Indeed, more parameters lead always to better fitting results, which does not mean that one has obtained a better model. So, lacking "parsimony" is a criticism that our model has to withstand. But, as demonstrated in the sequel, it is exactly the quest for robust routines that forces us to implement so many parameters. In earlier attempts one tried to press very differently moving forwards into only one single volatility "corset". Instabilities are most likely the consequence. The parameters introduced here are neccesary and just enough.

3.2 Dynamics of the Model

Consider a fixed sequence of equidistant tenor dates $0 =: T_0 < T_1 < ... < T_n$, called a tenor structure, where the distance is the so called day-count fraction $\delta := T_{i+1} - T_i\,, i = 1, ..., n-1$. With respect to this tenor structure we consider zerobond processes $B_i\,, i = 1, ..., n$, where each B_i lives on the interval $[0, T_i]$ and ends up with its face value $B_i(T_i) = 1$. As before, we deduce from these a system of forward LIBOR rates defined through

$$L_i = \frac{1}{\delta}\left(\frac{B_i(t)}{B_{i+1}(t)} - 1\right)\,, \quad 0 \leq t \leq T_i\,, \quad 1 \leq i \leq n-1. \tag{51}$$

Note that L_i is the annualized effective forward rate to be contracted at the date t, for a loan over a forward period $[T_i, T_{i+1}]$. Based on this rate one has to pay at T_{i+1} an interest amount of $\$\,\delta L_i(T_i)$ on a $\$$ 1 notional.

For a pre-specified volatility process $\gamma_i \in \mathbb{R}^m$, adapted to the filtration generated by some standard Brownian motion $W \in \mathbb{R}^m$, the dynamics of the corresponding LIBOR model have the form,

$$\frac{dL_i}{L_i} = (...)\,dt + \gamma_i \cdot dW, \quad i = 1, ..., n-1. \tag{52}$$

The drift term, adumbrated by the dots, is known under different measures, such as the risk-neutral, spot, terminal and all measures induced by individual bonds taken as numeraire. Recall that, for example, under the terminal measure P_n we have

$$\frac{dL_i}{L_i} = -\sum_{j=i+1}^{n-1} \frac{\delta_j\, L_j\, \gamma_i \cdot \gamma_j}{(1 + \delta_j L_j)} dt + \gamma_i \cdot dW.$$

If the processes $t \to \gamma_i(t)$ in (52) are deterministic, one speaks of a LIBOR *market* model.

In this work we study extensions of a LIBOR market model, which is given via a deterministic volatility structure γ, with respect to an extended Brownian filtration. These extensions have the following structure,

$$\frac{dL_i}{L_i} = (...)dt + \sqrt{1 - r_i^2}\gamma_i \cdot dW + r_i \beta_i \cdot dU, \quad 1 \leq i < n\,, \tag{53}$$

$$dU_k = \sqrt{v_k} d\widetilde{W}_k\,, \quad 1 \leq k \leq d\,,$$

$$dv_k = \kappa_k(\theta_k - v_k)dt + \sigma_k \sqrt{v_k}\left(\rho_k d\widetilde{W}_k + \sqrt{1 - \rho_k^2} d\overline{W}_k\right), \tag{54}$$

again for $k = 1, ..., d$. $\widetilde{W}$ and $\overline{W}$ are mutually independent d-dimensional standard Brownian motions, both independent of W. In (53), $\beta_i \in \mathbb{R}^d$ are arbitrary deterministic vector functions. They will be specified later. The r_i are constants that may be considered "allotment" or "proportion" factors, quantifying how much of the original input model should be in play. For $r_i = 0$, for all i, it is easily seen from (53) that the classical market model is retrieved. As such the extended model may be regarded as a perturbation of the former.

Since we intend to fully specify the dynamics of this extended model under the terminal measure P_n, the most challenging part of this section is the determination of the drift. Before we tackle this problem two subparagraphs are due. In the first we give an alternative formulation of the model which may help to receive a better intuition for the stochastic dependencies and the proper embedding of the bond structure. The second covers a parameter redundancy that can be solved by a transformation reducing the parameter space by d parameters.

3.2.1 Alternative Formulation

In the above specification of the model in (53) and (54) we sought a setup with independent Brownian motions, $W \in \mathbb{R}^m$, $\widetilde{W} \in \mathbb{R}^d$ and $\overline{W} \in \mathbb{R}^d$. This is indeed more appropriate regarding the required independent generation of random numbers in the final Monte Carlo analysis when derivatives are priced. However, it may serve to clarify stochastic dependencies, if we set

$$d\widehat{W}_k := \rho_k d\widetilde{W}_k + \sqrt{1-\rho_k^2}\, d\overline{W}_k\,, \quad k = 1, ..., m.$$

We then obtain

$$\frac{dL_i}{L_i} = (...)dt + \sqrt{1-r_i^2}\gamma_i \cdot dW + r_i\beta_i \cdot dU$$

$$dU_k = \sqrt{v_k}\, d\widetilde{W}_k$$

$$dv_k = \kappa_k(\theta_k - v_k)dt + \sigma_k\sqrt{v_k}\, d\widehat{W}_k$$

$$d\widetilde{W}_k \cdot d\widehat{W}_k = \rho_k dt, \quad k = 1, ..., d,$$

Again $\widetilde{W}$ and $\widehat{W}$ are both chosen to be independent of W. Likewise the k-th component $\widetilde{W}_k$ should be independent of $\widetilde{W}_l$ and $\widehat{W}_l$, for $l \neq k$. Consequently,

$\widehat{W}_k$ is independent of $\widetilde{W}_l$ and $\widehat{W}_l$, for $l \neq k$.
In this setup the stochastic dependencies are therefore not that easily described, however, the last equation nicely confirms the non-zero correlation between same components.

It is helpful to think of the extended LIBOR model as a vector-valued stochastic differential equation of dimension $n - 1 + d$ with $m + 2d$ factors in the diffusion term. Rewrite

$$\begin{aligned}
\frac{dL_i}{L_i} &= (...)dt + \Gamma_i \cdot d\mathcal{W}, \quad \text{for} \quad i = 1, ..., n-1 \\
dv_1 &= \kappa_1(\theta_1 - v_1)dt + \sigma_1\sqrt{v_1}d\widehat{W}_1, \\
dv_2 &= \kappa_2(\theta_2 - v_2)dt + \sigma_2\sqrt{v_2}d\widehat{W}_2, \\
&... \\
dv_d &= \kappa_d(\theta_d - v_d)dt + \sigma_d\sqrt{v_d}d\widehat{W}_d,
\end{aligned}$$

with

$$\Gamma_i = \begin{pmatrix} \sqrt{1-r_i^2}\gamma_{i1} \\ \cdot \\ \cdot \\ \sqrt{1-r_i^2}\gamma_{im} \\ r_i\beta_{i1}\sqrt{v_1} \\ \cdot \\ \cdot \\ r_i\beta_{id}\sqrt{v_d} \end{pmatrix} \qquad d\mathcal{W} = \begin{pmatrix} dW_1 \\ \cdot \\ \cdot \\ dW_m \\ d\widetilde{W}_1 \\ \cdot \\ \cdot \\ d\widetilde{W}_d \end{pmatrix}. \tag{55}$$

Arbitrage conditions are formulated in terms of coefficient functions in Itô stochastic differential equations (SDE), which represent the dynamics of traded assets. For an embedding into the corresponding zerobond processes, we thus need to specify an SDE describing them. Consider a vector-valued stochastic process X satisfying the following Itô SDE

$$\frac{dX}{X} = \mu(t, X(t))dt + \Sigma(t, X(t))d\mathbb{Z},$$

where $X \in \mathbb{R}^{n+d}$, $\mu \in \mathbb{R}^{n+d}$ and $\Sigma \in \mathbb{R}^{(n+d)\times(m+2d)}$. Furthermore

$$\mathbb{Z} := (W_1, ..., W_m, \widetilde{W}_1, ..., \widetilde{W}_d, \overline{W}_1, ..., \overline{W}_d).$$

The dynamics of the corresponding n zerobonds $B_i(t)$ will be described by the first n components. The remaining d components contain the CIR volatility processes. Many entries of Σ, such as the upper $d \times d$-matrix, will be zero. More specifically, defining $B := (B_1, ..., B_n)$ and $v := (v_1, ..., v_d)$, the above SDE can be written

$$\frac{dB_i}{B_i} = \mu_i(t, B, v)dt + \sum_{j=1}^{m+d} \overline{\sigma}_{ij}(t, B, v)dZ_j, \quad i = 1, ..., n \tag{56}$$
$$dv_1 = \kappa_1(\theta_1 - v_1)dt + \sigma_1\sqrt{v_1}d\widehat{W}_1,$$
$$dv_2 = \kappa_2(\theta_2 - v_2)dt + \sigma_2\sqrt{v_2}d\widehat{W}_2,$$
$$...$$
$$dv_d = \kappa_d(\theta_d - v_d)dt + \sigma_d\sqrt{v_d}d\widehat{W}_d.$$

Note that instantaneous volatilities $\overline{\sigma}_{ij}$ of zerobonds are denoted by an upper bar to distinguish them from volatility parameters $\sigma_k, k = 1, ..., d$, in the square root diffusions.

Along the lines presented in Section 1.1.1, one may in principle determine the drift from this SDE and (51) by an application of the general Itô Lemma. This is, however, very involved and can fortunately be concluded by easier arguments that have been proved.

3.2.2 Parameter Redundancy

A closer look at the vector volatility process v in (53) will disclose an over-parametrization of the model. In the way we defined the process, an interpretation of parameters θ, κ and σ as mean reversion level, speed and "volatility of volatility" is allowed. A transformation of the process may obscure this interpretation. Nevertheless, we can dispose an m-dimensional parameter vector already at this early stage of analysis by normalization. It will certainly facilitate the calibration procedures.

For this purpose consider for each component the scaled process $\widetilde{v}_k := \alpha(k)v_k$, for $k = 1, .., m$, with normalizing constants $\alpha(k)$. Its dynamics are

$$\begin{aligned} d\widetilde{v}_k &= \kappa_k(\alpha(k)\theta_k - \alpha(k)v_k)dt + \sigma_k\alpha(k)\sqrt{v_k}d\widehat{W}_k \\ &= \kappa_k(\alpha(k)\theta_k - \widetilde{v}_k)dt + \widetilde{\sigma}_k\sqrt{\widetilde{v}_k}d\widehat{W}_k, \end{aligned}$$

where we define $\widetilde{\sigma}_k := \sigma_k\sqrt{\alpha(k)}$, for $k = 1, ..., m$.

If we now choose $\alpha(k) := 1/\theta_k$, we obtain

$$d\widetilde{v}_k = \kappa_k(1 - \widetilde{v}_k)dt + \widetilde{\sigma}_k\sqrt{\widetilde{v}_k}d\widehat{W}_k. \tag{57}$$

By setting $\beta_{ik} =: \sqrt{\alpha(k)}\widetilde{\beta}_{ik}$, we see that we indeed have the degrees of freedom to choose the normalizing constants $\alpha(k)$ the way we did. After normalization we retrieve exactly the same model. Without loss of generality, we therefore set $\theta_k = 1$ for all $k = 1, ..., d$.

Remark 35 *After the calibration of $\widetilde{\sigma}_k := \sigma_k/\sqrt{\theta_k}$ for $k = 1, ..., m$, we will not be able to recover the original σ_k and θ_k. However, for the latter analysis these will not be needed. Their exact value is of no interest for the valuation of derivatives by Monte Carlo methods. In the sequel we switch back to the non-tilde notation.*

Remark 36 *The stochastic volatility process v is stationary. It is therefore natural to take the limit expectation as the starting value of the process $v_k(0) = \theta_k$. In this case we obtain consistently*

$$\widetilde{v}_k(0) = \alpha(k)v_k(0) = \alpha(k)\theta_k = 1.$$

See Section 1.5.1 for details regarding the distribution of v.

3.2.3 The Drift under P_n

We intend to work mainly in the terminal measure P_n. In specifying the dynamics of a model given in terms of an Itô diffusion process, we have to determine its drift for the prespecified volatility structure. Recall that we assume an arbitrage-free zerobond market in which the existence of a state price deflator is ensured. By taking the last bond B_n as numeraire we switch into the terminal measure. Under P_n all deflated zerobonds are martingales. Consequently, we can deduce conditions on the bond drifts which in the follow translate into a condition for the i^{th}-LIBOR drift. A major result employed in the actual derivation of the drift is the application of Itô's Lemma on a vector-valued stochastic differential equation.

Lemma 37 *Under the terminal measure P_n the drift of the i^{th}-LIBOR in (53) is given by*

$$-\sum_{j=i+1}^{n-1} \frac{\delta L_j}{1 + \delta L_j}\left(\sum_{k=1}^{m+d} \Gamma_{jk}\Gamma_{ik}\right) = -\sum_{j=i+1}^{n-1} \frac{\delta L_j}{1 + \delta L_j}\Gamma_j \cdot \Gamma_i.$$

Proof. The above drift can immediately be derived from formula (41) in Theorem 7 on LIBOR market models from Jamshidian (2001). All assumptions required for its application are satisfied. The author proves the results on the dynamics of processes in a more general setting, namely for arbitrary semimartingales.

For a more instructive version consider the SDE (56) of traded assets. Applying Itô's Lemma for semimartingales results in the same arbitrage condition that we have encountered before in (7). Here namely

$$L_i \Gamma_i = \delta^{-1} (1 + \delta L_i) (\sigma_i - \sigma_{i+1}).$$

As in Section 1.1.2, this leads to

$$dW^{(i+1)} = dW^{(n)} - \sum_{j=i+1}^{n-1} \frac{\delta L_j \Gamma_j}{1 + \delta L_j} dt,$$

which proves the lemma.
One may also apply Itô's Lemma in its version for diffusion processes. Most partial derivatives involving components of the volatility part $v = (v_1, ..., v_d)$ vanish and thus result in a similar arbitrage condition. ■

The considerations above identify the process under the terminal measure as

$$\frac{dL_i}{L_i} = - \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1 + \delta L_j} \Gamma_j \cdot \Gamma_i \, dt + \Gamma_i \cdot d\mathcal{W}^{(n)}, \qquad 1 \leq i \leq n-1, \tag{58}$$

where the entries of Γ_i are given in (55). More precisely

$$\frac{dL_i}{L_i} = - \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1 + \delta L_j} \left(\sum_{k=1}^{m+d} \Gamma_{jk} \Gamma_{ik} \right) dt + \Gamma_i \cdot d\mathcal{W}^{(n)} \qquad 1 \leq i \leq n-1.$$

It is often technically more convenient to work with the dynamics of the natural logarithm of LIBORS, i.e. $\ln L_i$. In the sequel of this work, we will refer mostly to the following version.

$$\begin{aligned} d \ln L_i &= -\frac{1}{2} |\Gamma_i|^2 dt - \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1 + \delta L_j} \left(\sum_{k=1}^{m+d} \Gamma_{jk} \Gamma_{ik} \right) dt + \Gamma_i \cdot d\mathcal{W}^{(n)} \\ &= -\frac{1}{2} |\Gamma_i|^2 dt - \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1 + \delta L_j} \Gamma_j \cdot \Gamma_i dt + \Gamma_i \cdot d\mathcal{W}^{(n)}, \end{aligned} \tag{59}$$

for $1 \leq i < n$. A straightforward application of Itô's Lemma to (58) shows (59).

3.3 Comparison to Wu/Zhang - Heston Model

In principle the extended model (53) cannot be compared to the Wu/Zhang model, because already the volatility processes are given under different measures. Only if the volatility process is independent of the Brownian motions driving the underlying, a change of measure will in general not affect the dynamics of the volatility process. The non-zero correlation between the two is, however, the crucial ingredient we seek and do not want to miss. If we consider a natural *alternative* model to Wu/Zhang, which specifies the CIR process under the terminal measure, we can reasonably analyze the difference between this modified Wu/Zhang model and (53).

In this section we demonstrate that the extended model (53) is indeed a generalization of the model by Wu/Zhang, if the volatility process of the latter is considered under the terminal measure. Under a specific choice of volatility parameters β_i, we then retrieve a similar alternative Heston model. Recall that the original Wu/Zhang model under the risk-neutral measure is given by

$$\begin{aligned} \frac{dL_i}{L_i} &= v \sum_{j=1}^{i} \frac{\delta L_j}{1+\delta L_j} \eta_j \cdot \eta_i \, dt + \sqrt{v}\, \eta_i \cdot d\widetilde{W}, \\ dv &= \kappa(\theta - v)dt + \epsilon\sqrt{v}\, d\widehat{W}. \end{aligned}$$

An *alternative* Wu/Zhang model, now under the terminal measure, is of the form

$$\begin{aligned} \frac{dL_i}{L_i} &= -v \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1+\delta L_j} \eta_j \cdot \eta_i \, dt + \sqrt{v}\, \eta_i \cdot d\widetilde{W}^{(n)}, \\ dv &= \kappa(\theta - v)dt + \epsilon\sqrt{v}\, d\widehat{W}^{(n)}, \end{aligned} \tag{60}$$

where we deviate from the notation in Section 2.3 to better identify corresponding features.

Lemma 38 *The extended* LIBOR *market model (53) is a generalization of the alternative Wu/Zhang model (60).*

Proof. In the extended model choose $r = 1$, $d = m$ and $v_k \equiv v$, for all $k = 1, ..., m$. The last equivalence is necessary to match Wu/Zhang's postulate of a one-dimensional CIR process v for the stochastic volatility. If we further set

$$\beta_i = \eta_i \, ,$$

or componentwise

$$\beta_{ik} = \eta_{ik}\,,$$

for all k, we in fact obtain the alternative version (60). In choosing $r = 1$ the first m components of Γ_i will be zero. The inner product in (58) reduces to a sum starting at $m + 1$:

$$\frac{dL_i}{L_i} = -\sum_{j=i+1}^{n-1} \frac{\delta L_j}{1+\delta L_j} \left(\sum_{k=m+1}^{2m} \Gamma_{jk}\Gamma_{ik} \right) dt + \Gamma_i \cdot d\widetilde{W}^{(n)}$$

$$= -v \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1+\delta L_j} \eta_j \cdot \eta_i dt + \sqrt{v} \sum_{k=1}^{m} \eta_{ik} d\widetilde{W}_k^{(n)}$$

$$= -v \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1+\delta L_j} \eta_j \cdot \eta_i \, dt + \sqrt{v}\, \eta_i \cdot d\widetilde{W}^{(n)}.$$

■

In the next section we will discuss a covariance assumption that will render the extended model (53) consistent with the standard LIBOR model, in the sense of equal covariance structures. As a warm-up exercise for the technique, we find the process η that will ensure the same for the alternative Wu/Zhang model.

Lemma 39 *Define $Ev(t) = \lambda(t) \in \mathbb{R}$. If*

$$\eta = \gamma/\sqrt{\lambda},$$

then (60) has the same covariance structure as the standard LIBOR *model.*

Proof. Define

$$\xi_i^{WZ}(t) := \int_0^t \sqrt{v}\, \eta_i \cdot d\widetilde{W}$$

$$= \sum_{k=1}^{m} \int_0^t \sqrt{v}\, \eta_{ik} d\widetilde{W}_k.$$

We have to show that

$$E(\xi_i^{WZ} \xi_j^{WZ}) = \int_0^t \gamma_i \cdot \gamma_j \, ds.$$

But

$$\begin{aligned}
E(\xi_i^{WZ}\xi_j^{WZ}) &= E\left(\sum_{k=1}^{m}\int_0^t \sqrt{v}\,\eta_{ik}d\widetilde{W}_k \sum_{l=1}^{m}\int_0^t \sqrt{v}\,\eta_{jl}d\widetilde{W}_l\right) \\
&= \sum_{k=1}^{m}\int_0^t E(v)\eta_{ik}\eta_{jk}\,ds \\
&= \sum_{k=1}^{m}\int_0^t \lambda\frac{\gamma_{ik}}{\sqrt{\lambda}}\frac{\gamma_{jk}}{\sqrt{\lambda}}\,ds \\
&= \sum_{k=1}^{m}\int_0^t \gamma_{ik}\gamma_{jk}ds \\
&= \int_0^t \gamma_i\cdot\gamma_j\,ds.
\end{aligned}$$

Note that many cross products in the second equality can be neglected, because the components of $\widetilde{W}$ are uncorrelated. ∎

If we go a step further and consider the extended model (53) in a simplified version with just one volatility process v we obtain

$$\begin{aligned}
\frac{dL_i}{L_i} &= (...)dt + \sqrt{1-r_i^2}\gamma_i\cdot dW^{(n)} + r_i\sqrt{v}\beta_i d\widetilde{W}^{(n)} \\
dv &= \kappa(\theta - v)dt + \sqrt{v}d\widehat{W}^{(n)}.
\end{aligned} \tag{61}$$

This model can also be interpreted as an extended alternative Wu/Zhang model (60). In this case we have a similar result.

Lemma 40 *Let again $Ev(t) = \lambda(t) \in \mathbb{R}$. If $r_i \equiv r$ and*

$$\beta = \gamma/\sqrt{\lambda}$$

then (61) has the same covariance structure as the standard LIBOR *model.*

Proof. Define

$$\xi_i(t) := \sqrt{1-r_i^2}\int_0^t \gamma_i\cdot dW + r_i\int_0^t \sqrt{v}\beta_i\cdot d\widetilde{W}.$$

We have to show that

$$E(\xi_i\xi_j) = \int_0^t \gamma_i\cdot\gamma_j\,ds.$$

But

$$\begin{aligned}
E(\xi_i\xi_j) = &\sqrt{1-r_i^2}\sqrt{1-r_j^2}\int_0^t \gamma_i \cdot \gamma_j ds \\
&+ r_i r_j E\left(\int_0^t \sqrt{v}\beta_i \cdot d\widetilde{W} \cdot \int_0^t \sqrt{v}\beta_j \cdot d\widetilde{W}\right) \\
&= \sqrt{1-r_i^2}\sqrt{1-r_j^2}\int_0^t \gamma_i \cdot \gamma_j ds + r_i r_j \int_0^t E\left(v\right)\beta_i \cdot \beta_j ds \\
&= \sqrt{1-r_i^2}\sqrt{1-r_j^2}\int_0^t \gamma_i \cdot \gamma_j ds + r_i r_j \int_0^t \lambda(s)\beta_i \cdot \beta_j ds \\
&= (1-r^2)\int_0^t \gamma_i \cdot \gamma_j ds + r^2 \int_0^t \lambda(s)\beta_i \cdot \beta_j ds \\
&= \int_0^t \gamma_i \cdot \gamma_j ds + r^2\left(\int_0^t \lambda(s)\beta_i \cdot \beta_j ds - \int_0^t \gamma_i \cdot \gamma_j ds\right) \\
&= \int_0^t \gamma_i \cdot \gamma_j ds.
\end{aligned}$$

Again most cross products in the second equality can be neglected, since the components of $\widetilde{W}$ are uncorrelated. ■

We have seen two examples in which a normalizing quantity $\sqrt{\lambda}$ is introduced into a model in order to establish the same correlation structure as in the standard LIBOR model. In the next section we analyze the most general case (53) with a vector volatility process. The normalizing quantity λ will be replaced by a *normalizing* matrix.

Remark 41 *We still allow λ to be a deterministic function of time. We will see shortly that if we take the limiting case of a stationary distribution, λ can be considered constant $\lambda(t) = \theta$. If $\theta \equiv 1$, (60) and (61) will indeed resemble a perturbed standard* LIBOR *model.*

3.4 Reduction of Parameters by a Covariance Assumption

Within the particular framework constructed above, one could interpret the second diffusion part in (53), namely $r_i\beta_i \cdot dU$, as an extension or perturbation of a given LIBOR market model. In the paper of Schoenmakers/Belomestny (2006) a similar extension consisted of a jump process. They also require the local covariance structure of their jump-diffusion model to coincide with the local covariance structure of the underlying classical LIBOR market model. In employing the same assumption for the model analyzed in this work, we construct an interesting *continuous* alternative to the jump-diffusion process considered before. We thus have a comparison of two quite different approaches — jump diffusion and stochastic volatility.
Recall that the application of the covariance assumption may be justified by three reasons:

1. Cap prices do not depend on the (local) correlation structure of forward LIBORS in a LIBOR market model, but, typically, do depend only weakly on this in a more general model. Since the correlation structure contains important information about prices of ATM swaptions, we do not want to destroy this *rich* correlation structure while calibrating the extended model to the cap(let)-strike matrix.

2. The lack of smile behavior of a LIBOR model is considered a consequence of Gaussianity of the driving random sources (Wiener processes). Therefore we want to perturb this Gaussian randomness to a non-Gaussian one by incorporating a CIR volatility process, while maintaining the (local) correlation structure of the standard LIBOR market model we started with.

3. Preserving the correlation structure allows for robust calibration, since it significantly reduces the number of parameters to be calibrated while holding a realistic correlation structure.

Two concepts have to be defined before we proceed to the main result of this section. Firstly, let us integrate the diffusion part of (53) from zero to t and define the resulting random variable by

$$\xi_i(t) := \sqrt{1-r_i^2}\int_0^t \gamma_i \cdot dW + \int_0^t r_i\beta_i \cdot dU = \int_0^t \Gamma_i \cdot d\mathcal{W}.$$

Recall that $\gamma_i \in \mathbb{R}^m$ is the (given) deterministic volatility structure of the input market model, determined by some calibration procedure to ATM caps

and ATM swaptions. We assume further that the matrix $(\gamma_{i,k}(t))_{1\leq i<n,1\leq k\leq m}$ has full rank m for all t. The deterministic vector functions $\beta_i \in \mathbb{R}^m$ allow additional degrees of freedom for upcoming fittings to volatility curves. The covariance assumption, however, restricts the choice for β_i.
Without mention we always have $i,j \in \{1,...,n-1\}$ in the sequel of this section. In the most general form we can even consider the proportion factors r_i to be different for every LIBOR.

Secondly, we refer to the results of Subsection 1.5.1. The square-root diffusions in (54) have a limiting stationary distribution. The transition law of the general CIR process

$$v(t) = v(u) + \int_u^t \left(\kappa(\theta - v(s))ds + \sigma\sqrt{v(s)}dW(s)\right)$$

is known. In particular, we have the representation

$$v(t) = \frac{\sigma^2\left(1 - e^{-\kappa(t-u)}\right)}{4\kappa}\chi^2_{\alpha,c}, \qquad t > u,$$

where $\chi^2_{\alpha,c}$ is a noncentral chi-square random variable with α degrees of freedom and noncentrality c. We have

$$\alpha := \frac{4\theta\kappa}{\sigma^2}, \quad c := \frac{4\kappa e^{-\kappa(t-u)}}{\sigma^2\left(1 - e^{-\kappa(t-u)}\right)}v(u).$$

For the expected value we obtain

$$E[v(t) \mid \mathcal{F}_u] = (v(u) - \theta)e^{-\kappa(t-u)} + \theta, \quad t \geq u.$$

See Glasserman (2003) for details. Letting $t \to \infty$, we find that $v(t)$ converges in distribution to $\sigma^2/4\kappa$ times a noncentral chi-square random variable with α degrees of freedom and noncentrality parameter zero. This limit ordinary chi-square random variable has therefore a stationary distribution in the sense that if $v(0)$ is drawn from this distribution, then $v(t)$ has the same distribution for all t. We use these facts to support the following argument.

As the distributions of v_k converge to a distribution of χ^2-type, for all $k = 1,..,d$, it is natural to take the expectation of this limiting stationary distribution as the starting value for this process.
We know that, in particular,

$$\lambda_k(t) = Ev_k = v_k(0)e^{-\kappa_k t} + \theta_k(1 - e^{-\kappa_k t}), \tag{62}$$

as can be found in, for example, Feller (1977), Cox, Ingersoll and Ross (1985) or Glassermann (2003). We are interested in the special limiting case where we let

$$v_k(0) = \theta_k, \quad \text{for} \quad k = 1, ..., d.$$

Consequently we obtain a constant expectation $Ev_k \equiv \theta_k$. Hence

$$\Lambda := \mathrm{diag}(\theta_1, ..., \theta_d)$$

is a constant diagonal matrix. The following theorem is proved in a more general setting.

Theorem 42 *Let $Ev_k(t) = \lambda_k(t) \in \mathbb{R}$, for $k = 1, ..., d$.*
Set $r_i \equiv r$ and

$$\beta_i = A(t)\gamma_i$$

for some coupling matrix $A(t) \in \mathbb{R}^{d\times m}$ that satisfies

$$A(t)^\mathsf{T}\Lambda(t)A(t) = I\,,$$

where $\Lambda(t)$ denotes a diagonal matrix in $\mathbb{R}^{d\times d}$, whose elements are the expected values $\lambda_k(t)$.
Then (53) has the same covariance structure as the standard LIBOR *model.*

Proof. For the covariance function of $\xi_i(t)$ in the terminal measure we obtain

$$\begin{aligned}
E_n(\xi_i(t)\xi_j(t)) &= \sqrt{1-r_i^2}\sqrt{1-r_j^2}\int_0^t \gamma_i^\top\gamma_j ds + r_i r_j E_n\left(\int_0^t \beta_i^\top dU \cdot \int_0^t \beta_j^\top dU\right) \\
&= \left(1-r^2\right)\int_0^t \gamma_i^\top\gamma_j ds + r^2\sum_{k=1}^d E_n\left(\int_0^t \beta_{ik}\beta_{jk}\, d\langle U_k\rangle\right) \\
&= \left(1-r^2\right)\int_0^t \gamma_i^\top\gamma_j ds + r^2\sum_{k=1}^d \int_0^t \beta_{ik}\beta_{jk}\, E_n\left(v_k\right)\, ds \\
&= \left(1-r^2\right)\int_0^t \gamma_i^\top\gamma_j ds + r^2\int_0^t \beta_i^\top\Lambda(t)\beta_j\, ds \\
&= \left(1-r^2\right)\int_0^t \gamma_i^\top\gamma_j ds + r^2\int_0^t \gamma_i^\top A(t)^\top\Lambda(t)A(t)\gamma_j\, ds \\
&= \left(1-r^2\right)\int_0^t \gamma_i^\top\gamma_j ds + r^2\int_0^t \gamma_i^\top\gamma_j\, dsr \\
&= \int_0^t \gamma_i^\top\gamma_j\, ds.
\end{aligned}$$

■

Remark 43 *The reverse direction also holds, in the sense that* $A(t)^\mathsf{T}\Lambda(t)A(t) = I$ *is indeed necessary. Consequently it are exactly the orthogonal transformations,* $Q(t) = (\Lambda)^{1/2} A(t)$*, that classify the allowed* β_i*.*

We now introduce the forementioned covariance equality that is require to hold throughout the remaining analysis. At first in a more general setting.

Assumption:

$$\sqrt{1-r_i^2}\sqrt{1-r_j^2}\int_0^t \gamma_i^\top \gamma_j dt + r_i r_j \int_0^t \beta_i^\top \Lambda \beta_j \, ds = \int_0^t \gamma_i^\top \gamma_j ds. \tag{63}$$

As a first simplification towards a reasonable model with not too many degrees of freedom, let us set: $r_i \equiv r$, and $\beta_i = A\gamma_i$, for some coupling matrix $A \in \mathbb{R}^{d\times m}$ that is allowed to be a function of t just as Λ is. When possible, we drop the independent variable to lighten notation and only insist on it, if it seems necessary to remind us of the functional relation.
By the theorem the assumption reduces to

$$0 = \gamma_i^\mathsf{T}(-r^2 I + r^2 A^\mathsf{T}\Lambda(t)A)\gamma_j,$$

for any matrix A that satisfies $A(t)^\mathsf{T}\Lambda(t)A(t) = I$.
In particular we can therefore choose $d = m$ and

$$A = \Lambda^{-1/2}(t),$$

where

$$\Lambda^{-1/2}(t) := \begin{pmatrix} \frac{1}{\sqrt{\lambda_1(t)}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sqrt{\lambda_m(t)}} \end{pmatrix}.$$

By an argument justified before, we set all starting values $v_k(0) = \theta_k$ for the processes v_k. From (62) we then have simply $\lambda_k(t) \equiv \theta_k$. As a consequence, our matrix Λ is time-independent and reduces to the form

$$\Lambda^{-1/2} = \begin{pmatrix} \frac{1}{\sqrt{\theta_1}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sqrt{\theta_m}} \end{pmatrix}.$$

Under the terminal measure our model is thus given by

$$\frac{dL_i}{L_i} = -\sum_{j=i+1}^{n-1} \frac{\delta L_j}{1+\delta L_j}\Gamma_j \cdot \Gamma_i dt + \sqrt{1-r^2}\gamma_i \cdot dW + r\ \gamma_i \cdot \Lambda^{-1/2} dU$$

$$= -\sum_{j=i+1}^{n-1} \frac{\delta L_j}{1+\delta L_j}\left(\sum_{k=1}^{2m} \Gamma_{jk}\Gamma_{ik}\right) dt + \sum_{k=1}^{m} \gamma_{ik}\left(\sqrt{1-r^2}dW_k + r\sqrt{\frac{v_k}{\theta_k}}d\widetilde{W}_k\right),$$

where Γ_i have the form

$$\Gamma_i = \begin{pmatrix} \sqrt{1-r^2}\gamma_{i1} \\ \cdot \\ \cdot \\ \sqrt{1-r^2}\gamma_{im} \\ r\gamma_{i1}\sqrt{\frac{v_1}{\theta_1}} \\ \cdot \\ \cdot \\ r\gamma_{im}\sqrt{\frac{v_m}{\theta_m}} \end{pmatrix}.$$

By the parameter redundancy argument of Section 3.2.2 we can even assume, without loss of generality, $\theta_k \equiv 1$. In that case we have $\Lambda = I \in \mathbb{R}^{m\times m}$ and (53) reduces to

$$\frac{dL_i}{L_i} = -\sum_{j=i+1}^{n-1} \frac{\delta L_j}{1+\delta L_j}\Gamma_j \cdot \Gamma_i dt + \sqrt{1-r^2}\gamma_i \cdot dW + r\ \gamma_i \cdot dU \qquad (64)$$

$$= -\sum_{j=i+1}^{n-1} \frac{\delta L_j}{1+\delta L_j}\left(\sum_{k=1}^{2m} \Gamma_{jk}\Gamma_{ik}\right) dt + \sum_{k=1}^{m} \gamma_{ik}\left(\sqrt{1-r^2}dW_k + r\sqrt{v_k}d\widetilde{W}_k\right),$$

with simplified

$$\Gamma_i = \begin{pmatrix} \sqrt{1-r^2}\gamma_{i1} \\ \cdot \\ \cdot \\ \sqrt{1-r^2}\gamma_{im} \\ r\gamma_{i1}\sqrt{v_1} \\ \cdot \\ \cdot \\ r\gamma_{im}\sqrt{v_m} \end{pmatrix}.$$

The last model (64) represents the case $\beta = \gamma$.

Though the choice $\beta = \gamma$ is appealing and (64) the most natural perturbation of the LIBOR market model, we demonstrate in the next section that without further assumptions on β we run into problems. Specifically, the determination of the characteristic functions causes difficulties. Fortunately there can be found relaxed conditions under which such solutions can be determined. The covariance condition then only holds on grid points $T_j, j = 1, ..., n-1$. As a consequence, in the times between, the covariance structure is preserved only on average. But numerically this causes only minor disturbances.

3.5 Calibration Problem and Resolution: Case L_{n-1}

3.5.1 The Problem

From this section on we assume a *full-factor* input LIBOR market model, hence $m = n-1$. Together with the convention $d = m$ stipulated in the last section, we therefore consider the most general case. Other cases have to be analyzed with dimension reductions via a principle component analysis. Furthermore, we always put $\theta = 1$. In a first step we restate the dynamics of the log-process of the $n-1^{st}$ forward LIBOR rate. From (59) we obtain

$$
\begin{aligned}
d \ln L_{n-1} &= -\frac{1}{2}|\Gamma_{n-1}|^2 dt + \sqrt{1-r^2}\gamma_{n-1} \cdot dW^{(n)} + r\ \gamma_{n-1} \cdot \Lambda^{-1/2} dU^{(n)} \\
&= -\frac{1}{2}(1-r^2)|\gamma_{n-1}|^2 dt - \frac{1}{2} r^2 \sum_{k=1}^{n-1} \gamma_{n-1,k}^2\, v_k\, dt \\
&\quad + \sum_{k=1}^{n-1} \gamma_{n-1,k} \left(\sqrt{1-r^2}\, dW_k^{(n)} + r\sqrt{v_k}\, d\widetilde{W}_k^{(n)} \right).
\end{aligned}
$$

Note that L_{n-1} is a martingale under the terminal measure P_n and therefore the remaining part of the drift vanishes.
As indicated in the work of Schoenmakers (2005), Schoenmakers/Coffey (2005) and in Section 1.2.2, we can, without loss of generality, assume an upper triangular $(n-1) \times m$ matrix $(\gamma_{j,l})$, in the sense that

$$
\gamma_{n-j,l} = 0 \quad \text{for} \quad 1 \le l < n-j, \quad j = 1, ..., n-1,
$$

for the loadings of a full factor model. Recall that this can be presumed, because an orthogonal transformation of the space spanned by the factor loadings does not affect the probability distributions. We obtain

$$d \ln L_{n-1} = -\frac{1}{2}(1-r^2)\gamma^2_{n-1,n-1}dt - \frac{1}{2}r^2\gamma^2_{n-1,n-1}\, v_{n-1}\, dt \qquad (65)$$

$$+ \gamma_{n-1,n-1}\left(\sqrt{1-r^2}dW^{(n)}_{n-1} + r\sqrt{v_{n-1}}d\widetilde{W}^{(n)}_{n-1}\right).$$

This expression depends only on the last volatility process v_{n-1}. Unfortunately, it turns out that due to the time-dependent coefficients in the two terms involving v_{n-1}, it is not possible to derive the characteristic function explicitly. Define a process

$$X_{n-1}(t) := \ln L_{n-1}(t) - \ln L_{n-1}(0)$$

and note that after rearranging terms we obtain

$$X_{n-1}(t) = \int_0^t \left(-\frac{1}{2}(1-r^2)\gamma^2_{n-1,n-1}ds + \gamma_{n-1,n-1}\sqrt{1-r^2}dW^{(n)}_{n-1}\right) \qquad (66)$$

$$+ \int_0^t \left(-\frac{1}{2}r^2\gamma^2_{n-1,n-1}\, v_{n-1}\, ds + \gamma_{n-1,n-1}\, r\sqrt{v_{n-1}}d\widetilde{W}^{(n)}_{n-1}\right)$$

$$=: X^{(1)}_{n-1}(t) + X^{(2)}_{n-1}(t).$$

The first integral, namely $X^{(1)}_{n-1}$, does not pose a problem, since apart from W_{n-1} only deterministic quantities are involved.

What causes a difficulty, however, is the Fourier transform of $X^{(2)}_{n-1}$

$$\Phi(z;t) = E_n\left\{\exp\left(izX^{(2)}_{n-1}(t)\right)\right\}.$$

To see this, simplify the expressions and set $V := \gamma^2_{n-1,n-1}\, r^2\, v_{n-1}$. The SDE for $X^{(2)}_{n-1}$ is then

$$dX^{(2)}_{n-1} = -\frac{1}{2}V dt + \sqrt{V}d\widetilde{W}^{(n)}_{n-1},$$

which resembles a special case of Bates' model without jumps. Or, in other words, a one-dimensional Heston model, as illustrated in Cont/Tankov (2003, pp 479), for example. In a simplified setup, with $X = \ln L$, consider a Heston model of the form

$$dX = -\frac{1}{2}V dt + \sqrt{V}dW^L,$$

$$dV = \alpha_1(\alpha_2 - V)dt + \alpha_3\sqrt{V}dW.$$

It is essential that the volatility process V has constant coefficients α_j, for $j = 1, 2, 3$. Only then a solution for the characteristic function is known. It can be determined by the method introduced in Section 2. Feynman-Kac incorporates the expected value above and the solution of a partial differential equation. See also Cont/Tankov (2003, equation (15.14)).

Due to the time-dependence of the function $\gamma_{n-1,n-1}$, however, the corresponding coefficients are not constant in our case:

With, again simplified, $V := \gamma^2(t)r^2v = U(t,v)$ and an application of Itô's Lemma, in the form

$$dV = \left(\frac{\partial U}{\partial t} + \frac{\partial U}{\partial v} f(t,v) + \frac{1}{2}\frac{\partial^2 U}{\partial v^2} g^2(t,v)\right) dt + \frac{\partial U}{\partial v} g(t,v) dW\,,$$

where

$$\begin{aligned} dv &= \kappa(1-v)dt + \sigma\sqrt{v}dW \\ &=: f(t,v)dt + g(t,v)dW, \end{aligned}$$

we obtain

$$\begin{aligned} dV &= \left(2\,r^2\,\gamma\gamma'\,v + r^2\,\gamma^2\kappa(1-v)\right) dt + r^2\,\gamma^2\sigma\sqrt{v}\,dW \\ &= \left(r^2\gamma^2\kappa - \left(\kappa - \frac{2\gamma'}{\gamma}\right) V\right) dt + r\,\gamma\sigma\sqrt{V}dW. \end{aligned}$$

Hence through $\gamma = \gamma(t)$ the process V has time-dependent coefficients. Moreover, even differentiable γ are needed to proceed in a meaningful way.

Wu/Zhang (2006) ran into this problem before. As a way out of this dilemma, they propose to assume piecewise constant $\gamma(t)$ to find a solution. This assumption is very restrictive, as we explained in Section 1.2.2.
In this context, note also that Itô processes with time-dependent coefficients $f(t,v)$ and $g(t,v)$ are not necessarily Levy processes. The time homogeneity assumption is violated. The whole apparatus of Levy-Khinchin results and their implications are not applicable.

A way to tackle the illustrated problem is to modify the above model in an appropriate fashion, so that an application of the methods explained in earlier chapters is possible.

3.5.2 A Way Out

In order to avoid the above problems, we relax the covariance restriction (63) and require it to hold just for specific points in time, namely the fixing dates of the forwards L_i, for $i = 1, ..., n-1$:

$$\left(1-r^2\right)\int_0^{\min(T_i,T_j)} \gamma_i\cdot\gamma_j ds + r^2\int_0^{\min(T_i,T_j)} \beta_i\cdot\beta_j\, dt = \int_0^{\min(T_i,T_j)} \gamma_i\cdot\gamma_j ds. \quad (67)$$

Whereas assumption (63) requested covariance structures of standard and extended model to be identical at every t, condition (67) holds at $\min(T_i, T_j)$ for $i, j \in \{1, .., n-1\}$ only. Since instantaneous volatilities are naturally zero after expiry dates, integration up to $\min(T_i, T_j)$ suffices.

What are the implications of such a modification? — Admittedly, we cannot claim any longer that covariance structures are preserved over the *complete* time interval. It is rather only provided on an equidistant grid. In times between, the structure is preserved on average, in a sense. However, it can not deviate significantly from the original, since it is pulled back to full consistency at the end of the tenor period.
In our view it is therefore justified to say that this extended model is sufficiently tied to the covariance structure of the classical model to be of meaning in practice.

A natural choice for the constant vectors $\beta_i \in \mathbb{R}^m$ would be for those satisfying

$$\beta_i \cdot \beta_j = \frac{1}{\min(T_i, T_j)} \int_0^{\min(T_i,T_j)} \gamma_i \cdot \gamma_j\, dt. \quad (68)$$

If the symmetric matrix on the right hand side of (68) were positive definite, such β would exist. Therefore the matrix would have to be of rank m, but this is already postulated. Moreover we would have to consider the case $d \geq m$, a conclusive argument for our already assumed $d = m$.
The (generalized) Cholesky decomposition would return some constant $\overline{\gamma}_i \in \mathbb{R}^m$, $i = 1, ..., n-1$, satisfying

$$\overline{\gamma}_i \cdot \overline{\gamma}_j = \frac{1}{\min(T_i, T_j)} \int_0^{\min(T_i,T_j)} \gamma_i \cdot \gamma_j\, dt.$$

The decomposition is in general not unique, but a suitable representation in upper or lower triangular form could be found.

Unfortunately, the right hand side of (68) is not positive definite. So alternative solutions are required. However, in order to obtain closed-form expressions for characteristic functions of (log-) LIBORS later on, we need $\beta(t)$ to be at least piecewise constant. Since the ansatz $\beta(t) = A(t)\gamma(t)$ will not succeed, we have to find reasonable approximations for (63) or its reduced form

$$\int_0^t \beta_i \cdot \beta_j \, ds = \int_0^t \gamma_i \cdot \gamma_j \, ds \,, \quad 1 \le i, j < n. \tag{69}$$

One such is to set

$$\beta_i(t) = \gamma_i(T_{m(t)}) \,,$$

with $m(t) := \inf \{j : T_j \ge t\}$ for $0 \le t \le T_i$. Equation (69) will hold in a good approximation, as the integral on the right is in fact approximated by a Riemann sum. If an even simpler structure is sought, with time-independent β, we propose to choose

$$\beta_i = \sigma_i^{Black} \mathbf{e}_i \,, \tag{70}$$

where

$$\left(\sigma_i^{Black}\right)^2 := \frac{1}{T_i} \int_0^{T_i} |\gamma_i(s)|^2 \, ds \,,$$

$$\mathbf{e}_i \cdot \mathbf{e}_j := \frac{\gamma_i \cdot \gamma_j}{|\gamma_i||\gamma_j|}(0).$$

We denote these constant vectors by $\overline{\gamma}_i := \beta_i$, for $i = 1, ..., n-1$.
With these specifications the dynamics under P_n are

$$\begin{aligned} \frac{dL_i}{L_i} &= - \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1+\delta L_j} \Gamma_j \cdot \Gamma_i \, dt + \sqrt{1-r^2}\gamma_i \cdot dW + r\, \overline{\gamma}_i \cdot dU \\ &= - \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1+\delta L_j} \Gamma_j \cdot \Gamma_i \, dt + \sum_{k=1}^{m} \gamma_{ik} \sqrt{1-r^2} dW_k + \sum_{k=1}^{m} \overline{\gamma}_{ik} r \sqrt{v_k} d\widetilde{W}_k \,, \end{aligned}$$

with Γ_i given by

$$\Gamma_i = \begin{pmatrix} \sqrt{1-r^2}\gamma_{i1} \\ \cdot \\ \cdot \\ \sqrt{1-r^2}\gamma_{im} \\ r\overline{\gamma}_{i1}\sqrt{v_1} \\ \cdot \\ \cdot \\ r\overline{\gamma}_{im}\sqrt{v_m} \end{pmatrix}.$$

Note that coefficients of stochastic components in the lower half of the vector are now constant. This feature will enable us to determine the characteristic function explicitly in the next paragraph.

3.5.3 Characteristic Function of $\ln L_{n-1}$

With the specification of $\overline{\gamma}$, the problem of Subsection 3.5.1 above is resolved. Consider the dynamics of the last log-LIBOR L_{n-1} as given in (65). We obtain

$$\begin{aligned} d\ln L_{n-1} = & -\frac{1}{2}(1-r^2)\gamma^2_{n-1,n-1}dt - \frac{1}{2}r^2\overline{\gamma}^2_{n-1,n-1}v_{n-1}dt \\ & + \gamma_{n-1,n-1}\sqrt{1-r^2}dW_{n-1} + \overline{\gamma}_{n-1,n-1}r\sqrt{v_{n-1}}d\widetilde{W}_{n-1}, \end{aligned} \tag{71}$$

which comprises an expression depending on the last volatility process v_{n-1} only. In this case of constant $\overline{\gamma}$ the characteristic function of $\ln L_{n-1}(t)$ can be derived explicitly. Set $X_{n-1}(t) := \ln L_{n-1}(t) - \ln L_{n-1}(0)$ and integrate (71) to obtain

$$\begin{aligned} X_{n-1}(t) = & -\frac{1}{2}(1-r^2)\int_0^t \gamma^2_{n-1,n-1}ds + \sqrt{1-r^2}\int_0^t \gamma_{n-1,n-1}dW_{n-1} \\ & - \frac{1}{2}r^2\overline{\gamma}^2_{n-1,n-1}\int_0^t v_{n-1}ds + \overline{\gamma}_{n-1,n-1}r\int_0^t \sqrt{v_{n-1}}d\widetilde{W}_{n-1} \\ =: & \, X^{(1)}_{n-1}(t) + X^{(2)}_{n-1}(t), \end{aligned}$$

where $X^{(1)}_{n-1}$ and $X^{(2)}_{n-1}$ are independent.
The characteristic function of a normally distributed random variable $X^{(1)}_{n-1}$ is

$$\varphi^{(1)}_n(z;t) := E_n\left\{\exp\left(\mathrm{i}\, z X^{(1)}_{n-1}(t)\right)\right\} = \exp\left(\psi^{(1)}_n(z;t)\right),$$

where we define

$$\begin{aligned} \psi^{(1)}_n(z;t) := & -\frac{z^2}{2}(1-r^2)\int_0^t \gamma^2_{n-1,n-1}ds - \mathrm{i}\, z\int_0^t \frac{1}{2}(1-r^2)\gamma^2_{n-1,n-1}ds \\ = & -\frac{1}{2}\left(z^2 + \mathrm{i}\, z\right)\left(1-r^2\right)\int_0^t \gamma^2_{n-1,n-1}ds. \end{aligned}$$

The Fourier transform of $X^{(2)}_{n-1}$ is more involved. Setting

$$V = r^2\overline{\gamma}^2_{n-1,n-1}v_{n-1},$$

we obtain

$$
\begin{aligned}
dV &= d\left(r^2\overline{\gamma}^2_{n-1,n-1}v_{n-1}\right)\\
&= \kappa_{n-1}(r^2\overline{\gamma}^2_{n-1,n-1} - V)dt + \sigma_{n-1}r^2\overline{\gamma}^2_{n-1,n-1}\sqrt{v_{n-1}}\,d\widehat{W}_{n-1}\\
&= \kappa_{n-1}(r^2\overline{\gamma}^2_{n-1,n-1} - V)dt + \sigma_{n-1}r\overline{\gamma}_{n-1,n-1}\sqrt{V}\,d\widehat{W}_{n-1}.
\end{aligned}
$$

We therefore face an SDE of the following form

$$
\begin{aligned}
dX^{(2)}_{n-1} &= -\frac{1}{2}V dt + \sqrt{V}\,d\widetilde{W},\\
dV &= \kappa(\eta - V)dt + \theta\sqrt{V}\,d\widehat{W}.
\end{aligned}
$$

The characteristic function of the solution $X^{(2)}_{n-1}$ is known, see Heston (1993) or Cont/Tankov (2003):

$$
\begin{aligned}
\varphi^{(2)}_n(z;t) &= E_n\left\{\exp\left(\mathfrak{i}\,zX^{(2)}_{n-1}(t)\right)\right\}\\
&= \exp\left(\kappa\sigma^{-2}\cdot t\,(\kappa - \mathfrak{i}\rho\sigma r\overline{\gamma}\cdot z) - \frac{(z^2+\mathfrak{i}\,z)V_0}{\lambda\coth\frac{\lambda\cdot t}{2} + \kappa - \mathfrak{i}\rho\sigma r\overline{\gamma}\cdot z}\right)\\
&\quad\cdot\left(\cosh\frac{\lambda\cdot t}{2} + \frac{\kappa - \mathfrak{i}\rho\sigma r\overline{\gamma}\cdot z}{\lambda}\sinh\frac{\lambda\cdot t}{2}\right)^{-2\kappa\sigma^{-2}},
\end{aligned}
$$

where

$$
\lambda = \sqrt{\sigma^2 r^2\overline{\gamma}^2(z^2+\mathfrak{i}\,z) + (\kappa - \mathfrak{i}\rho\sigma r\overline{\gamma}\cdot z)^2}.
$$

Obviously $X^{(2)}_{n-1}(0) = 0$. Note how $\varphi^{(2)}_n(z;t)$ allows a rather compact expression compared to $\varphi^{(1)}_n(z;t)$, which requires the whole history of $\gamma_{n-1,n-1}$ for the integrals. This is again owe to a constant $\overline{\gamma}_{n-1,n-1}$ over the considered time interval.

Since $X^{(1)}_{n-1}$ and $X^{(2)}_{n-1}$ are independent, we can compute the characteristic function of X_{n-1} as the product of the two

$$\begin{aligned}\varphi_n(z;t) &= E_n \left\{\exp\left(\mathfrak{i}\, z X_{n-1}(t)\right)\right\}\\ &= E_n \left\{\exp\left(\mathfrak{i}\, z X^{(1)}_{n-1}(t) + \mathfrak{i}\, z X^{(2)}_{n-1}(t)\right)\right\}\\ &= E_n \left\{\exp\left(\mathfrak{i}\, z X^{(1)}_{n-1}(t)\right)\right\} \cdot E_n \left\{\exp\left(\mathfrak{i}\, z X^{(2)}_{n-1}(t)\right)\right\}\\ &= \varphi^{(1)}_n(z;t) \cdot \varphi^{(2)}_n(z;t).\end{aligned}$$

In the next section we demonstrate how caplet prices on $L_{n-1}(T_{n-1})$ can be expressed as an integral over $\varphi_n(z;t)$. This brings us into the position to calibrate the first three parameters κ, σ and ρ. The allotment parameter r can be chosen appropriately.

Remark 44 *For the characteristic function $\varphi^{(2)}_n(z;t)$ identify in Cont/Tankov (pp 477):*

$$\gamma = \lambda, \eta = r^2 \overline{\gamma}^2_{n-1,n-1}, \theta = \sigma_{n-1} r \overline{\gamma}_{n-1,n-1}, \rho = \rho_{n-1}, \kappa = \kappa_{n-1}, \sigma = \sigma_{n-1}.$$

Then drop the subscript $n-1$.

3.6 Pricing Caplets

In this section we derive an expression for a general caplet price in terms of the characteristic function. A caplet for the period $[T_j, T_{j+1}]$ with strike K is an option that pays $(L_j(T_j) - K)^+\delta$ at time T_{j+1}, where $1 \leq j < n$. It is well-known that under the T_{j+1}-forward measure P_n the caplet price has the following simple representation. Writing E_{j+1} for the expectation with respect to this measure, we have

$$C_j(K) = \delta B_{j+1}(0) E_{j+1}(L_j(T_j) - K)^+$$

for the price of the j-th caplet at time zero. Consequently the j-th caplet price is determined by the dynamics of L_j under P_{j+1} only. Before recalling the FFT method of Carr/Madan, let us transform the strike variable for a fixed j into a log-forward moneyness variable defined by

$$v := \ln \frac{K}{L_j(0)}.$$

This definition simplifies the following derivations. In terms of this transformation the j-th caplet price is then given by

$$\mathcal{C}_j(v) := C_j(e^v L_j(0)) = \delta B_{j+1}(0) L_j(0) E_{j+1}\left(e^{X_j(T_j)} - e^v\right)^+,$$

where $X_j(T_j) = \ln L_j(T_j) - \ln L_j(0)$. We further introduce an auxiliary function

$$\begin{aligned}
\mathcal{O}_j(v) &:= \delta^{-1} B_{j+1}^{-1}(0) L_j^{-1}(0) \mathcal{C}_j(v) - (1 - e^v)^+ \\
&= E_{j+1}\left(e^{X_j(T_j)} - e^v\right)^+ - (1 - e^v)^+ \\
&= 1_{\{v \geq 0\}} \cdot E_{j+1}\left(e^{X_j(T_j)} - e^v\right)^+ + 1_{\{v \leq 0\}} \cdot E_{j+1}\left(e^v - e^{X_j(T_j)}\right)^+,
\end{aligned}$$

where the third expression is basically due to the put-call parity and follows from the identity $(a-b)^+ = a-b+(b-a)^+$ and the fact that $E_{j+1}e^{X_j(T_j)} = 1$. For an important property of $\mathcal{O}_j(v)$ we need the following proposition.

Proposition 45 *For the Fourier transform of the function $\mathcal{O}_j$ defined above and $\varphi_{j+1}(\cdot\,; t)$ denoting the characteristic function of the process $X_j(t)$ under P_{j+1} we have*

$$\mathcal{F}\{\mathcal{O}_j\}(z) = \int_{-\infty}^{\infty} \mathcal{O}_j(v) e^{\mathrm{i}vz} dv = \frac{1 - \varphi_{j+1}(z - \mathrm{i}\,; T_j)}{z(z - \mathrm{i})}. \qquad (72)$$

Proof. Consider the integration over $[0, \infty]$ first. For this region we have

$$\begin{aligned}
\int_0^\infty \mathcal{O}_j(v) e^{\mathrm{i}\, vz} dv &= \int_0^\infty e^{\mathrm{i}\, vz} E_{j+1} \left(e^{X_j(T_j)} - e^v\right)^+ dv \\
&= \int_0^\infty e^{\mathrm{i}\, vz} \int_v^\infty \left(e^x - e^v\right) P_{j+1}(X_j(T_j) \in dx) dv \\
&= \int_0^\infty \int_0^x e^{\mathrm{i}\, vz} \left(e^x - e^v\right) dv P_{j+1}(X_j(T_j) \in dx) \\
&= \int_0^\infty \left(e^{(\mathrm{i}\, z+1)x} \left(\frac{1}{\mathrm{i}\, z} - \frac{1}{\mathrm{i}\, z+1} \right) + \frac{1}{\mathrm{i}\, z+1} - \frac{e^x}{\mathrm{i}\, z} \right) P_{j+1}(dx).
\end{aligned}$$

Here and in the sequel $P_{j+1}(dx)$ is to be interpreted as $P_{j+1}(X_j(T_j) \in dx)$. On the other hand,

$$\begin{aligned}
\int_{-\infty}^0 \mathcal{O}_j(v) e^{\mathrm{i}\, vz} dv &= \int_{-\infty}^0 e^{\mathrm{i}\, vz} E_{j+1} \left(e^v - e^{X_j(T_j)}\right)^+ dv \\
&= \int_{-\infty}^0 e^{\mathrm{i}\, vz} \int_{-\infty}^v \left(e^v - e^x\right) P_{j+1}(X_j(T_j) \in dx) dv \\
&= \int_{-\infty}^0 \int_x^0 e^{\mathrm{i}\, vz} \left(e^v - e^x\right) dv P_{j+1}(X_j(T_j) \in dx) \\
&= \int_{-\infty}^0 \left(e^{(\mathrm{i}\, z+1)x} \left(\frac{1}{\mathrm{i}\, z} - \frac{1}{\mathrm{i}\, z+1} \right) + \frac{1}{\mathrm{i}\, z+1} - \frac{e^x}{\mathrm{i}\, z} \right) P_{j+1}(dx).
\end{aligned}$$

Addition of both sides gives

$$\begin{aligned}
\int_{-\infty}^\infty \mathcal{O}_j(v) e^{\mathrm{i}\, vz} dv &= \int_{-\infty}^\infty \left(e^{(\mathrm{i}\, z+1)x} \left(\frac{1}{\mathrm{i}\, z} - \frac{1}{\mathrm{i}\, z+1} \right) + \frac{1}{\mathrm{i}\, z+1} - \frac{e^x}{\mathrm{i}\, z} \right) P_{j+1}(dx) \\
&= \left(\frac{1}{\mathrm{i}\, z} - \frac{1}{\mathrm{i}\, z+1} \right) \varphi_{j+1}(z - \mathrm{i}\,; T_j) + \frac{1}{\mathrm{i}\, z+1} - \frac{1}{\mathrm{i}\, z}.
\end{aligned}$$

The final equality holds, because

$$\int_{-\infty}^\infty P_{j+1}(X_j(T_j) \in dx) = 1,$$

and from the martingale property of $X_j(T_j)$, from which we have

$$\int_{-\infty}^{\infty} e^x P_{j+1}(X_j(T_j) \in dx) = E_{j+1} e^{X_j(T_j)} = 1.$$

From here (72) follows. ■

We are now in the position to present an expression for the caplet price. It can be computed by

$$C_j(K) = \delta B_{j+1}(0)\,(L_j(0) - K)^+ + \frac{\delta_j B_{j+1}(0) L_j(0)}{2\pi} \int_{-\infty}^{\infty} \frac{1 - \varphi_{j+1}(z - \mathfrak{i}\,; T_j)}{z(z - \mathfrak{i})} e^{-\mathfrak{i}\, z \ln \frac{K}{L_j(0)}} dz, \tag{73}$$

where φ_{j+1} is the characteristic function of $X_j(T_j)$ under P_{j+1}. Numerical quadrature is necessary to evaluate the indefinite integral approximately. Once an analytic expression for $\varphi_{j+1}(z - \mathfrak{i}\,; T_j)$ is found, all parameters involved may be calibrated. Before we approach this, however, we need to determine dynamics under relevant forward measures.

3.7 Dynamics under Various Measures

3.7.1 Dynamics under Forward Measures

In the last section we have seen, how the caplet price C_i can be determined in terms of the characteristic function φ_{i+1}. Therefore it has been necessary to work in the measure P_{i+1}. A representation under the terminal measure P_n also exists,

$$C_i(K) = \delta B_n(0) E_n \left(\frac{(L_i(T_i) - K)^+}{B_n(T_{i+1})} \right),$$

but does not allow a simple evaluation. A measure change into P_{i+1} is thus a prerequisite. Consequently the representation we are looking for is:

$$C_i(K) = \delta B_{i+1}(0) E_{i+1} \left(\frac{(L_i(T_i) - K)^+}{B_{i+1}(T_{i+1})} \right) = \delta B_{i+1}(0) E_{i+1} \left((L_i(T_i) - K)^+ \right).$$

Recall that the dynamics of all LIBOR processes are so far given in the terminal measure P_n only. See (53) or (59). Moreover, the CIR stochastic volatility processes, v_k for $k = 1, ..., n-1$, have also been specified in this measure only, see (54). To price caplets, however, we need to represent the above processes in the various forward measures P_{j+1}.

Note that the symmetric matrix $\gamma_i \cdot \gamma_j$, and thus $\overline{\gamma}_i \cdot \overline{\gamma}_j$, is decomposed into matrices of triangular structure. Therefore we have

$$\overline{\gamma}_{n-j,l} = 0 \quad \text{for} \quad 1 \le l < n - j \, , \quad j = 1, ..., n-1. \tag{74}$$

Condition (74) may in fact be found in the simpler, more common, form:

$$\overline{\gamma}_{k,l} = 0 \quad \text{for} \quad 1 \le l \le k - 1 \, , \quad k = 1, ..., n-1.$$

However, for implementation purposes we prefer to work with (74). The specific choice for $\overline{\gamma}$ allows sums to be started from i instead of 1. We then have

$$\begin{aligned} d \ln L_i = & -\frac{1}{2} \left[(1 - r^2) \left| \gamma_i \right|^2 + r^2 \sum_{k=i}^{n-1} \overline{\gamma}_{ik}^2 v_k \right] dt \\ & - \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1 + \delta L_j} \left[(1 - r^2) \gamma_i \cdot \gamma_j + r^2 \sum_{k=i}^{n-1} \overline{\gamma}_{ik} \overline{\gamma}_{jk} v_k \right] dt \\ & + \sqrt{1 - r^2} \gamma_i \cdot dW^{(n)} + r \sum_{k=i}^{n-1} \sqrt{v_k} \overline{\gamma}_{ik} d\widetilde{W}_k^{(n)}, \end{aligned} \tag{75}$$

with corresponding volatility processes

$$dv_k = \kappa_k(1 - v_k)dt + \sigma_k\sqrt{v_k}\left(\rho_k d\widetilde{W}_k^{(n)} + \sqrt{1-\rho_k^2}d\overline{W}_k^{(n)}\right). \tag{76}$$

Rearrangement of terms results in

$$\begin{aligned} d\ln L_i = &-\frac{1}{2}\left[(1-r^2)\left|\gamma_i\right|^2 + r^2\sum_{k=i}^{n-1}\overline{\gamma}_{ik}^2 v_k\right]dt \\ &+\sqrt{1-r^2}\gamma_i\cdot\left(dW^{(n)} - \sqrt{1-r^2}\sum_{j=i+1}^{n-1}\frac{\delta L_j}{1+\delta L_j}\gamma_j dt\right) \\ &+ r\sum_{k=i}^{n-1}\overline{\gamma}_{ik}\sqrt{v_k}\left(d\widetilde{W}_k^{(n)} - r\sum_{j=i+1}^{n-1}\frac{\delta L_j}{1+\delta L_j}\overline{\gamma}_{jk}\sqrt{v_k}dt\right) \\ = &(...)dt + \sqrt{1-r^2}\gamma_i\cdot dW^{(i+1)} + r\overline{\gamma}_i\cdot dU^{(i+1)}. \end{aligned}$$

In an arbitrage free setting L_i is a martingale under P_{i+1}. Therefore the evident change of measure is achieved by defining

$$\begin{aligned} dW^{(i+1)} &= dW^{(n)} - \sqrt{1-r^2}\sum_{j=i+1}^{n-1}\frac{\delta L_j}{1+\delta L_j}\gamma_j dt, \\ d\widetilde{W}_k^{(i+1)} &= d\widetilde{W}_k^{(n)} - r\sum_{j=i+1}^{n-1}\frac{\delta L_j}{1+\delta L_j}\overline{\gamma}_{jk}\sqrt{v_k}dt. \end{aligned} \tag{77}$$

We then obtain the shorter SDE

$$\begin{aligned} d\ln L_i = &-\frac{1}{2}\left[(1-r^2)\left|\gamma_i\right|^2 + r^2\sum_{k=i}^{n-1}\overline{\gamma}_{ik}^2 v_k\right]dt \\ &+\sqrt{1-r^2}\gamma_i\cdot dW^{(i+1)} \\ &+ r\sum_{k=i}^{n-1}\overline{\gamma}_{ik}\sqrt{v_k}d\widetilde{W}_k^{(i+1)}, \end{aligned} \tag{78}$$

with transformed volatility processes

$$dv_k = \kappa_k(1-v_k)dt + r\sigma_k\rho_k \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1+\delta L_j}\overline{\gamma}_{jk} v_k dt$$

$$+ \rho_k\sigma_k\sqrt{v_k}d\widetilde{W}_k^{(i+1)} + \sqrt{1-\rho_k^2}\sigma_k\sqrt{v_k}d\overline{W}_k^{(n,i+1)}.$$

$$\approx \kappa_k^{(i+1)}\left(\frac{\kappa_k}{\kappa_k^{(i+1)}} - v_k\right)dt + \sigma_k\sqrt{v_k}\left(\rho_k d\widetilde{W}_k^{(i+1)} + \sqrt{1-\rho_k^2}d\overline{W}_k^{(i+1)}\right).$$

Hereby we redefine the reversion-speed parameters of the CIR-processes for this different measure as

$$\kappa_k^{(i+1)} := \left(\kappa_k - r\sigma_k\rho_k \sum_{j=i+1}^{n-1} \frac{\delta L_j(0)}{1+\delta L_j(0)}\overline{\gamma}_{jk}\right). \tag{79}$$

The final approximation stems from freezing the LIBOR processes at $t = 0$. Furthermore, we used the fact that

$$d\overline{W}_k^{(n)} = d\overline{W}_k^{(i+1)},$$

which we indicated by $d\overline{W}_k^{(n,i+1)}$. The proof is not evident. We reserve a lemma in an own paragraph at the end of this section for it.

If we further let

$$\theta_k^{(i+1)} := \frac{\kappa_k}{\kappa_k^{(i+1)}}, \tag{80}$$

we obtain

$$dv_k \approx \kappa_k^{(i+1)}\left(\theta_k^{(i+1)} - v_k\right)dt + \sigma_k\sqrt{v_k}\left(\rho_k d\widetilde{W}_k^{(i+1)} + \sqrt{1-\rho_k^2}d\overline{W}_k^{(i+1)}\right). \tag{81}$$

Parameters are constants. So it is *not* correct to speak of a parameter having some value *under a certain measure*, P_n for example. We see, however, from (81) that the volatility processes v_k keep their structure as CIR processes after the change of measure, if redefined by (79) and (80). The *new* parameters retain their interpretation as reversion speed or reversion level. In that sense, let us remember, for example, κ_k in (76) as the reversion speed parameters *for* P_n. To avoid confusion, redefine $\kappa_k =: \kappa_k^{(n)}$.

Volatility processes v_k of this form are used in the last paragraph of the following section, where characteristic functions of forward LIBOR processes are determined under P_{j+1}.

Two facts should be noted:

1. The parameter ρ_k and σ_k are not affected by the measure change.
2. From (79) we see that $\kappa_k^{(i+1)} = \kappa_k^{(n)} = \kappa_k$ for $i \geq k$. Therefore $\theta_k^{(i+1)} = 1$ for $i \geq k$. In particular we have $\theta_i^{(i+1)} = 1$ immediately, without calibrating $\theta_i^{(i+1)}$.

3.7.2 Dynamics under Swap Measures*

In this paragraph we demonstrate that a similar parameter transformation can be obtain for a Swap market model, where the *swap measure* is considered. In the sequel we intend to work in a LIBOR market model, so the result is included for a reference purpose only.

An interest rate swap is a contract to exchange a series of floating interest payments in return for a series of fixed rate payments. Hence, consider a series of equidistant payment dates between T_{p+1} and T_q, $q > p$. The fixed leg of the swap pays δK at each time T_{j+1}, $j = p, \ldots, q-1$ where $\delta = T_{j+1} - T_j$. In return, the floating leg pays $\delta L_j(T_j)$ at time T_{j+1}, where $L_j(T_j)$ is the rate, set at time T_j for payment at T_{j+1}. Hence given the set of forward rate reset dates T_j, $j = p, \ldots, q-1$ and and the series of payment dates T_j, $j = p+1, \ldots, q$ the time t value of the interest rate swap is:

$$\sum_{j=p}^{q-1} \delta B_{j+1}(t)(L_j(t) - K).$$

The par/fair forward swap rate $S_{p,q}(t)$ is the value of the fixed rate K, such that the present value of the contract is zero, hence:

$$S_{p,q}(t) = \frac{\sum_{j=p}^{q-1} \delta B_{j+1}(t) L_j(t)}{\sum_{j=p}^{q-1} \delta B_{j+1}(t)}. \tag{82}$$

We rearrange (82) to obtain

$$S_{p,q}(t) \sum_{j=p}^{q-1} \delta B_{j+1}(t) = B_p(t) - B_q(t).$$

The right hand side is a price of a traded asset, so the left hand side must be. Its value process, expressed in terms of an appropriate numeraire, is a martingale under the associated measure. We may therefore write

$$dS_{p,q}(t) = \sigma_{p,q}(t) S_{p,q}(t) d\mathcal{W}^{(p,q)}(t), \tag{83}$$

where $dW^{(p,q)}(t)$ is a Brownian motion under probability measure $P_{p,q}$ associated with the annuity numeraire $B_{p,q} = \sum_{j=p}^{q-1} \delta B_{j+1}$. The swap rate may be expressed as a weighted sum of the constituent forward rates:

$$S_{p,q}(t) = \sum_{j=p}^{q-1} w_j(t) L_j(t),$$

with

$$w_j(t) = \frac{\delta B_{j+1}(t)}{B_{p,q}}.$$

An application of Itô's Lemma shows

$$\begin{aligned} dS_{p,q}(t) &= \sum_{j=p}^{q-1} \frac{\partial S_{p,q}(t)}{\partial L_j(t)} dL_j(t) + \sum_{j=p}^{q-1} \sum_{i=p}^{q-1} \frac{\partial^2 S_{p,q}}{\partial L_j(t) \partial L_i(t)} dL_j(t) dL_i(t) \\ &= \sum_{j=p}^{q-1} \frac{\partial S_{p,q}(t)}{\partial L_j(t)} L_j(t) \Gamma_j \cdot \left[dW^{(n)} - (\ldots) dt \right]. \end{aligned} \tag{84}$$

Equating (83) and (84), we have

$$dS_{p,q}(t) = S_{p,q}(t) \left[\sum_{j=p}^{q-1} \nu_j(t) \Gamma_j \right] \cdot d\mathcal{W}^{(p,q)}(t),$$

with $\mathcal{W}^{(p,q)} = (W^{(p,q)}, \widetilde{W}^{(p,q)})$ and

$$\nu_j(t) = \frac{\partial S_{p,q}(t)}{\partial L_j(t)} \frac{L_j(t)}{S_{p,q}(t)}.$$

The change of measure from $\mathcal{W}^{(n)}$ to $\mathcal{W}^{(p,q)}$ can be found in Schoenmakers (2005). In particular,

$$dW^{(p,q)} = dW^{(n)} - \sqrt{1 - r^2} \sum_{i=p}^{q-1} w_i(t) \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1 + \delta L_j} \gamma_j dt$$

and

$$d\widetilde{W}_k^{(p,q)} = d\widetilde{W}_k^{(n)} - r \sum_{i=p}^{q-1} w_i(t) \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1 + \delta L_j} \overline{\gamma}_{jk} \sqrt{v_k} dt.$$

In terms of these new Brownian motions the dynamics of the volatility processes are

$$dv_k = \kappa_k(1-v_k)dt + r\sigma_k\rho_k \sum_{i=p}^{q-1} w_i(t) \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1+\delta L_j}\overline{\gamma}_{jk} v_k dt$$
$$+ \rho_k\sigma_k\sqrt{v_k}d\widetilde{W}_k^{(p,q)} + \sqrt{1-\rho_k^2}\sigma_k\sqrt{v_k}d\overline{W}_k^{(p,q,n)}. \tag{85}$$

The process $\overline{W}^{(p,q,n)}$ in (85) is a standard Brownian motion under both measures $P_{p,q}$ and P_n. The proof is similar to the one given in the next paragraph for $\overline{W}^{(n,i+1)}$. Assuming that

$$\frac{\partial S_{p,q}(t)}{\partial L_j(t)} \quad \text{and} \quad \frac{L_j(t)}{S_{p,q}(t)}$$

vary little with time, freeze the weights at their time $t=0$ values. The swap rate dynamics are then approximately given by

$$dS_{p,q}(t) \approx S_{p,q}(t)\left[\sum_{j=p}^{q-1} \nu_j(0)\Gamma_j\right] \cdot d\mathcal{W}^{(p,q)}(t). \tag{86}$$

Similarly, freezing LIBORS in the drift of (85) leads to an approximated volatility process v_k given by

$$dv_k \approx \kappa_k^{(p,q)}\left(\theta_k^{(p,q)} - v_k\right)dt + \sigma_k\sqrt{v_k}\left(\rho_k d\widetilde{W}_k^{(p,q)} + \sqrt{1-\rho_k^2}d\overline{W}_k^{(p,q,n)}\right), \tag{87}$$

with reversion speed parameter

$$\kappa_k^{(p,q)} := \left(\kappa_k - r\sigma_k\rho_k \sum_{i=p}^{q-1} w_i(0) \sum_{j=i+1}^{n-1} \frac{\delta_j L_j(0)}{1+\delta_j L_j(0)}\overline{\gamma}_{jk}\right),$$

and mean reversion level

$$\theta_k^{(p,q)} := \frac{\kappa_k}{\kappa_k^{(p,q)}}.$$

3.7.3 A Lemma

Let us recapitulate and prove a result that we used in showing (81).

Lemma 46 *For $k = 1, ..., n-1$, we have*

$$d\overline{W}_k^{(n)} = d\overline{W}_k^{(i+1)}.$$

In other words, $d\overline{W}_k^{(n,i+1)}$ is invariant under the various measures, namely P_{i+1} and P_n.

Proof. See Jamshidian (1997) for the difference between compensators. It is given by

$$\mu^{i+1}_{\overline{W}_k^{(n)}} = \langle \overline{W}_k^{(n)}, \ln M \rangle,$$

with

$$M = \prod_{j=i+1}^{n-1} (1 + \delta L_j).$$

That is, we have

$$\langle \overline{W}_k^{(n)}, \ln M \rangle = d\overline{W}_k^{(n)} d \ln M = d\overline{W}_k^{(n)} d \left(\sum_{j=i+1}^{n-1} \ln (1 + \delta L_j) \right)$$

$$= \sum_{j=i+1}^{n-1} d\overline{W}_k^{(n)} d \ln(1 + \delta L_j)$$

$$= \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1 + \delta L_j} d\overline{W}_k^{(n)} d \ln L_j.$$

A closer look at (75) reveils that all terms are negligible, since of higher order than dt, or zero due to independence of $\overline{W}$ and W or $\widetilde{W}$, respectively. We thus have

$$\langle \overline{W}_k^{(n)}, \ln M \rangle = 0$$

or in other words, as indicated by $d\overline{W}_k^{(n,i+1)}$:

$$d\overline{W}_k^{(n)} = d\overline{W}_k^{(i+1)}.$$

■

Remark 47 *By exchanging* $\overline{W}_k$ *with* $\widetilde{W}_k$ *we obtain by an analog argument that*

$$\begin{aligned}\langle \widetilde{W}_k^{(n)}, \ln M\rangle &= d\widetilde{W}_k^{(n)} d\ln M \\ &= \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1+\delta L_j} d\widetilde{W}_k^{(n)} d\ln L_j \\ &= \sum_{j=i+1}^{n-1} \frac{r\delta L_j}{1+\delta L_j} \beta_{jk}\sqrt{v_k} dt\,,\end{aligned}$$

the compensator of $\widetilde{W}_k^{(n)}$ *under the measure* P_{i+1}*, which we already subtracted in (77).*

3.8 The Conditional Characteristic Function

3.8.1 General Aspects

In this section we determine the characteristic functions needed for the parameter estimation in (73). We will do this for various measures. In the first paragraph we start with the characteristic function of $\log(L_{n-1})$ under P_n, which leaves the proof, an application of Feynman-Kac, not that complicated. In the second paragraph we demonstrate, how the other characteristic functions can be found recursively. This recursion can in principle be performed while staying in the terminal measure P_n. However, it will be more convenient to calibrate the model in the corresponding measures P_{i+1} before changing back to P_n. This is illustrated in the third paragraph.

For a more convenient reference at this point, we state again our Heston LIBOR model under the terminal measure P_n. Its log-version is by Itô's Lemma:

$$d\ln L_i = -\frac{1}{2}\left|\Gamma_i\right|^2 dt - \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1+\delta L_j}\Gamma_i \cdot \Gamma_j dt + \Gamma_i \cdot d\mathcal{W}, \tag{88}$$

for $i = 1, ..., n-1$, where

$$\Gamma_i = \begin{pmatrix} \sqrt{1-r^2}\gamma_{i1} \\ . \\ . \\ \sqrt{1-r^2}\gamma_{i,n-1} \\ r\overline{\gamma}_{i1}\sqrt{v_1} \\ . \\ . \\ r\overline{\gamma}_{i,n-1}\sqrt{v_{n-1}} \end{pmatrix}, \quad \mathcal{W} = \begin{pmatrix} W_1 \\ . \\ . \\ W_{n-1} \\ \widetilde{W}_1 \\ . \\ . \\ \widetilde{W}_{n-1} \end{pmatrix}.$$

From $d\widetilde{W}_k \cdot d\widehat{W}_k = \rho_k dt$ the dynamics of the volatility processes v_k, can be written as

$$dv_k = \kappa_k(1-v_k)dt + \sigma_k\rho_k\sqrt{v_k}d\widetilde{W}_k + \sigma_k\sqrt{(1-\rho_k^2)}\sqrt{v_k}d\overline{W}_k\,,$$

where $W_k\,, \widetilde{W}_k\,, \overline{W}_k$ are independent Brownian motions. For the parameters we have: $\kappa_k > 0$, $\sigma_k > 0$ and $\rho_k \geq 0$. Naturally, everywhere $k = 1, ..., n-1$. The vectors γ_i and $\overline{\gamma}_i$ determine the covariance structure of the forward LIBOR

rates.
Let us write (88) in the more explicit form

$$
\begin{aligned}
d \ln L_i &= -\frac{1}{2}\left[(1-r^2)\,|\gamma_i|^2 + r^2 \sum_{k=1}^{n-1} \overline{\gamma}_{ik}^2 v_k\right] dt \\
&\quad - \sum_{j=i+1}^{n-1} \frac{\delta L_j}{1+\delta L_j}\left[(1-r^2)\sum_{k=1}^{n-1} \gamma_{ik}\gamma_{jk} + r^2 \sum_{k=1}^{n-1} \overline{\gamma}_{ik}\overline{\gamma}_{jk} v_k\right] dt \\
&\quad + \sqrt{(1-r^2)} \sum_{k=1}^{n-1} \gamma_{ik} dW_k + r \sum_{k=1}^{n-1} \overline{\gamma}_{ik} \sqrt{v_k}\, d\widetilde{W}_k. \qquad (89)
\end{aligned}
$$

The following arguments are easier comprehensible in this form.

We need to determine the conditional characteristic function of $\ln L_i(T)$ given $L_i(0)$ for all $i = 1, ..., n-1$.

3.8.2 The Last Libor

In the main body of this paragraph two theorems are proved. A result of Feynman-Kac is needed for the first. It will be introduced here in the version found in Duffie (1996). For another reference, see Kloeden/Platen (1992).

Theorem 48 *Feynman-Kac*
The Cauchy problem consists in finding a function $f \in \mathcal{C}^{2,1}(\mathbb{R}^N \times [0,T))$*, for given* $T \geq 0$*, solving*

$$\mathcal{D}f(x,t) - r(x,t)f(x,t) + h(x,t) = 0\,, \quad (x,t) \in \mathbb{R}^N \times [0,T),$$

with boundary condition

$$f(x,T) = g(x)\,, \quad x \in \mathbb{R}^N\,,$$

where

$$\mathcal{D}f(x,t) = f_t(x,t) + f_x(x,t)\mu(x,t) + \frac{1}{2} tr[\sigma(x,t) \cdot \sigma(x,t)^\mathsf{T} f_{xx}(x,t)]\,.$$

For the functions involved we have

$$r : \mathbb{R}^N \times [0,T) \to \mathbb{R}\,, \quad h : \mathbb{R}^N \times [0,T) \to \mathbb{R}\,, \quad g : \mathbb{R}^N \to \mathbb{R}\,,$$

and

$$\mu : \mathbb{R}^N \times [0,T) \to \mathbb{R}^N \quad \sigma : \mathbb{R}^N \times [0,T) \to \mathbb{R}^{N \times d}.$$

A probabilistic solution to this problem is given by

$$f(x,t) = E^{x,t}\left[\int_t^T \psi_{t,s} h(X_s,s)ds + \psi_{t,T}\, g(X_T)\right],$$

where

$$\psi_{t,s} = \exp\left[-\int_t^s r(X_\tau,\tau)d\tau\right].$$

$E^{x,t}$ *indicates that* X *is assumed to solve the SDE*

$$dX_t = \mu(X_t,t)dt + \sigma(X_t,t)dB_t\,.$$

Theorem 49 *With* $v \in \mathbb{R}^{n-1}$*, the characteristic function of* $\ln L_{n-1}(T) - \ln L_{n-1}(0)$ *is given by*

$$\begin{aligned}\varphi_n(z\,;T,v) &= E_n\left[\exp\left(\mathrm{i}\, z\ln\left(\frac{L_{n-1}(T)}{L_{n-1}(0)}\right)\right)\middle|\, v_k(0) = v_k, k = 1,...,n-1\right] \\ &= \varphi_{n,0}\,(z\,;T)\prod_{k=1}^{n-1}\varphi_{n,k}\,(z\,;T,v_k)\,, \end{aligned} \tag{90}$$

where

$$\varphi_{n,0}(z\,;T) = \exp\left(-\frac{1}{2}(1-r^2)\eta^2_{n-1}(T)\left(z^2+\mathrm{i}\,z\right)\right), \quad \eta^2_{n-1}(T) = \int_0^T |\gamma_{n-1}|^2\,dt$$

and each $\varphi_{n,k}\,(z\,;T,v_k)$ *is of the form* $\varphi_{n,k}\,(z\,;T,v_k) = p_{n,k}\,(z\,;T,y_k,v_k)\,|_{y_k=0}$*, with* $p_{n,k}$ *satisfying the parabolic equations*

$$\begin{aligned}\frac{\partial p_{n,k}}{\partial T} &= \kappa_k\,(1-v_k)\,\frac{\partial p_{n,k}}{\partial v_k} - \frac{1}{2}r^2\overline{\gamma}^2_{n-1,k}v_k\frac{\partial p_{n,k}}{\partial y_k} + \frac{1}{2}\sigma_k^2 v_k\frac{\partial^2 p_{n,k}}{\partial v_k^2} \\ &\quad + \frac{1}{2}r^2\overline{\gamma}^2_{n-1,k}v_k\frac{\partial^2 p_{n,k}}{\partial y_k^2} + \sigma_k\rho_k r\overline{\gamma}_{n-1,k}v_k\frac{\partial^2 p_{n,k}}{\partial v_k\partial y_k},\end{aligned} \tag{91}$$

with boundary condition

$$p_{n,k}(z\,;0,y_k,v_k) = e^{\mathrm{i}\,zy_k}.$$

Proof. In (89) set $i = n-1$. Representation (90) is immediate after integrating (89) and noting that the middle sum involving the other LIBORS is zero, since L_{n-1} is a martingale under P_n.

Define

$$p_{n,k}\left(z\,;T,y_k,v_k\right):=$$

$$=E_n^{x,0}\left[\exp\left(\mathrm{i}\,z\left(Y_k(0)-\frac{1}{2}r^2\overline{\gamma}_{n-1,k}^2\int_0^T v_k ds+r\overline{\gamma}_{n-1,k}\int_0^T\sqrt{v_k}d\widetilde{W}_k\right)\right)\right]$$

$$=E_n\left[\exp\left(\mathrm{i}z\left(Y_k(T)\right)\right)|Y_k(0)=y_k,v_k(0)=v_k\right],$$

with a random variable $Y_k(T)$ defined in the obvious way. In the first expected value we introduce $x:=(y_k,v_k)$, for $t=0$, in analogy with the Feynman-Kac Theorem 48 .

Additionally set $r\equiv 0$ and $h\equiv 0$ in Theorem 48. Then we have $\psi_{t,s}\equiv 1$, for all $s\geq t$. Further define a stochastic vector

$$X_t=\left(Y_k(t),v_k(t)\right),\quad \text{with initial value}\quad X_0=x=(y_k,v_k)$$

and

$$\begin{pmatrix} dY_k \\ dv_k \end{pmatrix}:=\begin{pmatrix} -\frac{1}{2}r^2\overline{\gamma}_{n-1,k}^2 v_k dt+r\overline{\gamma}_{n-1,k}\sqrt{v_k}\,d\widetilde{W}_k \\ \kappa_k(1-v_k)dt+\sigma_k\rho_k\sqrt{v_k}\,d\widetilde{W}_k+\sigma_k\sqrt{(1-\rho^2)}\sqrt{v_k}\,d\overline{W}_k \end{pmatrix}.$$

Set

$$g(X_T)=g\begin{pmatrix} Y_k(T) \\ v_k(T) \end{pmatrix}=e^{\mathrm{i}\,zY_k(T)}.$$

With $f(x,t)=p_{n,k}(z\,;T-t,y_k,v_k)$ we observe that we have indeed

$$f(x,T)=p_{n,k}(z\,;0,y_k,v_k)=e^{\mathrm{i}\,zy_k}=g(X_0)=g(x).$$

Further identify

$$\begin{pmatrix} \mu_1(X_t,t) \\ \mu_2(X_t,t) \end{pmatrix}=\begin{pmatrix} -\frac{1}{2}r^2\overline{\gamma}_{n-1,k}^2 v_k(t) \\ \kappa_k(1-v_k(t)) \end{pmatrix}$$

and

$$\sigma(X_t,t)=\begin{pmatrix} r\overline{\gamma}_{n-1,k}\sqrt{v_k(t)} & 0 \\ \sigma_k\rho_k\sqrt{v_k(t)} & \sigma_k\sqrt{(1-\rho_k^2)}\sqrt{v_k(t)} \end{pmatrix}.$$

Then we obtain

$$\sigma(X_t,t)\cdot\sigma(X_t,t)^{\mathsf{T}} = \begin{pmatrix} r^2\overline{\gamma}^2_{n-1,k}v_k(t) & \sigma_k\rho_k r\overline{\gamma}_{n-1,k}v_k(t) \\ \sigma_k\rho_k r\overline{\gamma}_{n-1,k}v_k(t) & \sigma^2 v_k(t) \end{pmatrix}.$$

Compute now

$$\frac{1}{2}tr\left(\sigma(X_t,t)\cdot\sigma(X_t,t)^{\mathsf{T}}\ \frac{\partial p_{n,k}}{\partial X}\right),$$

where

$$\frac{\partial p_{n,k}}{\partial X} = \begin{pmatrix} \dfrac{\partial^2 p_{n,k}}{\partial y_k^2} & \dfrac{\partial^2 p_{n,k}}{\partial y_k \partial v_k} \\ \\ \dfrac{\partial^2 p_{n,k}}{\partial y_k \partial v_k} & \dfrac{\partial^2 p_{n,k}}{\partial v_k^2} \end{pmatrix}.$$

Conclude the result with Theorem 48. The boundary condition for $\varphi_{n,k}$ translates through $p_{n,k}$ by

$$\varphi_{n,k}\left(z\,;0,v_k\right) = p_{n,k}\left(z\,;0,y_k,v_k\right)|_{y_k=0} = 1.$$

■

Since $\overline{\gamma}_i$ are constant, the above equation can be solved explicitly.

Theorem 50 *Solutions for* $\varphi_{n,k}$ *are given by*

$$\varphi_{n,k}(z\,;T,v_k) = \exp\left(A_{n-1,k}(z;T) + v_k B_{n-1,k}(z;T)\right),$$

with

$$\begin{aligned} A_{n-1,k}(z;T) &= \frac{\kappa_k}{\sigma_k^2}\left[(a_{n-1,k}+d_{n-1,k})T - 2\ln\left(\frac{1-g_{n-1,k}e^{d_{n-1,k}T}}{1-g_{n-1,k}}\right)\right] \\ B_{n-1,k}(z;T) &= \frac{(a_{n-1,k}+d_{n-1,k})(1-e^{d_{n-1,k}T})}{\sigma_k^2(1-g_{n-1,k}e^{d_{n-1,k}T})}, \end{aligned}$$

and

$$\begin{aligned} a_{n-1,k} &= \kappa_k - \mathfrak{i} r\rho_k\sigma_k\overline{\gamma}_{n-1,k}\,z \\ d_{n-1,k} &= \sqrt{a_{n-1,k}^2 + r^2\overline{\gamma}^2_{n-1,k}\sigma_k^2(z^2+\mathfrak{i}\,z)} \\ g_{n-1,k} &= \frac{a_{n-1,k}+d_{n-1,k}}{a_{n-1,k}-d_{n-1,k}}. \end{aligned}$$

Proof. It is well known that equation (91) can be solved explicitly by the ansatz

$$p_{n,k}(z\,;T,y_k,v_k) = \exp\left(A_{n-1,k}(z;T) + v_k B_{n-1,k}(z;T) + \mathfrak{i}\, z y_k\right),$$

which gives a Riccati equation in $A_{n-1,k}$ and $B_{n-1,k}$ with the solutions given above. We thus obtain

$$\varphi_{n,k}(z\,;T,v_k) = p_{n,k}\left(z\,;T,y_k,v_k\right)|_{y_k=0} = \exp\left(A_{n-1,k}(z;T) + v_k B_{n-1,k}(z;T)\right).$$

■

Note that the first lower index n of characteristic functions refers to a measure, whereas the first lower index $n-1$ introduced at coefficients refers to a forward LIBOR. The second lower index refers to the component.

Remark 51 *Recall that as many CIR-processes perturb the stochastic volatility component as we have forward* LIBOR *rates,* $d = n-1$*. Note that in Theorem 49 we showed more than necessary. Due to the appropriate choice of* $\overline{\gamma}_i$ *, only the last* log-LIBOR *contributes a nontrivial factor to the characteristic function. For all others we actually have*

$$\varphi_{n,k} \equiv 1\,, \quad k = 1, ..., n-2\,.$$

Consequently for the characteristic function of $\ln L_{n-1}$ *only the parameters* κ_{n-1} , σ_{n-1} *and* ρ_{n-1} *need to be determined. Later we will take advantage of this fact in our recursion algorithm.*

3.8.3 Recursion Back in Time

Now that we determined the characteristic function φ_n^{n-1} of the (normalized) $(n-1)$ - LIBOR, we could in principle work our way back to the front recursively, while staying in the same measure P_n. For the m-th LIBOR, with $m < n-1$, we can find $\varphi_n^m(z;T)$ by the following routine. To begin with

$$
\begin{aligned}
\varphi_n^m(z\,;T,v) &= E_n\left[\exp\left(\mathrm{i}\,z\ln\left(\frac{L_m(T)}{L_m(0)}\right)\right)|v_k(0)=v_k\right] \\
&= \varphi_{n,0}^m\,(z\,;T)\exp\left(\sum_{k=1}^{n-1}[A_{m,k}(z;T)+v_kB_{m,k}(z;T)]\right),
\end{aligned}
$$

where the ansatz for $\varphi_{n,k}^m\,(z\,;T)$ is chosen according to the latest results. For $\varphi_{n,0}^m\,(z\,;T)$ in this case we have

$$
\varphi_{n,0}^m(z\,;T) =
$$

$$
\exp\left(-(1-r^2)\left[\frac{\theta_m^2(T)\,(z^2+\mathrm{i}\,z)}{2}+\mathrm{i}\,z\sum_{j=m+1}^{n-1}\frac{\delta_jL_j(0)}{1+\delta_jL_j(0)}\sum_{k=1}^{n-1}\chi_{mjk}(T)\right]\right),
$$

where

$$
\theta_m^2(T)=\int_0^T|\gamma_m|^2\,dt,\quad \chi_{mjk}(T)=\int_0^T\gamma_{mk}(t)\gamma_{jk}(t)dt.
$$

Since the middle term in (89) can not be ignored here, an approximation by freezing the stochastic LIBOR rates at the initial time zero is needed.

As to $A_{m,k}$ and $B_{m,k}$, they are given by

$$
\begin{aligned}
A_{m,k}(z;T) &= \frac{\kappa_k}{\sigma_k^2}\left[(a_{m,k}+d_{m,k})T-2\ln\left(\frac{1-g_{m,k}e^{d_{m,k}T}}{1-g_{m,k}}\right)\right] \\
B_{m,k}(z;T) &= \frac{(a_{m,k}+d_{m,k})(1-e^{d_{m,k}T})}{\sigma_k^2(1-g_{m,k}e^{d_{m,k}T})},
\end{aligned}
$$

where

$$
\begin{aligned}
a_{m,k} &= \kappa_k-\mathrm{i}\,r\rho_k\sigma_k\overline{\gamma}_{m,k}z, \\
d_{m,k} &= \sqrt{a_{m,k}^2+r^2\left(\overline{\gamma}_{m,k}^2+\overline{\gamma}_{m,k}\sum_{j=m+1}^{n-1}\frac{\delta_jL_j(0)}{1+\delta_jL_j(0)}\overline{\gamma}_{jk}\right)\sigma_k^2(z^2+\mathrm{i}\,z)}, \\
g_{m,k} &= \frac{a_{m,k}+d_{m,k}}{a_{m,k}-d_{m,k}}.
\end{aligned}
$$

By means of the special pattern in (74), it is possible to determine only one parameter triple $(\kappa_m\,, \sigma_m\,, \rho_m\,)$ at a time, while going recursively backwards in time. The formerly evaluated parameters are used as fixed scalars in the next step. This procedure of optimizing over only three parameters at one step will deliver a remarkably robust calibration algorithm.

We freezed LIBOR rates at $t = 0$ in the drift terms. In the sequel we therefore just obtained an approximation. The term

$$\sum_{j=m+1}^{n-1} \frac{\delta_j L_j(0)}{1 + \delta_j L_j(0)} \overline{\gamma}_{jk}\,,$$

however, will come across again in a more refined analysis. It seems that a freezing can not be avoided.

Remark 52 *This paragraph contains a nice result, but is of no significance for the calibration process. It demonstrates that principally we can find, or better approximate, characteristic functions under the measure P_n. Unfortunately, the caplet prices (73) are given in terms of φ_{j+1}, and are as such only available under P_{j+1}.*

3.8.4 The Real Thing

In order to find the conditional characteristic function φ_{j+1}, we will basically follow the same steps as above. This time, however, the dynamics are considered under the measure P_{j+1}.

Theorem 53 *With $v \in \mathbb{R}^{n-1}$, under the measure P_{j+1} the characteristic function of $\ln L_j(T) - \ln L_j(0)$ is given by*

$$\begin{aligned} \varphi_{j+1}(z\,; T, v) &= E_{j+1}\left[\exp\left(\mathrm{i} z \ln\left(\frac{L_j(T)}{L_j(0)}\right)\right)\Big|\, v_k(0) = v_k\,,\, k = 1, ..., n-1\right] \\ &= \varphi_{j+1,0}\,(z\,; T) \prod_{k=j}^{n-1} \varphi_{j+1,k}\,(z\,; T, v_k)\,, \end{aligned} \quad (92)$$

where

$$\varphi_{j+1,0}(z\,; T) = \exp\left(-\frac{1}{2}(1-r^2)\eta_j^2(T)\left(z^2 + \mathrm{i}\,z\right)\right), \quad \eta_j^2(T) = \int_0^T |\gamma_j|^2\,dt$$

and each $\varphi_{j+1,k}\,(z\,; T, v_k)$ is of the form
$\varphi_{j+1,k}\,(z\,; T, v_k) = p_{j+1,k}\,(z\,; T, y_k, v_k)\,|_{y_k=0}$, with $p_{j+1,k}$ satisfying the parabolic

equations

$$
\begin{aligned}
\frac{\partial p_{j+1,k}}{\partial T} &= \kappa_k^{(j+1)}\left(\theta_k^{(j+1)}-v_k\right)\frac{\partial p_{j+1,k}}{\partial v_k}-\frac{1}{2}r^2\overline{\gamma}_{j,k}^2 v_k\frac{\partial p_{j+1,k}}{\partial y_k}+\frac{1}{2}\sigma_k^2 v_k\frac{\partial^2 p_{j+1,k}}{\partial v_k^2}\\
&\quad+\frac{1}{2}r^2\overline{\gamma}_{j,k}^2 v_k\frac{\partial^2 p_{j+1,k}}{\partial y_k^2}+\sigma_k\rho_k r\overline{\gamma}_{j,k}v_k\frac{\partial^2 p_{j+1,k}}{\partial v_k\partial y_k}, \qquad (93)
\end{aligned}
$$

with boundary condition

$$p_{j+1,k}(z\,;0,y_k,v_k)=e^{\mathrm{i}\,zy_k}.$$

Proof. The dynamics of $\ln L_j$ under P_{j+1} are given by (78) and (81). The arguments are now analog to those in Theorem 49. ■

Since $\overline{\gamma}_i$ are constant, the above equation can be solved explicitly.

Theorem 54 *Solutions for $\varphi_{j+1,k}$ are given by*

$$\varphi_{j+1,k}(z\,;T,v_k)=\exp\left(A_{j,k}(z;T)+v_kB_{j,k}(z;T)\right)$$

with

$$
\begin{aligned}
A_{j,k}(z;T) &= \frac{\kappa_k^{(j+1)}\theta_k^{(j+1)}}{\sigma_k^2}\left[(a_{j,k}+d_{j,k})T-2\ln\left(\frac{1-g_{j,k}e^{d_{j,k}T}}{1-g_{j,k}}\right)\right]\\
B_{j,k}(z;T) &= \frac{(a_{j,k}+d_{j,k})(1-e^{d_{j,k}T})}{\sigma_k^2(1-g_{j,k}e^{d_{j,k}T})},
\end{aligned}
$$

and

$$
\begin{aligned}
a_{j,k} &= \kappa_k^{(j+1)}-\mathrm{i}\,r\rho_k\sigma_k\overline{\gamma}_{j,k}\,z\\
d_{j,k} &= \sqrt{a_{j,k}^2+r^2\overline{\gamma}_{j,k}^2\sigma_k^2(z^2+\mathrm{i}\,z)}\\
g_{j,k} &= \frac{a_{j,k}+d_{j,k}}{a_{j,k}-d_{j,k}}.
\end{aligned}
$$

Proof. Analog to Theorem 50. Recall that the dynamics of v_k are now given by (81). We again obtain

$$\varphi_{j+1,k}(z\,;T,v_k)=p_{j+1,k}\left(z\,;T,y_k,v_k\right)|_{y_k=0}=\exp\left(A_{j,k}(z;T)+v_kB_{j,k}(z;T)\right).$$

■

Remark 55 *It is again the particular choice of $\overline{\gamma}$ that enables the product in (92) to be started at j. This fundamental feature will prove beneficial in the following calibration.*

3.9 Calibration

The calibration procedures regarding the two product groups will be explained in this section. The calibration to caplet prices will complete the parameter set for the LIBOR market model. The respectice method for swaptions is given for the sake of completeness.

3.9.1 Calibration to Caplet Prices

With the preparatory work of the last three sections, we can now outline a calibration procedure for the LIBOR structure (53). We do this under assumptions already introduced and discussed .

(i) The input market LIBOR volatility structure $\gamma \in \mathbb{R}^{(n-1)\times m}$ is of full rank, that is $m = n - 1$.

(ii) The terminal log-LIBOR increment $d\ln L_{n-1}$ is influenced by a single stochastic volatility shock dU_{n-1}. The one but last, that is $d\ln L_{n-2}$, only by dU_{n-1} and dU_{n-2}, and so forth.
Put differently, $\beta \in \mathbb{R}^{(n-1)\times d}$ is a square upper triangular matrix of rank $n-1$, hence $d = n - 1$.

(iii) The r_i are constant, that is $r_i \equiv r$, and the matrix β is determined as the time-independent upper triangular solution $\overline{\gamma}$ of the covariance condition (67), as given in (70).

(iv) Recall that $v_k(0) \equiv \theta_k \equiv 1$, $1 \le k < n$.

For the LIBOR dynamics structured in the above way we thus have

$$
\begin{aligned}
d\ln L_i = {} & -\frac{1}{2}\left((1-r^2)\,|\gamma_i|^2 + r^2 \sum_{k=i}^{n-1} \overline{\gamma}_{ik}^2\, v_k \right) dt \\
& + \sqrt{1-r^2}\,\gamma_i \cdot dW^{(i+1)} \\
& + r \sum_{k=i}^{n-1} \overline{\gamma}_{ik} \sqrt{v_k}\, d\widetilde{W}_k^{(i+1)}, \quad 1 \le i < n,
\end{aligned}
$$

where for $i = n-1$ the dynamics of v_{n-1} is given by (76), and for $i < n-1$ the dynamics of v_k, $i \le k < n$, is approximated by (81).

We will calibrate the structure to prices of caplets according to the following roadmap.

1. First step $i = n-1$. Calibrate r and the parameter set $(\kappa_{n-1}\,, \theta_{n-1} = 1\,, \sigma_{n-1}\,,\, \rho_{n-1}\,)$ to the T_{n-1} column of the cap-strike matrix via (73) using the explicitly known characteristic function φ_n of

$$\ln\left(\frac{L_{n-1}(T_{n-1})}{L_{n-1}(0)}\right),$$

as given in Theorem 50.

2. For $i = n-2$ down to 1 carry out the next iteration step:

3. The k-th step: $i = n-k$. Transform the yet known parameter set $(\kappa_j,\, \sigma_j\,,\, \rho_j)$ $i < j < n$, with (79) and (80) into the corresponding set

$$(\kappa_j^{(i+1)},\, \sigma_j^{(i+1)},\, \rho_j^{(i+1)}, \theta_j^{(i+1)})\,, \quad i < j < n.$$

By the upper triangular structure of square matrix $\overline{\gamma}$, we obviously have
$\kappa_i^{(i+1)} = \kappa_i$. Hence by (80), $\theta_i^{(i+1)} = 1$. Then calibrate the at this stage unknown parameter set $(\kappa_i,\, \sigma_i,\, \rho_i\,)$ to the T_i column of the cap-strike matrix via (73) using the explicitly known characteristic function φ_{i+1} of $\ln[L_i(T_i)/L_i(0)]$, as given in Theorem 54.

Remark 56 *Except for the first step, where the parameter r may be optimized too, only three parameters of a one-dimensional CIR process are calibrated at a time.*

In a monograph Mikhailov/Nögel (2003) point out that the calibration of a (one-dimensional) Heston model turns out to be “very robust and reliable”, even with an Excel solver, which is based on the Generalized Reduced Gradient (GRG) method. By the proposed algorithm above, the calibration of the multi-dimensional Heston volatility part is reduced to a one-dimensional problem. This suggests that the all over procedure is stable too.

3.9.2 Calibration to Swaption Prices*

A European swaption over a period $[T_p, T_q]$ gives the right to enter at T_p into an interest rate swap with strike rate K. The swaption value at time $t \leq T_p$ is given by

$$Swpn_{p,q}(t) = B_{p,q}(t) E_{p,q}^{\mathcal{F}_t}(S_{p,q}(T_p) - K)^+.$$

Since our approximative model (86)-(87) for $S_{p,q}$ has an affine structure with constant coefficients we can find the characteristic function of $S_{p,q}$ under $P_{p,q}$ and follow the lines of the previous section to calibrate the model.

Remark 57 *Due to the covariance restriction* (70)*, we can expect model prices of ATM swaptions to deviate not too far from market prices. Recall that our model employs the same covariance structure as the LMM calibrated to the market prices of ATM swaptions.*

3.10 Calibration to Real Data

In this section we calibrate the model (137) and (81) to market data available on 19.06.2008. The caplet-strike volatility matrix is shown in Table 3. The

T/K	2.00	3.00	4.00	5.00	6.00	8.00
1	0.325	0.244	0.19	0.165	0.174	0.22
1.5	0.372	0.295	0.237	0.196	0.198	0.223
2	0.374	0.299	0.246	0.208	0.205	0.224
3	0.347	0.283	0.241	0.213	0.205	0.212
4	0.325	0.266	0.228	0.204	0.196	0.201
5	0.307	0.252	0.217	0.196	0.189	0.192
6	0.294	0.241	0.208	0.189	0.182	0.184
7	0.283	0.232	0.201	0.183	0.176	0.176
8	0.274	0.225	0.194	0.177	0.17	0.169
9	0.267	0.219	0.189	0.172	0.164	0.162
10	0.262	0.215	0.184	0.167	0.159	0.156
12	0.251	0.206	0.177	0.16	0.151	0.147
15	0.238	0.195	0.167	0.151	0.142	0.137
20	0.226	0.184	0.157	0.141	0.133	0.13

Table 3: Subset out of 195 caplet volatilities σ_T^K(in %) for different strikes and different tenor dates (in years), 19.06.2008.

corresponding implied volatility surface is shown in Figure 2.
Pronounced smiles are clearly observable. Due to the structure of the given data, we are going to calibrate the stochastic volatility model based on semi-annual tenors, i.e. $\delta_j \equiv 0.5$, with $n = 41$, and where the initial calibration date 19.06.2008 is identified with $T_0 = 0$.
In a pre-calibration a standard market model is calibrated to ATM caps and ATM swaptions using Schoenmakers (2005). However, we emphasize that

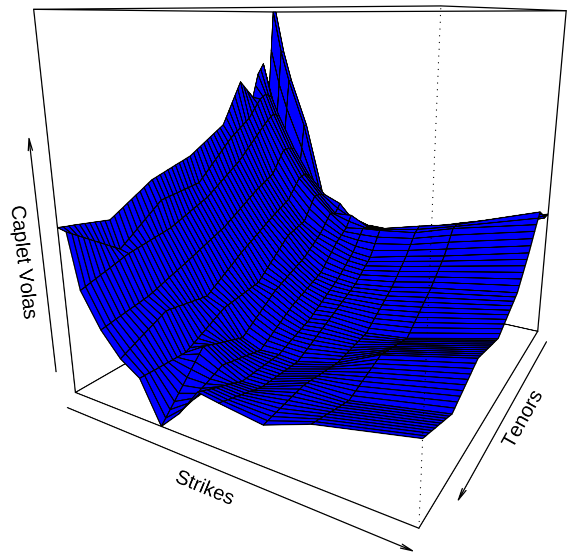

Figure 2: Caplet implied volatility surface σ_T^K.

the method by which this input market model is obtained is not essential nor a discussion point for this paper. For the pre-calibration we have used a volatility structure of the form

$$\gamma_i(t) = c_i g(T_i - t)\mathbf{e}_i, \quad 0 \le t \le \min(T_i, T_j), \quad 1 \le i, j < n,$$

where g is a simple parametric function and $\mathbf{e}_i$ are unit vectors. The pre-calibration routine returned $\mathbf{e}_i \in \mathbb{R}^{40}$ with

$$\mathbf{e}_i \cdot \mathbf{e}_j = \rho_{ij} =$$
$$\exp\left[-\frac{|i-j|}{n-1}\left(-\ln\rho_\infty + \eta\frac{i^2 + j^2 + ij - 3ni - 3nj - 3i - 3j + 3n + 2}{(n-2)(n-3)}\right)\right]$$
$$1 \le i, j < n$$

and $\rho_\infty = 0.23$, $\eta = 1.42$ such that the matrix $(\mathbf{e}_{i,j})$ is upper triangle. The function g is given by

$$g(s) = g_\infty + (1 - g_\infty + as)e^{-bs}.$$

with $a = 0.32$, $b = 0.07$, and $g_\infty = 0.58$. The loading factors c_i can be readily computed from

$$(\sigma_{T_i}^{ATM})^2 T_i = c_i^2 \int_0^{T_i} g^2(s)\, ds, \quad i = 1, \ldots, n-1,$$

using the initial LIBOR curve, which is obtained by a standard stripping procedure from the yield curve at 19.06.2008. Table 4 shows the calibrated values of c_i. Our calibration procedure delivers the following parameter values: $r = 0.24$ and ρ, σ, κ varying across maturities as shown in Table 5.

1	2	3	4	5	6	7	8	9	10
0.096	0.090	0.101	0.111	0.106	0.101	0.099	0.097	0.092	0.087
11	12	13	14	15	16	17	18	19	20
0.084	0.081	0.078	0.076	0.073	0.071	0.068	0.066	0.064	0.062
21	22	23	24	25	26	27	28	29	30
0.060	0.059	0.058	0.057	0.056	0.055	0.054	0.0534	0.0526	0.0518
31	32	33	34	35	36	37	38	39	40
0.051	0.050	0.050	0.049	0.049	0.048	0.049	0.048	0.047	0.047

Table 4: The values of loadings factors c_i calibrated to ATM caplets volatilities.

Tenor	20	19	18	17
ρ	-0.7832	-0.7832	-0.7832	-0.7832
σ	7.4920	7.4920	6.2427	5.0198
κ	2.3376	2.3376	3.9385	4.5590

Table 5: Parameters estimates for chosen tenors.

The quality of the calibration can be seen in Figure 3, where calibrated volatility curves are shown for several caplet periods together with the market caplet volas. The overall root-mean-square fit we have reached shows to be 0.5%-5%, when the caplet maturity ranges from 0.5 to 20.

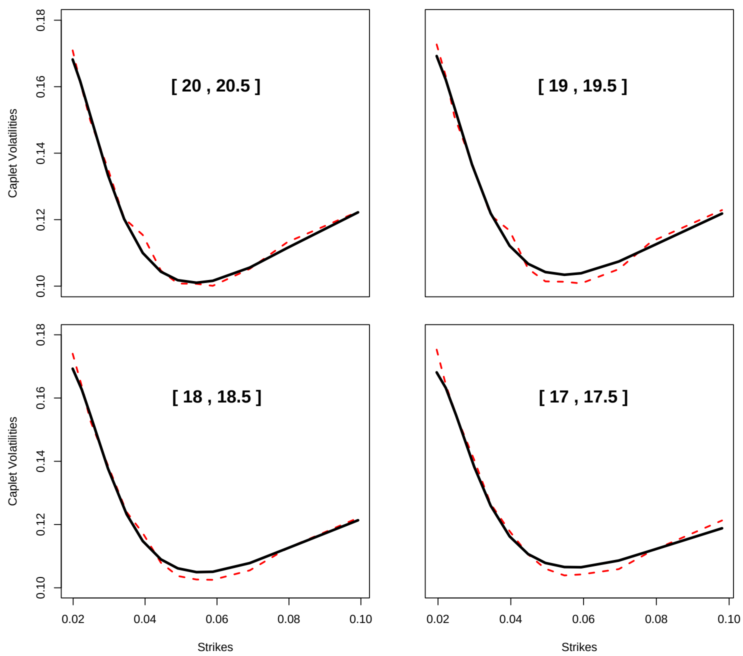

Figure 3: Caplet volas from the calibrated model (solid lines) and market caplets volas σ_T^K (dashed lines) for different caplet periods.

.

4 Multicurrency Extension of a Multiple Stochastic Volatility Libor Market Model

4.1 Introduction

Various authors have been actively working on the pricing of dual-currency and quanto style interest rate derivatives in the framework of a *non-stochastic* volatility LIBOR market model. See for example Pedersen and Miltersen (2000), Mikkelsen (2002) and Schloegl (2002). Of particular interest in their work were pricing formulae, which are notably appealing when lognormality can be assumed. In a non-stochastic volatility framework this requires that volatility processes are chosen deterministic. These processes are, specifically, the domestic and the foreign forward LIBOR volatilities, denoted by γ and γ^*, respectively, and an additional volatility process for the forward exchange rate, in the sequel denoted by σ_X. Schloegl (2002) and Mikkelsen (2002) show that assuming lognormal domestic and foreign forward LIBOR dynamics, the forward foreign exchange rate of *one maturity only* can be modeled consistently by a lognormal variable. In simulations we work exclusively in the terminal measure. Our approach is therefore in accordance with their result.

Since we intend to apply a stochastic volatility LIBOR model of our recent work, see Belomestny, Mathew and Schoenmakers (2007), to a multicurrency setting, lognormality is, however, not an issue in this case. The Feller/CIR-processes, we employ, render it a stochastic volatility model and unfortunately destroy any hope for lognormality. Recall that our goal is not to find analytic or semi-analytic solutions, but rather to generate a better smile and skew. Ultimately, prices are determined by a Monte Carlo simulation.

Apart from finding a better fitting model, our interest is to have one that can be calibrated in an efficient way. The calibration problem has been resolved in our recent work, see Belomestny/Mathew/Schoenmakers (2007). We want to profit from its demonstrated convenience also in the multicurrency setting.

A general modeling aspect for multicurrency products should not stay without mention. On one hand the model should be *realistic* in the sense that it incorporates attributes which are reflecting the economical circumstances. On the other hand we seek a model which allows a robust calibration procedure. The philosophical dilemma in our case comes down to decide for the dimension of the standard Brownian motion. Or, in other words, for the number of independent one-dimensional Brownian motions assumed as factors determining the stochastic dependence of the yield curve. Essentially the problem is to choose not too many of these in order to leave the model tractable and inline with empirical evidence, while having enough degrees of freedom for calibration purposes.

As an illustration, assume that interest rate curves are generally driven by a d-dimensional Brownian motion, where d is small compared to $2n$. For example $d = 3$, a value favoured by industry since over 90% of the variation is evidentally explained by these few factors. The SDE representing a market consisting of a domestic and a foreign yield curve would thus be of the form

$$\frac{dL_i}{L_i} = (...)dt + \gamma_i \cdot dW \,,$$
$$\frac{dL_i^*}{L_i^*} = (...)dt + \gamma_i^* \cdot dW \,, \quad i = 1, ..., n-1, \tag{94}$$

where the * adumbrates foreign interest LIBOR rates, as generally it will indicate the foreign market throughout this article. Mikkelsen and Schloegl use this set up. The usual inner market calibrations of domestic and foreign market will provide deterministic coefficients $\gamma \in \mathbb{R}^d$ and $\gamma^* \in \mathbb{R}^d$, respectively. If d is small, it is to be expected that both coefficient vectors will qualitatively resemble each other and thus lead to significant positive correlation among yield curves of two economies. A disputable stance!

On the other hand, correlation between the two markets is a necessary ingredient that can not be ignored, as supported by empirical evidence. Overall, we can state that in a unified market we need to incorporate *additional* sources of disturbance that perturb the foreign curve, but allow for correlation with the domestic curve. Evidentally, d should be chosen large enough. In order to calibrate efficiently and utilize the algorithm for the full factor model introduced in our former work, we have to choose $d = 2n - 2$ for the extension of the *standard* LIBOR model.

As we wish to extend the standard LIBOR model to one featuring stochastic volatility, in a setup we employed in our previous paper, the full factor model has to be enlarged by additional $2n - 2$ independent one-dimensional Brownian motions representing the correlated part of the CIR processes. Finally, the uncorrelated part will contribute another $2n - 2$ for both currencies. To cover the most general case, we therefore require the largest σ-field to be generated by $6(n - 1)$ independent one-dimensional Brownian motions.

The chapter is organized as follows. In Section 2 we supply the dynamics of the multicurrency extended standard LIBOR model and give a setup for which the primary goals, illustrated at the beginning of the section, are achieved. Sections 3 and 4 present some facts on the forward exchange rate and its parametrization.

In Section 5 we propose a sensible parametrization of the extended correlation matrix. Its (one) parameter is determined with the help of information from the currency market, as shortly recaped in Section 4. Its actual computation then is described in Section 6.

In the last three sections we deal with the stochastic volatility extended model. In Section 7 we first supply evidence that the arbitrage theory can be carried over to the stochastic volatility case. The arguments are similar to those in Sections 1.1.1 and 1.1.2. Full dynamics under various measures are presented in Section 8. Finally we illustrate in Section 9, how the calibration procedure from Belomestny/Mathew/Schoenmakers (2007) carries over to the multicurrency setting analyzed in this chapter.

4.2 Dynamics of the Extended Market

In this section we will exclusively refer to and work with a standard LIBOR model, not with the stochastic volatility version we actually intend to analyze in the sequel. The reason lies in the notational convenience we benefit from, when elaborating the point we wish to make.

Our goal for this section will be twofold:

(G.1) We will try to melt two markets, the domestic and the foreign market, into just one. That is, we wish to define the traded assets of this extended market and determine their *unified* dynamics described by an SDE.

(G.2) In doing this, however, we wish to obtain the feature that the volatility coefficients γ_i of the domestic LIBORS in this extended model are the same as the one from the domestic LIBOR model. In other words, when *restricting* the extended model to pricing domestic currency products, we find ourselves working within the original standard LIBOR model of the domestic currency.

The reason for the second requirement is that in practice we will have a fully calibrated home currency LIBOR model at our disposal and wish to utilize it to save time.

4.2.1 G.1

Regarding (G.1), as usual consider a tenor structure $0 = T_0 < T_1 < ... < T_n$, $\delta_i := T_{i+1} - T_i$. Let $(B_1, ..., B_n, B_1^*, ..., B_n^*)$ be an arbitrage free joint system of domestic zerobonds B_i and foreign zero bonds B_i^* expressed *in domestic currency*. We assume the coupled dynamics

$$\begin{aligned} \frac{dB_i}{B_i} &= \mu_i dt + \sigma_i \cdot dW \\ \frac{dB_i^*}{B_i^*} &= \mu_i^* dt + \sigma_i^* \cdot dW , \quad i = 1, ..., n. \end{aligned} \tag{95}$$

where W is a $2(n-1)$ dimensional standard Brownian motion, for reasons explained in the previous section.

Connected with (95) we have a general FX-LIBOR system $(L_1, ..., L_{n-1}, L_1^*, ..., L_{n-1}^*, X_n)$ defined by

$$L_i = \frac{1}{\delta_i}\left(\frac{B_i}{B_{i+1}} - 1\right), \quad L_i^* = \frac{1}{\delta_i}\left(\frac{B_i^*}{B_{i+1}^*} - 1\right), \quad X_n = \frac{B_n^*}{B_n}. \tag{96}$$

Set

$$\begin{aligned} L_i\gamma_i &:= \delta_i^{-1}\left(1+\delta_i L_i\right)\left(\sigma_i - \sigma_{i+1}\right), \\ L_i^*\gamma_i^* &:= \delta_i^{-1}(1+\delta_i L_i^*)\left(\sigma_i^* - \sigma_{i+1}^*\right) \\ \gamma_{X_n} &= \sigma_n^* - \sigma_n. \end{aligned} \tag{97}$$

Then with respect to B_n as numeraire we have the dynamics

$$\begin{aligned} \frac{dL_i}{L_i} &= -\sum_{j=i+1}^{n-1} \frac{\delta_j L_j}{1+\delta_j L_j}\gamma_j \cdot \gamma_i \, dt + \gamma_i \cdot dW^{(n)}, \quad 1 \le i < n, \\ \frac{dX_n}{X_n} &= \gamma_{X_n} \cdot dW^{(n)}, \end{aligned}$$

as can be seen from Schoenmakers (2005), equation (1.19). With respect to B_n^* as numeraire we have

$$\begin{aligned} \frac{dL_i^*}{L_i^*} &= -\sum_{j=i+1}^{n-1} \frac{\delta_j L_j^*}{1+\delta_j L_j^*}\gamma_j^* \cdot \gamma_i^* dt + \gamma_i^* \cdot dW^{(*n)}, \quad 1 \le i < n, \\ \frac{dX_n}{X_n} &= \left|\gamma_{X_n}\right|^2 dt + \gamma_{X_n} \cdot dW^{(*n)}, \end{aligned}$$

where $W^{(n)} \in \mathbb{R}^{2(n-1)}$ and $W^{(*n)} \in \mathbb{R}^{2(n-1)}$ are standard Brownian motions under P_n and P_{*n} respectively, with

$$dW^{(*n)} = dW^{(n)} - \gamma_{X_n} dt. \tag{98}$$

A more rigorous proof of (98) is provided in the next section. In an extended market we can therefore work in two possible settings. Under P_n the dynamics are

$$\begin{aligned} \frac{dL_i^*}{L_i^*} &= -\gamma_i^* \cdot \gamma_{X_n} dt - \sum_{j=i+1}^{n-1} \frac{\delta_j L_j^*}{1+\delta_j L_j^*}\gamma_j^* \cdot \gamma_i^* dt + \gamma_i^* \cdot dW^{(n)}, \\ \frac{dL_i}{L_i} &= -\sum_{j=i+1}^{n-1} \frac{\delta_j L_j}{1+\delta_j L_j}\gamma_j \cdot \gamma_i dt + \gamma_i \cdot dW^{(n)}, \quad 1 \le i < n, \\ \frac{dX_n}{X_n} &= \gamma_{X_n} \cdot dW^{(n)}. \end{aligned} \tag{99}$$

We could as well work under P_{*n} in which case

$$\begin{aligned}
\frac{dL_i^*}{L_i^*} &= -\sum_{j=i+1}^{n-1} \frac{\delta_j L_j^*}{1+\delta_j L_j^*} \gamma_j^* \cdot \gamma_i^* dt + \gamma_i^* \cdot dW^{(*n)}, \\
\frac{dL_i}{L_i} &= \gamma_i^* \cdot \gamma_{X_n} dt - \sum_{j=i+1}^{n-1} \frac{\delta_j L_j}{1+\delta_j L_j} \gamma_j \cdot \gamma_i dt + \gamma_i \cdot dW^{(*n)}, \quad 1 \le i < n, \\
\frac{dX_n}{X_n} &= \left|\gamma_{X_n}\right|^2 dt + \gamma_{X_n} \cdot dW^{(*n)}.
\end{aligned}$$

One conclusion of this is that the domestic zerobonds B_i are correlated with the foreign zerobonds B_i^* in domestic currency. Leaving this possibility out of consideration, would give rise to a very controversial discussion among practitioners. However, the correlation structures between the LIBORS need to be determined.

4.2.2 G.2

As a user of extended LIBOR models one will usually have a domestic market perspective and would like to find the feature that restricting the model to ones home currency yields consistent parameter values. We have in general no interest in the foreign market perspective, because accounting, balance sheet, financial and earnings reports are published in domestic currency. It will indeed be possible to satisfy these needs, when choosing the coefficients appropriately in the above setup. However, there will be no possibility in retrieving both, domestic and foreign, inner market calibrations at the same time, unless one assumes independence between the two markets. This will be the subject in a later section. For now consider the following setup:

Set the first $n-1$ coefficients of the domestic zerobonds to zero, $\sigma_{ij} = 0$ for $j = 1, ..., n-1$, and $i = 1, ..., n$. Explicitly we have

$$\begin{aligned}
\frac{dB_i}{B_i} &= \mu_i dt + \qquad 0 \qquad + \sum_{j=n}^{2n-2} \sigma_{ij} dW_j \,, \\
\frac{dB_i^*}{B_i^*} &= \mu_i^* dt + \sum_{j=1}^{n-1} \sigma_{ij}^* dW_j + \sum_{j=n}^{2n-2} \sigma_{ij}^* dW_j \,. \qquad (100)
\end{aligned}$$

This setup clearly allows for correlation between the two markets and contains the full $(n-1)$-factor LIBOR model when restricted to the domestic currency. From (97) we obtain

$$\frac{dL_i}{L_i} = (...)dt + \qquad 0 \qquad + \sum_{j=n}^{2n-2} \gamma_{ij} dW_j^{(n)} \,,$$

$$\frac{dL_i^*}{L_i^*} = (...)dt + \sum_{j=1}^{n-1} \gamma_{ij}^* dW_j^{(n)} + \sum_{j=n}^{2n-2} \gamma_{ij}^* dW_j^{(n)} \,,$$

for $1 \leq i < n$. We will demonstrate that a sensible parametrization can be determined which guarantees this setup and delivers that (G.2) is achieved.

4.3 The Forward Exchange Rate

In this section we consider a filtered probability space $(\Omega, \{\mathcal{F}\}_{t\in[0,T_n]}, P_n)$, where the underlying measure is immediately chosen to be the terminal measure. The model is set up on the basis of assumptions (BP.1) and (BP.2) of Musiela and Rutkowski (1997). We repeat them briefly:

(BP.1) For any date $T \in [0, T_n]$, the price process of a zero coupon bond $B(t, T)$, $t \in [0, T]$, is a strictly positive semimartingale under P_n.

(BP.2) For any fixed $T \in [0, T_n]$, the *forward process*

$$F_B(t, T, T_n) = \frac{B(t, T)}{B(t, T_n)} \,, \quad \forall t \in [0, T]$$

follows a martingale under P_n.

Assumption (BP.2) is justified in view of absence of arbitrage. Discounted by the last bond as numeraire, the securities are martingales. The discrete-tenor case has been shown in earlier sections. Analogously these assumptions have to hold for the foreign security market:

(BP.1) For any date $T \in [0, T_n]$, the price process of a zero coupon bond $B^*(t, T)$, $t \in [0, T]$, is a strictly positive semimartingale under P_{*n}.

(BP.2) For any fixed $T \in [0, T_n]$, the *forward process*

$$F_{B^*}(t, T, T_n) = \frac{B^*(t, T)}{B^*(t, T_n)} \,, \quad \forall t \in [0, T]$$

follows a martingale under P_{*n}.

P_{*n} denotes the terminal measure in the foreign currency.

The objects of interest in the fixed income market are the δ-compounded forward rates defined by

$$L(t,T) = \delta^{-1}\left(\frac{B(t,T)}{B(t,T+\delta)} - 1\right),$$

and

$$L^*(t,T) = \delta^{-1}\left(\frac{B^*(t,T)}{B^*(t,T+\delta)} - 1\right),$$

which are referred to as forward LIBOR rates. Note that by assumption (BP.2) $L(t, T_n - \delta)$ is a martingale under P_n. Returning to the discrete-tenor case from now, this means that the dynamics of the domestic and the foreign forward LIBOR rates with the longest maturity under their respective terminal measures are given by

$$dL_{n-1}(t) = L_{n-1}(t)\gamma_{n-1}(t)dW^{(n)}(t),$$

$$dL^*_{n-1}(t) = L^*_{n-1}(t)\gamma^*_{n-1}(t)dW^{(*n)}(t),$$

where $W^{(n)}$ and $W^{(*n)}$ are standard Brownian motions under P_n and P_{*n} respectively.

A fundamental challenge in pricing cross-currency products in LIBOR models is to find the dynamics of the foreign forward LIBOR processes under a unifying measure in the domestic currency. This can be accomplished in various ways, as demonstrated in the paper of Schloegl (2002). In his monograph he shows that the bridge between the two currencies can be built at any time point T_i. Thereby he employs that LIBOR dynamics can be given under any measure P_i, the martingale measure corresponding to the numeraire B_i. Since the stochastic model analyzed here is fully specified in the terminal measure, we consider only the case P_n.

Naturally, the first step in converting the foreign currency assets into the domestic currency unit is to introduce a *spot exchange rate* process ζ. In order to satisfy regularity conditions we require the following:

(S.1) The spot exchange rate process $\zeta(t)$, $t \in [0, T_n]$, is a strictly positive semimartingale under P_{*n}.

In this work we consider ζ to be given in terms of units of domestic currency per one unit of foreign currency.

If only zerobonds are used to define the market, ζ is not a tradable asset in either market, domestic or foreign. As such it will in general not be a martingale under any martingale measure when discounted by the corresponding numeraire. However, the foreign bond converted to domestic currency at the spot exchange rate can be regarded as a domestic asset when the market is defined appropriately. Specifically, let the market consist of domestic and converted foreign zerobonds, i.e.

$$\{B(t,T_i)\,,\; B^*(t,T_i) = \widetilde{B}(t,T_i)\zeta(t)\,, \quad i = 1,..,n\},$$

then

$$X(t,T_i) := \frac{B^*(t,T_i)}{B(t,T_i)} \tag{101}$$

is a martingale under P_i for all $i = 1,..,n$. The process $X(t,T_i)$ is the *time T_i forward exchange rate.* Since we intend to establish the link between the currencies with respect to the terminal measure, let us in particular set

$$dX(t,T_n) := X(t,T_n)\gamma_{X_n}(t,T_n)dW^{(n)}(t), \tag{102}$$

where the existence of a process γ_{X_n} (with appropriate properties) is ensured by the martingale representation theorem. Its parameterization and calibration is a problem of own relevance in currency markets and is based on data from currency options.
The following proposition is key for the measure change between markets.

Proposition 58 *Let P_n and P_{*n} be the respective measures with B_n and B_n^* as numeraires, then*

1. *their density is given by*

$$\frac{dP_{*n}}{dP_n}|_{\mathcal{F}_t} = \frac{X(t,T_n)}{X(0,T_n)}\,, \quad w.p.1,$$

2. *Brownian motions under the two measures are related by*

$$dW^{(*n)}(t) = dW^{(n)}(t) - \gamma_{X_n}(t,T_n)dt. \tag{103}$$

Proof. 1. For all assets A_t in the domestic market we have

$$E_n\left(\frac{A_t}{B(t,T_n)}|\mathcal{F}_0\right) = \frac{A_0}{B(0,T_n)}.$$

This holds in particular for $A_t^* = \widetilde{A}_t \zeta(t)$. Thus on one hand we have

$$E_n \left(\frac{A_t^*}{B(t,T_n)} | \mathcal{F}_0 \right) = \frac{A_0^*}{B(0,T_n)}$$

$$\Leftrightarrow E_n \left(\frac{A_t^*}{B^*(t,T_n)} \frac{B^*(t,T_n)}{B(t,T_n)} | \mathcal{F}_0 \right) = \frac{A_0^*}{B^*(0,T_n)} \frac{B^*(0,T_n)}{B(0,T_n)}$$

$$\Leftrightarrow E_n \left(\frac{A_t^*}{B^*(t,T_n)} \frac{B^*(t,T_n)}{B(t,T_n)} \frac{B(0,T_n)}{B^*(0,T_n)} | \mathcal{F}_0 \right) = \frac{A_0^*}{B^*(0,T_n)}$$

$$\Leftrightarrow E_n \left(\frac{A_t^*}{B^*(t,T_n)} \frac{X(t,T_n)}{X(0,T_n)} | \mathcal{F}_0 \right) = \frac{A_0^*}{B^*(0,T_n)}.$$

From the definition of a measure change density, on the other hand we have

$$E_{*n} \left(\frac{A_t^*}{B^*(t,T_n)} | \mathcal{F}_0 \right) = \frac{A_0^*}{B^*(0,T_n)}$$

$$\Leftrightarrow E_n \left(\frac{A_t^*}{B^*(t,T_n)} \frac{dP_{*n}}{dP_n} |_{\mathcal{F}_t} | \mathcal{F}_0 \right) = \frac{A_0^*}{B^*(0,T_n)}.$$

The only way

$$E_n \left(\frac{A_t^*}{B^*(t,T_n)} \left(\frac{dP_{*n}}{dP_n} |_{\mathcal{F}_t} - \frac{X(t,T_n)}{X(0,T_n)} \right) | \mathcal{F}_0 \right) = 0$$

can hold for all foreign assets A_t^*, is if

$$\frac{dP_{*n}}{dP_n} |_{\mathcal{F}_t} = \frac{X(t,T_n)}{X(0,T_n)}, \quad \text{w.p.1.}$$

In a sense, the uniqueness of Radon-Nikodym densities concludes the proof.
2. Apply Itô's Lemma to $1/X_n$. On one hand we have, since $1/X_n = B_n/B_n^*$,

$$d\left(\frac{1}{X_n} \right) = \frac{1}{X_n} (\sigma_n - \sigma_n^*) \cdot dW^{(*n)} = \frac{1}{X_n} (-\gamma_{X_n}) \cdot dW^{(*n)}$$

under P_{*n}. On the other hand, by Itô's product rule applied to $1/X_n$ and (102), we have

$$d\left(\frac{1}{X_n} \right) = \frac{1}{X_n} \left(\gamma_{X_n} \cdot \gamma_{X_n} \, dt - \gamma_{X_n} \cdot dW^{(n)} \right).$$

■

Relation (103) provides the final justification for (98).

4.4 Parametrization of the Extended Correlation Matrix

Recall the following facts from the parametrization of the correlation matrices in the standard $(n-1)$-full factor LIBOR models. Denote by $\gamma_i^s \in \mathbb{R}^{n-1}$ the deterministic volatility coefficients of the standard full factor LIBOR model in the domestic market, and by $\gamma_i^{*s} \in \mathbb{R}^{n-1}$ those in the foreign market.

(i) The scalar volatility processes from the inner calibrations are given by

$$\eta_i^s := |\gamma_i^s| \quad \text{and} \quad \eta_i^{*s} := |\gamma_i^{*s}|.$$

Their parameterization was established via

$$\eta_i^s(t) := c_i g(T_i - t)\,, \quad \text{and} \quad \eta_i^{*s}(t) := c_i^* g^*(T_i - t)\,,$$

where g and g^* are functions of the form $g(x) := g_\infty + (1 - g_\infty + ax)e^{-bx}$ with parameters a, b and g_∞, see Schoenmakers (2005), eqn. (2.25).

(ii) Correlation structures $\rho^s = \rho^s(t)$ and $\rho^{*s} = \rho^{*s}(t)$ are then given by

$$\rho_{ij}^s(t) := \frac{\gamma_i^s \cdot \gamma_j^s}{|\gamma_i^s||\gamma_j^s|}(t)\,, \quad \text{and} \quad \rho_{ij}^{*s}(t) := \frac{\gamma_i^{*s} \cdot \gamma_j^{*s}}{|\gamma_i^{*s}||\gamma_j^{*s}|}(t)\,, \tag{104}$$

for $0 \le t \le \min(T_i, T_j)$ and $1 \le i, j \le n-1$. Furthermore

$$\gamma_i^s = \eta_i^s \mathbf{e}_i^s\,, \quad \text{and} \quad \gamma_i^{*s} = \eta_i^{*s} \mathbf{e}_i^{*s},$$

where $\mathbf{e}_i^s \in \mathbb{R}^{n-1}$ and $\mathbf{e}_i^{*s} \in \mathbb{R}^{n-1}$ are sets of decomposing unit vectors. (Schoenmakers, Section 2.2.4.)

In order to guarantee G.2, two things are immediate:

1. In calibration of g, $g(x) := g_\infty + (1 - g_\infty + ax)e^{-bx}$, the parameters a, b and g_∞ remain only the same, if the sum in

$$\min \sum_{i=1}^{n-1} \left(\widehat{\sigma}_i^B T_i - c^2 \int_0^{T_i} g^2(x)dx \right)$$

is taken, as indicated, over $n-1$ *domestic* caplets. We thus stay within one currency, see Schoenmakers (2005, page 32). Accordingly, we need another optimization for the foreign caplets, say

$$\min \sum_{i=1}^{n-1} \left(\widehat{\sigma}_i^{*B} T_i - c^{*2} \int_0^{T_i} g^{*2}(x)dx \right).$$

As a result we have two functions, g and g^*, one for each currency.

2. The dimension of the vectors $\gamma_i \in \mathbb{R}^{2n-2}$ and $\gamma_i^* \in \mathbb{R}^{2n-2}$ is twice as large as the one from the inner $(n-1)$-dimensional full factor models within each particular market. Equality between the two γ-sets can thus be accomplished only by using unit vectors $\mathbf{e}_i$ that are "equal" in the sense that nontrivial entries agree. Other entries, particularly the ones accounting for the difference in dimension, must be zero.

The first point does not pose a problem. The calibration has to be done accordingly. Regarding the second point, we consider correlation matrices of the following form.

$$\mathcal{R}_\rho = \begin{pmatrix} R^* & | & \rho M \\ --- & & --- \\ \rho M & | & R \end{pmatrix} \in \mathbb{R}^{(2n-2)\times(2n-2)}.$$

The upper left blockmatrix R^* denotes the foreign, the lower right blockmatrix R the domestic correlation matrix. The off-diagonal blockmatrices represent the correlation between inter-market LIBOR rates. We will suggest two possible forms of the matrix M at the end of this subsection. For now any matrix M is feasible, as long as the over-all matrix $\mathcal{R}_\rho$ stays a correlation matrix.

Since we assume to always be working with full rank matrices, we can Cholesky decompose $\mathcal{R}_\rho$ into a product of an upper and a lower triangular matrix

$$\mathcal{R}_\rho = \begin{pmatrix} * & * & \cdots & * \\ & * & & * \\ & \mathbf{0} & \ddots & \vdots \\ & & & * \end{pmatrix} \cdot \begin{pmatrix} * & & & \\ * & * & \mathbf{0} & \\ \vdots & & \ddots & \\ * & * & \cdots & * \end{pmatrix}.$$

Denoting the i-th row vector of the upper triangular matrix by $\mathbf{e}_i^\top \in \mathbb{R}^{2n-2}$, we obtain the following representation

$$\mathcal{R}_\rho = \begin{pmatrix} \mathbf{e}_1^\top \\ \mathbf{e}_2^\top \\ \vdots \\ \mathbf{e}_{2n-2}^\top \end{pmatrix} \cdot (\mathbf{e}_1 \, , \, \mathbf{e}_2 \, , \cdots , \, \mathbf{e}_{2n-2}) \, .$$

If, for $i = 1, ..., n-1$, we set

$$\gamma_i^*(t) = \eta_i^*(t)\mathbf{e}_i(t) \, , \quad \text{and} \quad \gamma_i(t) = \eta_i(t)\mathbf{e}_{n-1+i}(t) \, ,$$

where $\mathbf{e}_i \in \mathbb{R}^{2n-2}$, for $i = 1, ..., 2n-2$, we indeed obtain a set of decomposing unit vectors as required. Furthermore, note that at least the first $n-1$ components of $\gamma_i \in \mathbb{R}^{2n-2}$ are zero. Truncating away these first $n-1$ components, the restricted $\gamma_i^{trunc}(t) \in \mathbb{R}^{n-1}$ must be the ones identical to those from the inner $(n-1)$-factor LIBOR model, $\gamma_i^{trunc} = \gamma_i^s$. This follows directly from the Cholesky decomposition and the setup chosen in subsection G.2. Whereas R is generated from vectors whose first $n-1$ entries are zero, the upper diagonal blockmatrix R^* does not have this feature. It is, of course, a correlation matrix also, but by the way it is decomposed, it can not have been generated only by the foreign *inner* LIBOR model. The coefficient vectors $\gamma^* \in \mathbb{R}^{2n-2}$ are "putting weight" on more than $n-1$ of the one-dimensional Brownian motions. The domestic *inner* LIBOR model interfers with its correlations.

At this stage the conjecture comes to mind, that in order to be able to utilize both inner calibrated models, we need to assume independence between the two markets. Only if $\rho = 0$, the correlation matrix of the foreign LIBOR rates R^* also can be generated by $n-1$-dimensional vectors $\eta_i^* \mathbf{e}_i$.

Two Examples: 1. Put $M = \mathbf{1} \in \mathbb{R}^{n-1 \times n-1}$, the matrix containing only ones. The resulting correlation matrix is the case, where for simplicity an average correlation ρ is assumed between all cross-currency rates. Thus $\mathcal{R}_\rho =$

$$\begin{pmatrix} \frac{\gamma_1^{*\top}\gamma_1^*}{|\gamma_1^*||\gamma_1^*|} & \cdots & \frac{\gamma_1^{*\top}\gamma_{n-1}^*}{|\gamma_1^*||\gamma_{n-1}^*|} & | & \rho & \cdots & \rho \\ \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\ \frac{\gamma_{n-1}^{*\top}\gamma_1^*}{|\gamma_{n-1}^*||\gamma_1^*|} & \cdots & \frac{\gamma_{n-1}^{*\top}\gamma_{n-1}^*}{|\gamma_{n-1}^*||\gamma_{n-1}^*|} & | & \rho & \cdots & \rho \\ --- & --- & --- & | & --- & --- & --- \\ \rho & \cdots & \rho & | & \frac{\gamma_1^{\top}\gamma_1}{|\gamma_1||\gamma_1|} & \cdots & \frac{\gamma_1^{\top}\gamma_{n-1}}{|\gamma_1||\gamma_{n-1}|} \\ \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\ \rho & \cdots & \rho & | & \frac{\gamma_{n-1}^{\top}\gamma_1}{|\gamma_{n-1}||\gamma_1|} & \cdots & \frac{\gamma_{n-1}^{\top}\gamma_{n-1}}{|\gamma_{n-1}||\gamma_{n-1}|} \end{pmatrix}.$$

For the "average"-ρ off-diagonal matrix, investigated here, it is shown in the appendix that if ρ is chosen carefully,

$$\rho \leq 1/(n-1) \cdot \min(\alpha_i, \beta_i),$$

where α_i and β_i are the eigenvalues of R^* and R respectively, we again obtain a correlation matrix for the extended market. Unfortunately, for large n or some small eigenvalue, this bound is very restrictive in allowing only a very small ρ. The next example is therefore more appropriate.

2. Set

$$M = \frac{R + R^*}{2} \in \mathbb{R}^{n-1 \times n-1}. \tag{105}$$

This is our favorite parametrization for the extended correlation matrix. In the appendix we indicate briefly, that matrices of this form (105) are indeed positive definite correlation matrices.

It is not the aim of this thesis to go into detail in this regard. In any case, if the correlation matrix $\mathcal{R}_\rho$ is not to be assumed exogeneously given, we have to find a sensible calibration algorithm for the parameter ρ from market data.

4.5 Determination of ρ

Before we proceed to give a formula for the yet unknown parameter ρ, a fundamental concept needs to be elaborated. Denote by $\gamma_i^{*s} \in \mathbb{R}^{n-1}$ the deterministic volatility coefficients of the standard full factor LIBOR model in the foreign currency. Its dimension is clearly exactly half of our $\gamma_i^* \in \mathbb{R}^{2(n-1)}$ in the extended market. Though no immediate relation among the entries is identifiable, it is plausible that:

1. The correlation structures of the foreign market must be model independent.
2. The absolute volatilities in the two models should coincide.

We therefore require

$$\frac{\gamma_k^* \cdot \gamma_l^*}{|\gamma_k^*||\gamma_l^*|} = \frac{\gamma_k^{*s} \cdot \gamma_l^{*s}}{|\gamma_k^{*s}||\gamma_l^{*s}|} \tag{106}$$

to hold for all $k, l = 1, ..., n-1$. From 2. we even have

$$|\gamma_k^*| = |\gamma_k^{*s}|, \quad \text{and thus} \quad \gamma_k^* \cdot \gamma_l^* = \gamma_k^{*s} \cdot \gamma_l^{*s}, \tag{107}$$

again for all $k, l = 1, ..., n-1$. Since the RHSs are given from a calibrated foreign standard LIBOR model, we can freely dispose of the quantities on the LHSs.

For a more general setting, set $X_i = B_i^*/B_i$ for $1 \leq i < n$,. Note that (see Glasserman (2004), page 168)

$$B_i(t) = B_{\eta(t)}(t) \prod_{j=\eta(t)}^{i-1} \frac{1}{1 + \delta_j L_j(t)},$$

for $0 \leq t \leq T_i$, and $\eta(t) := \min\{m : T_m \geq t\}$. Thus

$$\begin{aligned} \frac{dB_i}{B_i} &= (...)dt + d\ln B_i \\ &= (...)dt + \frac{dB_{\eta(t)}}{B_{\eta(t)}} - \sum_{j=\eta(t)}^{i-1} \frac{\delta_j L_j}{1 + \delta_j L_j} \gamma_j \cdot dW^{(\cdot)} \\ &=: (...)dt + \sigma_{\eta(t)} \cdot dW^{(\cdot)} - \sum_{j=\eta(t)}^{i-1} \frac{\delta_j L_j}{1 + \delta_j L_j} \gamma_j \cdot dW^{(\cdot)}. \end{aligned}$$

In the same way, for $t \leq T_i$,

$$B_i^*(t) = B_{\eta(t)}^*(t) \prod_{j=\eta(t)}^{i-1} \frac{1}{1 + \delta_j L_j^*(t)}$$

and

$$\begin{aligned} \frac{dB_i^*}{B_i^*} &= (...)dt + d\ln B_i^* \\ &= (...)dt + \sigma_{\eta(t)}^* \cdot dW^{(\cdot)} - \sum_{j=\eta(t)}^{i-1} \frac{\delta_j L_j^*}{1 + \delta_j L_j^*} \gamma_j^* \cdot dW^{(\cdot)}. \end{aligned}$$

Note that $B_i^*(t)$ is the foreign bond expressed in domestic currency. We may set for $t \leq T_i$, $B_i^*(t) =: \zeta(t)\widetilde{B}_i(t)$, where the $\widetilde{B}_i(t)$ are foreign bonds expressed in foreign currency and $\zeta(t)$ is the FX spot rate. In particular we have $B_i^*(T_i) = \zeta(T_i)$, and

$$\zeta(t) = \frac{B_i^*(t)}{\widetilde{B}_i(t)} = \frac{B_{\eta(t)}^*(t)}{\widetilde{B}_{\eta(t)}(t)}.$$

Let $X_i := B_i^*/B_i$, thus

$$\frac{dX_i}{X_i} = (...)dt + \left(\sigma_{\eta(t)}^* - \sigma_{\eta(t)}\right) \cdot dW^{(\cdot)} + \sum_{j=\eta(t)}^{i-1} \left(\frac{\delta_j L_j \gamma_j}{1 + \delta_j L_j} - \frac{\delta_j L_j^* \gamma_j^*}{1 + \delta_j L_j^*} \right) \cdot dW^{(\cdot)}$$

$$= (...)dt + \gamma_{X_i} \cdot dW^{(\cdot)} = \gamma_{X_i} \cdot dW^{(i)},$$

with

$$\gamma_{X_i} = \sigma_{\eta(t)}^* - \sigma_{\eta(t)} + \sum_{j=\eta(t)}^{i-1} \left(\frac{\delta_j L_j \gamma_j}{1 + \delta_j L_j} - \frac{\delta_j L_j^* \gamma_j^*}{1 + \delta_j L_j^*} \right) \tag{108}$$

and

$$\gamma_k^* \cdot \gamma_{X_i} = \gamma_k^* \cdot \left(\sigma_{\eta(t)}^* - \sigma_{\eta(t)}\right) + \sum_{j=\eta(t)}^{i-1} \left(\frac{\delta_j L_j \gamma_k^* \cdot \gamma_j}{1 + \delta_j L_j} - \frac{\delta_j L_j^* \gamma_k^* \cdot \gamma_j^*}{1 + \delta_j L_j^*} \right).$$

If the spot rate is assumed to follow $d\zeta/\zeta =: (...)dt + \sigma_\zeta \cdot dW^{(\cdot)}$, note that we have $\sigma_{\eta(t)}^* = \widetilde{\sigma}_{\eta(t)} + \sigma_\zeta$. By the definition of η, the bond $B_{\eta(t)}$ is always located shortly before its own maturity, never more then one δ- period. In a good approximation, we can thus set $\sigma_{\eta(t)} = \widetilde{\sigma}_{\eta(t)} = 0$. We then obtain

$$\gamma_k \cdot \sigma_{\eta(t)}^* =: r_k \left|\gamma_k\right| \left|\sigma_{\eta(t)}^*\right| \approx r_k \left|\gamma_k\right| \left|\sigma_\zeta\right| = \gamma_k \cdot \sigma_\zeta$$

and

$$\gamma_k^* \cdot \sigma_{\eta(t)}^* =: r_k^* \left|\gamma_k^*\right| \left|\sigma_{\eta(t)}^*\right| \approx r_k^* \left|\gamma_k\right| \left|\sigma_\zeta\right| = \gamma_k^* \cdot \sigma_\zeta.$$

The newly introduced parameters r_k and r_k^* can be interpreted as the correlation between the k-th, domestic respectively foreign, LIBOR and the spot rate ζ. For the correlation processes of X_i with the k-th foreign LIBOR rate, we conclude

$$\begin{aligned}\gamma_k^* \cdot \gamma_{X_i} &\approx \gamma_k^* \cdot \sigma_\zeta + \sum_{j=\eta(t)}^{i-1} \left(\frac{\delta_j L_j \gamma_k^* \cdot \gamma_j}{1+\delta_j L_j} - \frac{\delta_j L_j^* \gamma_k^* \cdot \gamma_j^*}{1+\delta_j L_j^*} \right) \\ &= \gamma_k^* \cdot \sigma_\zeta + \rho \sum_{j=\eta(t)}^{i-1} \frac{\delta_j L_j}{2\left(1+\delta_j L_j\right)} \left|\gamma_k^*\right| \left|\gamma_j\right| \left(R_{kj} + R_{kj}^*\right) \\ &\quad - \sum_{j=\eta(t)}^{i-1} \frac{\delta_j L_j^*}{1+\delta_j L_j^*} \left|\gamma_k^*\right| \left|\gamma_j^*\right| R_{kj}^*. \end{aligned} \tag{109}$$

For the variance process of X_i we have

$$\begin{aligned}\gamma_{X_i} \cdot \gamma_{X_i} = \left|\gamma_{X_i}\right|^2 \approx &\left|\sigma_\zeta\right|^2 + \\ &+ 2 \sum_{j=\eta(t)}^{i-1} \left(\frac{\delta_j L_j}{1+\delta_j L_j} \left|\gamma_j\right| \left|\sigma_\zeta\right| r_j - \frac{\delta_j L_j^*}{1+\delta_j L_j^*} \left|\gamma_j^*\right| \left|\sigma_\zeta\right| r_j^* \right) \\ &+ \sum_{k,j=\eta(t)}^{i-1} \frac{\delta_k \delta_j L_k L_j}{\left(1+\delta_k L_k\right)\left(1+\delta_j L_j\right)} \left|\gamma_k\right| \left|\gamma_j\right| R_{kj} \\ &+ \sum_{k,j=\eta(t)}^{i-1} \frac{\delta_k \delta_j L_k^* L_j^*}{\left(1+\delta_k L_k^*\right)\left(1+\delta_j L_j^*\right)} \left|\gamma_k^*\right| \left|\gamma_j^*\right| R_{kj}^* \\ &- \rho \sum_{k,j=\eta(t)}^{i-1} \frac{\delta_k \delta_j L_k^* L_j}{\left(1+\delta_k L_k^*\right)\left(1+\delta_j L_j\right)} \left|\gamma_k^*\right| \left|\gamma_j\right| \left(R_{kj} + R_{kj}^*\right). \end{aligned} \tag{110}$$

We propose to calibrate the remaining unknown parameter ρ in (110) to standard options in the FX market as follows. Since X_i is a martingale in the measure P_i, its dynamics are given by

$$\frac{dX_i}{X_i} = \gamma_{X_i} \cdot dW^{(i)}.$$

Suppose we have a call option to exchange one unit of foreign currency for K units of domestic currency at time T_i. Clearly, the net payoff of this option is $(K - B_i^*(T_i))^+$, and the option value in domestic currency at time $t \leq T_i$ is given by

$$\begin{aligned} C_i(t, K) &:= B_i(t) E(K - B_i^*(T_i))^+ \\ &= B_i(t) E(K - \zeta(T_i))^+. \end{aligned}$$

Based on a deterministic approximation $|\widetilde{\gamma}_X|$ of $|\gamma_X|$ to be determined below, it thus follows that $C_i(0, K)$ can be evaluated using Black's 76 formula with input volatility $\sigma_{X_i}(0, K)$, where

$$\sigma_{X_i}^2(0, K) := \frac{1}{T_i} \int_0^{T_i} |\widetilde{\gamma}_{X_i}|^2 (s) ds.$$

The deterministic approximation of $|\gamma_{X_i}|$ is obtained by freezing as usual, where $|\sigma_\zeta| \equiv \overline{\sigma}_\zeta$, $r_i \equiv r_d$, and $r_j^* \equiv r_f^*$ are assumed to be constant for simplicity,

$$\begin{aligned} |\widetilde{\gamma}_{X_i}|^2 (s) = \overline{\sigma}_\zeta^2 + & \\ & + 2\overline{\sigma}_\zeta \sum_{j=\eta(s)}^{i-1} \left(\left[\frac{\delta_j r_d L_j}{1 + \delta_j L_j} \right] (0) |\gamma_j| (s) - \left[\frac{\delta_j r_f^* L_j^*}{1 + \delta_j L_j^*} \right] (0) \left|\gamma_j^*\right| (s) \right) \\ & + \sum_{k,l=\eta(s)}^{i-1} \left[\frac{\delta_k \delta_l L_k L_l}{(1 + \delta_k L_k)(1 + \delta_l L_l)} \right] (0) \left(|\gamma_k| |\gamma_l| R_{kl} \right) (s) \\ & + \sum_{k,l=\eta(s)}^{i-1} \left[\frac{\delta_k \delta_l L_k^* L_l^*}{(1 + \delta_k L_k^*)(1 + \delta_l L_l^*)} \right] (0) \left(|\gamma_k^*| |\gamma_l^*| R_{kl}^* \right) (s) \qquad (111) \\ & - \rho \sum_{k,l=\eta(s)}^{i-1} \left[\frac{\delta_k \delta_l L_k^* L_l}{(1 + \delta_k L_k^*)(1 + \delta_l L_l)} \right] (0) \left(|\gamma_k^*| |\gamma_l| (R_{kl} + R_{kl}^*) \right) (s). \end{aligned}$$

Thus it is natural to infer $\sigma_{X_i}(0, K)$ from the FX market and then calibrate parameters $\rho, \overline{\sigma}_\zeta, r_d$, and r_f^* from (4.5) using (111) under the restriction that $\mathcal{R}_\rho$ remains a valid correlation matrix.

From (109), with $i = n$,

$$\begin{aligned} \gamma_k^* \cdot \gamma_{X_n} = r_f^* \overline{\sigma}_\zeta |\gamma_k^*| + \rho \sum_{l=\eta(t)}^{n-1} \frac{\delta_l L_l}{2(1 + \delta_l L_l)} |\gamma_k^*| |\gamma_l| (R_{kl} + R_{kl}^*) \\ - \sum_{l=\eta(t)}^{n-1} \frac{\delta_l L_l^*}{1 + \delta_l L_l^*} |\gamma_k^*| |\gamma_l^*| R_{kl}^*, \end{aligned}$$

we therefore obtain the dynamics of all LIBORS in the terminal measure:

$$\begin{aligned}
\frac{dL_i^*}{L_i^*} &= -r_f^* \overline{\sigma}_\zeta \left|\gamma_i^*\right| dt \\
&\quad - \rho \sum_{l=\eta(t)}^{n-1} \frac{\delta_l L_l}{2\left(1+\delta_l L_l\right)} \left|\gamma_i^*\right| \left|\gamma_l\right| \left(R_{il} + R_{il}^*\right) dt \\
&\quad + \sum_{l=\eta(t)}^{n-1} \frac{\delta_l L_l^*}{1+\delta_l L_l^*} \left|\gamma_i^*\right| \left|\gamma_l^*\right| R_{il}^* dt \\
&\quad - \sum_{j=i+1}^{n-1} \frac{\delta_j L_j^*}{1+\delta_j L_j^*} \gamma_j^* \cdot \gamma_i^* dt + \gamma_i^* \cdot dW^{(n)}, \\
\frac{dL_i}{L_i} &= -\sum_{j=i+1}^{n-1} \frac{\delta_j L_j}{1+\delta_j L_j} \gamma_j \cdot \gamma_i dt + \gamma_i \cdot dW^{(n)}, \quad 1 \leq i < n, \\
\frac{dX_n}{X_n} &= \gamma_{X_n} \cdot dW^{(n)}.
\end{aligned} \tag{112}$$

The quantity $-r_f^* \overline{\sigma}_\zeta \left|\gamma_i^*\right|$ may be considered a deterministic correction to the foreign yield under the terminal measure.

4.6 Arbitrage Theory for Stochastic Volatility

For the concepts in this section we work in a probability space $(\Omega, \mathcal{F}, P)$ with filtration $\mathcal{F} = (\mathcal{F}_t)_{0\le t\le T<\infty}$ generated by an $\mathbb{R}^{4(n-1)}$-valued Brownian motion $\mathcal{W}$ which satisfies the "usual conditions". On this space consider an $2n$-dimensional process $\mathcal{B} = (\mathcal{B}(t))_{0\le t\le T}$ of n domestic and n foreign tradable securities converted to domestic currency. Therefore define a fixed sequence of tenor dates $0 =: T_0 < T_1 < \ldots < T_n$, called a tenor structure, together with a sequence of day-count fractions $\delta_i := T_{i+1} - T_i,\ i = 1, \ldots, n-1$. With respect to this tenor structure we consider zerobond processes $B_i := B(\cdot, T_i),\ i = 1, \ldots, n$, where each B_i lives on the interval $[0, T_i]$ and ends up with face value $B_i(T_i) = 1$. Let the first n components are given by the SDE

$$\begin{aligned}\frac{dB_i}{B_i} &= \mu_i dt + \sigma_i \cdot d\mathcal{W}, \qquad B_i(0) > 0, \quad i = 1, ..., n, \\ &= \mu_i dt + \sum_{j=1}^{4(n-1)} \sigma_{ij} d\mathcal{W}_j, \end{aligned} \tag{113}$$

representing the domestic LIBOR market as proposed by Belomestny, Mathew, Schoenmakers (2007). The last n components are given by the SDE

$$\begin{aligned}\frac{dB_i^*}{B_i^*} &= \mu_{n+i} dt + \sigma_{n+i} \cdot d\mathcal{W}, \qquad B_i^*(0) > 0, \quad i = 1, ..., n, \\ &= \mu_{n+i} dt + \sum_{j=1}^{4(n-1)} \sigma_{n+i,j} d\mathcal{W}_j. \end{aligned} \tag{114}$$

Or simply, in a one block notation with straightforward μ and σ:

$$\frac{d\mathcal{B}}{\mathcal{B}} = \mu dt + \sigma \cdot d\mathcal{W},$$

where

$$d\mathcal{B} = \begin{pmatrix} dB_1 \\ \cdot \\ \cdot \\ dB_n \\ dB_1^* \\ \cdot \\ \cdot \\ dB_n^* \end{pmatrix}, \tag{115}$$

and

$$d\mathcal{W} = \begin{pmatrix} dW_1 \\ \cdot \\ \cdot \\ dW_{2(n-1)} \\ d\widetilde{W}_1 \\ \cdot \\ \cdot \\ d\widetilde{W}_{2(n-1)} \end{pmatrix} . \tag{116}$$

The coefficient vector, respectively, matrix processes $\mu \in \mathbb{R}^{2n}$ and $\sigma \in \mathbb{R}^{2n\times 4(n-1)}$ are assumed to be $\mathcal{F}$-predictable and to satisfy Lipschitz and growth conditions necessary to ensure existence and uniqueness of solutions.

Already in this setting it is obvious that an arbitrage free market can be obtained by an appropriate choice for the drift, since for $n \geq 1$ the relation of number of tradable bonds to sources of uncertainty, represented by the number of one-dimensional Brownian motions, provides enough degrees of freedom to define a price deflator as numeraire.

Introducing stochastic volatility to this SDE demands an enlargement of the given framework. In their full generality, the dynamics of domestic and foreign zerobonds $B_i(t)$, respectively $B_i^*(t)$, will be driven by additional $2(n-1)$ CIR processes. Specifically, the above SDE will turn into

$$\frac{dB_i}{B_i} = \mu_i(t)dt + \sum_{j=1}^{4(n-1)} \sigma_{ij}(t, v_1, ..., v_{2(n-1)})d\mathcal{W}_j\,, \quad i = 1, ..., n$$

for the domestic part, and

$$\frac{dB_i^*}{B_i^*} = \mu_{n+i}(t)dt + \sum_{j=1}^{4(n-1)} \sigma_{n+i,j}(t, v_1, ..., v_{2(n-1)})d\mathcal{W}_j\,, \quad i = 1, ..., n$$

for the foreign zerobonds. For the square-root processes we have

$$dv_k = \kappa_k(\theta_k - v_k)dt + \overline{\sigma}_k\sqrt{v_k}\left(\rho_k d\widetilde{W}_k + \sqrt{1-\rho_k^2}d\overline{W}_k\right)\,, \quad 1 \leq k \leq 2(n-1).$$

The processes $\widetilde{W}$ and $\overline{W}$ are mutually independent $2(n-1)$-dimensional standard Brownian motions, both independent of W. In the CIR processes

v_k we overlined $\overline{\sigma}_k$ only in this paragraph to distinguish them from the zerobond volatilities. For the coefficient processes $\mu \in \mathbb{R}^{2n}$ and $\sigma \in \mathbb{R}^{2n\times 4(n-1)}$ to be adapted to the generated σ-field, the underlying filtration has to be augmented to include $\overline{W}$ as generator. A market extended in such a way is incomplete but still arbitrage free. The chosen martingale measures are not unique.

Following Schoenmakers (2005), the bond prices and the price deflator can be given in the following integrated form

$$B_i(t) = B_i(0)\exp\left[\int_0^t \left(\mu_i - \frac{1}{2}|\sigma_i|^2\right)ds + \int_0^t \sigma_i \cdot d\mathcal{W}(s)\right],$$

$$\xi(t) = \exp\left[\int_0^t \left(-r - \frac{1}{2}|\lambda|^2\right)ds - \int_0^t \lambda \cdot d\mathcal{W}(s)\right], \qquad (117)$$

as an application of Itô's Lemma demonstrates. The form of the deflator follows from the martingale representation theorem The first equation holds for $i = 1, ..., 2n$, once we identify

$$B_{n+j} := B_j^* , \quad \text{for} \quad j = 1, ..., n.$$

From (117) we see that

$$\xi B_i = B_i(0)\exp\left[\int_0^t \left(\mu_i - \frac{1}{2}|\sigma_i|^2 - r - \frac{1}{2}|\lambda|^2\right)ds + \int_0^t (\sigma_i - \lambda) \cdot d\mathcal{W}(s)\right],$$

which implies the equality

$$\mu_i - r - \frac{1}{2}|\sigma_i|^2 - \frac{1}{2}|\lambda|^2 + \frac{1}{2}|\sigma_i - \lambda|^2 = 0,$$

for all i, or equivalently

$$\mu_i = r + \sigma_i \cdot \lambda , \quad \text{for} \quad i = 1, ..., 2n, \qquad (118)$$

because ξB_i are martingales for all i. From this and linear algebra, we see that a solution for $\lambda \in \mathbb{R}^{4(n-1)}$ and $r \in \mathbb{R}$ can be found as long as $4(n-1) \geq 2n$. An arbitrage free market is thus ensured.

4.7 Dynamics of Stochastic Volatility Model

4.7.1 Under Terminal Measure

In this section we shall enlarge the stochastic volatility model of the previous chapter to the multicurrency setting. For this purpose we need to consider additional $2(n-1)$ Brownian motions from the correlated part of the CIR processes. These render $\mathcal{W}$ into a $4(n-1)$-dimensional process. As in Belomestny, Mathew, Schoenmakers (2007), it is straightforward to show that the domestic part of the resulting model has the following dynamics under the terminal measure P_n:

$$\frac{dL_i}{L_i} = -\sum_{j=i+1}^{n-1} \frac{\delta L_j}{1+\delta L_j} \left(\sum_{k=1}^{4(n-1)} \Gamma_{jk}\Gamma_{ik} \right) dt + \Gamma_i \cdot d\mathcal{W}^{(n)}\,, \tag{119}$$

where

$$\Gamma_i = \begin{pmatrix} \sqrt{1-r^2}\gamma_{i1} \\ \cdot \\ \cdot \\ \sqrt{1-r^2}\gamma_{i,2(n-1)} \\ r\overline{\gamma}_{i1}\sqrt{v_1} \\ \cdot \\ \cdot \\ r\overline{\gamma}_{i,2(n-1)}\sqrt{v_{2(n-1)}} \end{pmatrix} \qquad d\mathcal{W} = \begin{pmatrix} dW_1 \\ \cdot \\ \cdot \\ dW_{2(n-1)} \\ d\widetilde{W}_1 \\ \cdot \\ \cdot \\ d\widetilde{W}_{2(n-1)} \end{pmatrix},$$

for $1 \le i < n$.

In the same fashion we deduce for the foreign market, under P_{*n},

$$\frac{dL_i^*}{L_i^*} = -\sum_{j=i+1}^{n-1} \frac{\delta L_j^*}{1+\delta L_j^*} \left(\sum_{k=1}^{4(n-1)} \Gamma^*_{jk}\Gamma^*_{ik} \right) dt + \Gamma_i^* \cdot d\mathcal{W}^{(*n)}\,, \tag{120}$$

with a similar Γ^*:

$$\Gamma_i^* = \begin{pmatrix} \sqrt{1-r^2}\gamma^*_{i1} \\ \cdot \\ \cdot \\ \sqrt{1-r^2}\gamma^*_{i,2(n-1)} \\ r\overline{\gamma}^*_{i1}\sqrt{v_1} \\ \cdot \\ \cdot \\ r\overline{\gamma}^*_{i,2(n-1)}\sqrt{v_{2(n-1)}} \end{pmatrix} \qquad d\mathcal{W} = \begin{pmatrix} dW_1 \\ \cdot \\ \cdot \\ dW_{2(n-1)} \\ d\widetilde{W}_1 \\ \cdot \\ \cdot \\ d\widetilde{W}_{2(n-1)} \end{pmatrix},$$

again for $1 \leq i < n$. For simplicity we directly chose a constant r.

The $2(n-1)$ square-root processes v_k are also immediately given in the terminal measure P_n

$$dv_k = \kappa_k(1 - v_k)dt + \sigma_k \sqrt{v_k} \left(\rho_k d\widetilde{W}_k^{(n)} + \sqrt{1 - \rho_k^2} d\overline{W}_k^{(n)} \right), \tag{121}$$

for $k = 1, ..., 2(n-1)$. Their dynamics under P_{*n} are of no significance yet. In the previous chapter we have seen that we can choose $\theta_k \equiv 1$, without loss of generality.

Note that $\mathcal{W} \in \mathbb{R}^{4(n-1)}$. Under appropriate definition of σ_n^* and σ_n in the stochastic bond model, we again set

$$X_n = \frac{B_n^*}{B_n},$$

and obtain, just as in the non-stochastic volatility case,

$$\gamma_{X_n} = \sigma_n^* - \sigma_n.$$

This time, however, $\gamma_{X_n} \in \mathbb{R}^{4(n-1)}$. But analogously as in (108) we have

$$\gamma_{X_n} = \sigma_n^* - \sigma_n \approx \widetilde{\sigma}_\zeta + \sum_{j=\eta(t)}^{n-1} \left(\frac{\delta_j L_j}{1 + \delta_j L_j} \Gamma_j - \frac{\delta_j L_j^*}{1 + \delta_j L_j^*} \Gamma_j^* \right), \tag{122}$$

with $\widetilde{\sigma}_\zeta \in \mathbb{R}^{4(n-1)}$. It is natural to assume that the absolute spot rate volatility is independent of the number of factors:

$$|\widetilde{\sigma}_\zeta| = |\sigma_\zeta|.$$

As before we conclude

$$d\mathcal{W}^{(*n)} = d\mathcal{W}^{(n)} - \gamma_{X_n} dt.$$

The proof is the same as in part 2. of Proposition 58. It is based on Itô's Lemma which holds for general semimartingales.

We can now write

$$
\begin{aligned}
\frac{dL_i^*}{L_i^*} &= -\sum_{j=i+1}^{n-1} \frac{\delta L_j^*}{1+\delta L_j^*} \left(\sum_{k=1}^{4n-4} \Gamma_{jk}^* \Gamma_{ik}^* \right) dt + \Gamma_i^* \cdot d\mathcal{W}^{(*n)} \\
&= -\sum_{j=i+1}^{n-1} \frac{\delta L_j^*}{1+\delta L_j^*} \left(\sum_{k=1}^{4n-4} \Gamma_{jk}^* \Gamma_{ik}^* \right) dt + \Gamma_i^* \cdot \left(d\mathcal{W}^{(n)} - \gamma_{X_n} dt \right) \\
&= -\left(\sum_{j=i+1}^{n-1} \frac{\delta L_j^*}{1+\delta L_j^*} \left(\sum_{k=1}^{4n-4} \Gamma_{jk}^* \Gamma_{ik}^* \right) + \Gamma_i^* \cdot \gamma_{X_n} \right) dt + \Gamma_i^* \cdot d\mathcal{W}^{(n)} \\
&= -\left(\sum_{j=i+1}^{n-1} \frac{\delta L_j^*}{1+\delta L_j^*} \Gamma_i^* \cdot \Gamma_j^* + \Gamma_i^* \cdot \gamma_{X_n} \right) dt + \Gamma_i^* \cdot d\mathcal{W}^{(n)} \\
&= -\Gamma_i^* \cdot \left(\sum_{j=i+1}^{n-1} \frac{\delta L_j^*}{1+\delta L_j^*} \Gamma_j^* + \gamma_{X_n} \right) dt + \Gamma_i^* \cdot d\mathcal{W}^{(n)} \\
&= -\Gamma_i^* \cdot \left(\widetilde{\sigma}_\zeta + \sum_{j=\eta(t)}^{n-1} \frac{\delta_j L_j}{1+\delta_j L_j} \Gamma_j - \sum_{j=\eta(t)}^{i} \frac{\delta_j L_j^*}{1+\delta_j L_j^*} \Gamma_j^* \right) dt + \Gamma_i^* \cdot d\mathcal{W}^{(n)},
\end{aligned}
$$

where γ_{X_n} is given by (122). Except for the process $\widetilde{\sigma}_\zeta$ and the parameters in the square root processes, all quantities are given. The parameters will be calibrated in the following sections. As for the process $\widetilde{\sigma}_\zeta$ we will seek a deterministic correction as in (112). Observe that

$$\Gamma_i^* \cdot \widetilde{\sigma}_\zeta = r_f^* \, |\Gamma_i^*| \, |\widetilde{\sigma}_\zeta| \approx r_f^* \, |\gamma_i^*| \, |\sigma_\zeta| = r_f^* \, \overline{\sigma}_\zeta \, |\gamma_i^*| \, . \tag{123}$$

But also

$$\Gamma_i^* \cdot \widetilde{\sigma}_\zeta = r_f^* \, |\Gamma_i^*| \, |\widetilde{\sigma}_\zeta| \approx r_f^* \, |\gamma_i^*| \, |\sigma_\zeta| = \gamma_i^* \cdot \sigma_\zeta . \tag{124}$$

In the first equality we used the fact that the correlation r_f^* should naturally be unaffected by introduction of stochastic volatility components. The approximation stems from exchanging $|\Gamma_i^*|$ by its *approximate* expected value $|\gamma_i^*|$, see appendix 5.4.

We will use the deterministic RHS in (123) for the dynamics under P_n and the less restrictive, but still stochastic, RHS in (124) for the dynamics under forward measures P_{*i+1} in the next section. For now we need the known quantity $\overline{\sigma}_\zeta$. It follows immediately

$$\frac{dL_i^*}{L_i^*} = -r_f^* \, \overline{\sigma}_\zeta \, |\gamma_i^*| \, dt - \sum_{j=\eta(t)}^{n-1} \frac{\delta_j L_j}{1+\delta_j L_j} \Gamma_i^* \cdot \Gamma_j dt \tag{125}$$

$$+ \sum_{j=\eta(t)}^{i} \frac{\delta_j L_j^*}{1+\delta_j L_j^*} \Gamma_i^* \cdot \Gamma_j^* \, dt + \Gamma_i^* \cdot d\mathcal{W}^{(n)}.$$

With (119) and (125) we have fully specified dynamics to price multicurrency products by simulation, once the CIR-process parameters are calibrated.

4.7.2 Under Forward Measures

Recall that from the parameterization

$$\gamma_i = \eta_i \mathbf{e}_{n-1+i}$$

and the setting

$$\overline{\gamma}_i = \sigma_i^{BS} \mathbf{e}_{n-1+i},$$

for $i = 1, ..., n-1$, see Belomestny, Mathew, Schoenmakers (2007), the first and third $n-1$ entries of Γ_i, for $i = 1, ..., n-1$, are zero. This implies that the extended stochastic volatility model preserves the designated property, that when constricted to the home market, it reduces to the one-currency stochastic volatility $(n-1)$-full factor model introduced in Belomestny, Mathew, Schoenmakers (2007). Therefore, the calibration of the parameters of the second $n-1$ of the v_k's, that is for $k = n, ..., 2n-2$, is performed as before in the single currency case. Moreover, already existing estimates can be directly employed. We can thus neglect the specification of (119) in terms of forward measures P_{i+1}. In order to determine the parameters of the first $n-1$ v_k's, that is for $k = 1, ..., n-1$, however, we need the dynamics of the foreign LIBORS L^* under the respective measures P_{*i+1}.

For a simpler notation define the process $\alpha \in \mathbb{R}^{4(n-1)}$ by

$$\alpha := \sum_{j=\eta(t)}^{n-1} \frac{\delta_j L_j}{1+\delta_j L_j} \Gamma_j - \sum_{j=\eta(t)}^{i} \frac{\delta_j L_j^*}{1+\delta_j L_j^*} \Gamma_j^*.$$

Rewriting (125) in full length with the definition of Γ^* yields

$$
\begin{aligned}
\frac{dL_i^*}{L_i^*} &= -\Gamma_i^* \cdot \left(\sum_{j=i+1}^{n-1} \frac{\delta L_j^*}{1+\delta L_j^*} \Gamma_j^* + \gamma_{X_n} \right) dt + \Gamma_i^* \cdot d\mathcal{W}^{(n)} \\
&\overset{(122)}{=} -\Gamma_i^* \cdot \left(\widetilde{\sigma}_\zeta + \sum_{j=\eta(t)}^{n-1} \frac{\delta_j L_j}{1+\delta_j L_j} \Gamma_j - \sum_{j=\eta(t)}^{i} \frac{\delta_j L_j^*}{1+\delta_j L_j^*} \Gamma_j^* \right) dt + \Gamma_i^* \cdot d\mathcal{W}^{(n)} \\
&\overset{(124)}{=} \sqrt{1-r^2} \sum_{k=i}^{2(n-1)} \gamma_{ik}^* \left(dW_k^{(n)} - \left(\alpha_k + \frac{\sigma_{\zeta_k}}{\sqrt{1-r^2}} \right) dt \right) \\
&\quad + r \sum_{k=1}^{2(n-1)} \overline{\gamma}_{ik}^* \sqrt{v_k} \left(d\widetilde{W}_k^{(n)} - \alpha_{2n-2+k}\, dt \right) \\
&=: \sqrt{1-r^2} \sum_{k=i}^{2(n-1)} \gamma_{ik}^* dW_k^{(*i+1)} + r \sum_{k=i}^{2(n-1)} \overline{\gamma}_{ik}^* \sqrt{v_k} d\widetilde{W}_k^{(*i+1)}. \qquad (126)
\end{aligned}
$$

Since L_i^* is a martingale under P_{*i+1}, we have that both $W^{(*i+1)}$ and $\widetilde{W}^{(*i+1)}$ in (126) are standard Brownian motions under P_{*i+1}. In terms of the new Brownian motion

$$
d\widetilde{W}_k^{(n)} = d\widetilde{W}_k^{(*i+1)} + r\sqrt{v_k} \left(\sum_{j=\eta(t)}^{n-1} \frac{\delta_j L_j}{1+\delta_j L_j} \overline{\gamma}_{jk} - \sum_{j=\eta(t)}^{i} \frac{\delta_j L_j^*}{1+\delta_j L_j^*} \overline{\gamma}_{jk}^* \right) dt
$$

the volatility dynamics are

$$
\begin{aligned}
dv_k &= \kappa_k(1-v_k)dt + r\sigma_k\rho_k v_k \left(\sum_{j=\eta(t)}^{n-1} \frac{\delta_j L_j}{1+\delta_j L_j} \overline{\gamma}_{jk} - \sum_{j=\eta(t)}^{i} \frac{\delta_j L_j^*}{1+\delta_j L_j^*} \overline{\gamma}_{jk}^* \right) dt \\
&\quad + \rho_k \sigma_k \sqrt{v_k} d\widetilde{W}_k^{(*i+1)} + \sqrt{1-\rho_k^2} \sigma_k \sqrt{v_k} d\overline{W}_k^{(n,*i+1)}. \qquad (127)
\end{aligned}
$$

As shown in the lemma below, the process $\overline{W}^{(n,*i+1)}$ in (127) is standard Brownian motion under both measures P_{*i+1} and P_n.

By freezing the LIBORS at their initial values in (127) and observing that $\eta(0) = 1$, we obtain an approximative CIR dynamics

$$dv_k \approx \kappa_k^{(*i+1)} \left(\theta_k^{(*i+1)} - v_k\right) dt + \sigma_k \sqrt{v_k} \left(\rho_k d\widetilde{W}_k^{(*i+1)} + \sqrt{1-\rho_k^2} d\overline{W}_k^{(*i+1)}\right) \tag{128}$$

with reversion speed parameter

$$\kappa_k^{(*i+1)} := \kappa_k - r\sigma_k\rho_k \left(\sum_{j=1}^{n-1} \frac{\delta_j L_j(0)}{1+\delta_j L_j(0)} \overline{\gamma}_{jk} - \sum_{j=1}^{i} \frac{\delta_j L_j^*(0)}{1+\delta_j L_j^*(0)} \overline{\gamma}_{jk}^*\right), \tag{129}$$

and mean reversion level

$$\theta_k^{(*i+1)} := \frac{\kappa_k}{\kappa_k^{(*i+1)}}. \tag{130}$$

The approximative dynamics (128) for the volatility process will be used in the next section. The following formulas and PDEs for pricing and calibration in a multicurrency setting will resemble in most instances those of the single-currency case, except for different parameters involving *'s. Note that, however, through (129) and (130) they contain all necessary information about γ_{X_n}, and thus about the foreign market.

Measure Invariance In the multicurrency case a modified version of Lemma 46 has to be shown. Its proof is similar in most respects.

Lemma 59 *For $k = 1, ..., n-1$, we have*

$$d\overline{W}_k^{(n)} = d\overline{W}_k^{(*i+1)}.$$

In other words, $d\overline{W}_k^{(n,*i+1)}$ is invariant under the various measures, namely P_{*i+1} and P_n.

Proof. See Jamshidian (1997) for the difference between compensators. He shows that it is given by

$$\mu_{\overline{W}_k^{(n)}}^{*i+1} = \langle \overline{W}_k^{(n)}, \ln M \rangle,$$

where now M is

$$M = X \prod_{j=i+1}^{n-1} (1 + \delta L_j^*).$$

From Section 1.1.3 we have

$$M_t = E_n\left(\frac{dP_{*i+1}}{dP_n}|\mathcal{F}_t\right) = C\,\frac{B^*_{i+1}}{B_n} = C\,X\,\frac{B^*_{i+1}}{B^*_n},$$

for some positive constant $C > 0$. From this we obtain

$$\begin{aligned}
\langle \overline{W}^{(n)}_k, \ln M\rangle &= d\overline{W}^{(n)}_k d\ln M \\
&= d\overline{W}^{(n)}_k d\left(\ln X + \sum_{j=i+1}^{n-1} \ln\left(1+\delta L^*_j\right)\right) \\
&= d\ln X d\overline{W}^{(n)}_k + \sum_{j=i+1}^{n-1} d\overline{W}^{(n)}_k d\ln(1+\delta L^*_j) \\
&= \left(-\frac{1}{2}|\gamma_X|^2 dt + \gamma_X dW^{(n)}\right) d\overline{W}^{(n)}_k + \sum_{j=i+1}^{n-1} \frac{\delta L^*_j}{1+\delta L^*_j} d\overline{W}^{(n)}_k d\ln L^*_j
\end{aligned}$$

A closer look at (125) or (126) reveils that all terms are negligible, since either of higher order than dt, or zero due to independence of $\overline{W}$ and W or $\widetilde{W}$, respectively. We thus have

$$\langle \overline{W}^{(n)}_k, \ln M\rangle = 0$$

or in other words, as indicated by $d\overline{W}^{(n,*i+1)}_k$:

$$d\overline{W}^{(n)}_k = d\overline{W}^{(*i+1)}_k.$$

■

4.8 Pricing and Calibration

Due to our parametrization, $\overline{\gamma}$ can be immediately taken from in the inner domestic calibration. The first $n-1$-steps will be a copy of the single-currency case, as explained in BMS (2008). We therefore start our procedure with the foreign caplet pricing, which involves *new* φ_{*j+1} depending on $\overline{\gamma}^*$. Let us take the opportunity to recap the caplet pricing for the foreign market.

4.8.1 Pricing Caplets

A foreign caplet for the period $[T_j, T_{j+1}]$ with strike K is an option that pays $(L_j^*(T_j) - K)^+ \delta_j$ units of *foreign* currency at time T_{j+1}, where $1 \leq j < n$. It is well-known that under the forward measure P_{*j+1} the j-th foreign caplet price in *domestic* currency at time zero is given by

$$C_j^*(K) = \delta_j B_{j+1}^*(0) E_{*j+1}(L_j^*(T_j) - K)^+.$$

Consequently under P_{*j+1} the j-th foreign caplet price is determined by the dynamics of L_j^* only. The FFT-method of Carr and Madan (1999) can be straightforwardly adapted to the caplet pricing problem as done in Belomestny and Schoenmakers (2006). In terms of the log-moneyness variable

$$v := \ln \frac{K}{L_j^*(0)} \tag{131}$$

the j-th caplet price can be expressed as

$$\mathcal{C}_j^*(v) := C_j^*(e^v L_j^*(0)) = \delta_j B_{j+1}^*(0) L_j^*(0) E_{*j+1} \left(e^{X_j(T_j)} - e^v \right)^+,$$

where $X_j(t) = \ln L_j^*(t) - \ln L_j^*(0)$. One then defines the auxiliary function

$$\mathcal{O}_j^*(v) := \frac{\mathcal{C}_j^*(v)}{\delta_j B_{j+1}^*(0) L_j^*(0)} - (1 - e^v)^+ \tag{132}$$

and can show the following proposition.

Proposition 60 *For the Fourier transform of the function $\mathcal{O}_j^*$ defined above and $\varphi_{*j+1}(\cdot; t)$ denoting the characteristic function of the process $X_j(t)$ under P_{*j+1} we have*

$$\mathcal{F}\left\{\mathcal{O}_j^*\right\}(z) = \int_{-\infty}^{\infty} \mathcal{O}_j^*(v) e^{\mathrm{i}vz} dv = \frac{1 - \varphi_{*j+1}(z - \mathrm{i}; T_j)}{z(z - \mathrm{i})}. \tag{133}$$

The proof can be found in Belomestny and Reiß (2006). Next, combining (131), (132), and (133) yields

$$C_j^*(K) = \delta B_{j+1}^*(0)\left(L_j^*(0) - K\right)^+ \tag{134}$$
$$+ \frac{\delta B_{j+1}^*(0) L_j^*(0)}{2\pi} \int_{-\infty}^{\infty} \frac{1 - \varphi_{*j+1}(z - \mathfrak{i}; T_j)}{z(z - \mathfrak{i})} e^{-\mathfrak{i}z \ln \frac{K}{L_j^*(0)}} dz.$$

4.8.2 Calibration Road Map

We now outline a calibration procedure for the LIBOR structure (119), (120) and (121) under the following additional assumptions.

(i) The input market LIBOR volatility structures of both currencies, $\gamma \in \mathbb{R}^{(n-1)\times(n-1)}$ and $\gamma^* \in \mathbb{R}^{(n-1)\times(n-1)}$, are assumed to be of full rank.

(ii) The terminal domestic log-LIBOR increment $d \ln L_{n-1}$ is influenced by a single stochastic volatility shock $\sqrt{v_{2n-2}} d\widetilde{W}_{2n-2}$, the one but last, hence $d \ln L_{n-2}$, by only $\sqrt{v_{2n-2}} d\widetilde{W}_{2n-2}$ and $\sqrt{v_{2n-1}} d\widetilde{W}_{2n-1}$, and backwards so forth through the foreign LIBORS. Put differently, we assume $\varrho \in \mathbb{R}^{(2n-2)\times(2n-2)}$ to be Cholesky-decomposable into a product of a squared upper triangular matrix and a lower triangular matrix, both of rank $2n-2$.

(iii) We obtain time-independent upper triangular solutions $\overline{\gamma}^*$ and $\overline{\gamma}$ through setting, for $i, j = 1, ..., n-1$,

$$\begin{aligned} \overline{\gamma}_i^* &= \sigma_i^{*\text{Black}} \mathbf{e}_i, \quad \text{where} \\ \left(\sigma_i^{*\text{Black}}\right)^2 &:= \frac{1}{T_i} \int_0^{T_i} |\gamma_i^*(s)|^2 \, ds, \\ \mathbf{e}_i^\top \mathbf{e}_j &:= \frac{\gamma_i^{*\top} \gamma_j^*}{|\gamma_i^*||\gamma_j^*|}(0) \end{aligned}$$

and

$$\overline{\gamma}_i = \sigma_i^{\text{Black}} \mathbf{e}_{n-1+i}, \quad \text{where} \tag{135}$$
$$\left(\sigma_i^{\text{Black}}\right)^2 := \frac{1}{T_i} \int_0^{T_i} |\gamma_i(s)|^2 \, ds,$$
$$\mathbf{e}_{n-1+i}^\top \mathbf{e}_{n-1+j} := \frac{\gamma_i^\top \gamma_j}{|\gamma_i||\gamma_j|}(0), \tag{136}$$

where $\mathbf{e}_i \in \mathbb{R}^{2n-2}$, for $i, j = 1, ..., 2n-2$.

(iv) Recall that $v_k(0) \equiv \theta_k \equiv 1$, $1 \leq k \leq 2(n-1)$.

The backward algorithm works for the last $n-1$ volatility components, i.e. for $k = n, ..., 2(n-1)$, just as in the single currency case, see BMS (2008). At the passage to, and then through, the foreign LIBORS the procedure is given now. For the foreign LIBOR dynamics structured in the above way we obtain after an application of Itô's Lemma

$$\begin{aligned} d \ln L_i^*(t) = &-\frac{1}{2}\left[(1-r^2)\left|\gamma_i^*\right|^2 + r^2 \sum_{k=i}^{2(n-1)} \left(\overline{\gamma}_{ik}^*\right)^2 v_k\right] dt \\ &+ \sqrt{1-r^2}\gamma_i^* \cdot dW^{(*i+1)} \\ &+ r \sum_{k=i}^{2(n-1)} \overline{\gamma}_{ik}^* \sqrt{v_k} d\widetilde{W}_k^{(*i+1)}, \quad 1 \leq i < n. \end{aligned} \tag{137}$$

For $i = n-1$ the dynamics of v_{n-1} is given by (121), whereas for $i < n-1$ the dynamics of v_k, $i \leq k < n$, is approximately given by (128).

We will calibrate the structure to prices of caplets according to the following roadmap.

1. First step $i = 2(n-1)$ back through all domestic LIBORS to step $i = n$ proceed as in single-currency case. Calibrate the parameter set $(\kappa_i, \sigma_i, \rho_i, \theta_i \equiv 1)$, for $i = n, ..., 2n-2$. Optionally calibrate r with it, or just fix r.

2. Step $i = n-1$. Calibrate the parameter set $(\kappa_{n-1}^{(*n)}, \theta_{n-1}^{(*n)}, \sigma_{n-1}, \rho_{n-1})$ to the T_{n-1} column of the foreign cap-strike matrix via (134) using the explicitly known characteristic function φ_{*n} of $\ln[L_{n-1}^*(T_{n-1})/L_{n-1}^*(0)]$ (see Appendix (5.2)).

3. For $i = n-2$ down to 1 carry out the next iteration step:

4. Step $i = n-k$. Transform the yet known parameter set $(\kappa_j, \theta_j, \sigma_j, \rho_j)$, $i < j \leq 2(n-1)$, via (129) and (130) into the corresponding set $(\kappa_j^{(*i+1)}, \theta_j^{(*i+1)}, \sigma_j, \rho_j)$, $i < j \leq 2(n-1)$. Then calibrate the at this stage unknown parameter set $(\kappa_i^{(*i+1)}, \theta_i^{(*i+1)}, \sigma_i, \rho_i)$ to the T_i column

of the foreign cap-strike matrix via (134) using the explicitly known characteristic function φ_{*i+1} of $\ln[L_i^*(T_i)/L_i^*(0)]$ under the approximation (126)-(128) (see Appendix (5.2)).

Unfortunately the drift correction in (129) does not disappear when $k \leq i$ in $\kappa_k^{(*i+1)}$, as it is the case for single-currency models, see BMS (2008). Therefore, parameters $\kappa_k^{(*i+1)}$ have to be retransformed by (129) to κ_k for simulation in P_n. The parameter $\theta_k^{(*i+1)}$ transform naturally back to $\theta_k \equiv 1$.

The above calibration algorithm includes at each step, as usual, the minimization of some objective function. As such function we take the weighted sum of squares of the corresponding differences between observed market prices and prices induced by the model. The weights are taken to be proportional to Black-Scholes vegas. As initial values for the local optimization routine at time step $i+1$ the values of estimated parameters at time step i are used.

4.8.3 Application to Differential Swap

As we mentioned in the introduction, in our volatility model the random variables under consideration are in general not lognormaly distributed. Although essential coefficients will be assumed to be of deterministic form, either directly, as in the case of γ and γ^*, or by approximation, like γ_{X_n}, closed formulae were not our goal. In the end, prices of *exotic* derivatives are determined by a Monte Carlo Simulation. Of interest is, however, that a model can be calibrated in an efficient and stable way, a problem often neglected or not even addressed in former works.

Here is an example of such an exotic product whose price may be approximated once all parameters are specified. The value of a (domestic) floating to (foreign) floating *Differential Swap* can be expressed by the following expectation under the terminal measure

$$V_{DS}(t) = N \cdot \delta \sum_{i=1}^{n-1} B_{i+1}(t) E_n \left(L_i^*(T_i) - L_i(T_i) | \mathcal{F}_t \right),$$

where N denotes the notional amount and δ the day count fraction.

The price of an *option* on a Differential Swap with expiration time $T < T_i$, for all $i = 1, ..., n$, is given by

$$B(0,T) E_T \left[\max \left(0, N \cdot \delta \sum_{i=1}^{n-1} B_{i+1}(T) E_n \left(L_i^*(T_i) - L_i(T_i) | \mathcal{F}_T \right) \right) \right].$$

5 Appendix

5.1 The Conditional Characteristic Function for L_j

For $j = 1, ..., n-1$, we need to determine the characteristic function of $\ln L_j(T) - \ln L_j(0)$ under the relevant measure P_{j+1}. For each component $k = 1, ..., n-1$ the Heston CIR-process has the general form

$$dv_k = \kappa_k^{(j+1)} \left(\theta_k^{(j+1)} - v_k \right) dt + \sigma_k \rho_k \sqrt{v_k} d\widetilde{W}_k^{(j+1)} + \sigma_k \sqrt{(1-\rho_k^2)} \sqrt{v_k} d\overline{W}_k^{(j+1)}.$$

In this case and a forward LIBOR dynamic given by (78) , with general $v \in \mathbb{R}^{n-1}$, the solution is of the form

$$\begin{aligned} \varphi_{j+1}(z\,; T, v) &= E_{j+1}\left[\exp\left(\mathrm{i} z \ln\left(\frac{L_j(T)}{L_j(0)} \right) \right) \middle| v_k(0) = v_k\,,\ k = 1, ..., n-1 \right] \\ &= \varphi_{j+1,0}\left(z\,; T\right) \prod_{k=j}^{n-1} \varphi_{j+1,k}\left(z\,; T, v_k\right), \end{aligned} \tag{138}$$

where

$$\varphi_{j+1,0}(z\,; T) = \exp\left(-\frac{1}{2}(1-r^2)\eta_j^2(T)\left(z^2 + \mathrm{i}\, z\right) \right), \quad \eta_j^2(T) = \int_0^T |\gamma_j|^2\, dt.$$

Each $\varphi_{j+1,k}\left(z\,; T, v_k\right)$ is of the form

$$\varphi_{j+1,k}\left(z\,; T, v_k\right) = p_{j+1,k}\left(z\,; T, y_k, v_k\right)|_{y_k=0},$$

with $p_{j+1,k}$ satisfying the parabolic equations

$$\begin{aligned} \frac{\partial p_{j+1,k}}{\partial T} &= \kappa_k^{(j+1)}\left(\theta_k^{(j+1)} - v_k\right)\frac{\partial p_{j+1,k}}{\partial v_k} - \frac{1}{2} r^2 \overline{\gamma}_{j,k}^2 v_k \frac{\partial p_{j+1,k}}{\partial y_k} + \frac{1}{2}\sigma_k^2 v_k \frac{\partial^2 p_{j+1,k}}{\partial v_k^2} \\ &\quad + \frac{1}{2} r^2 \overline{\gamma}_{j,k}^2 v_k \frac{\partial^2 p_{j+1,k}}{\partial y_k^2} + \sigma_k \rho_k r \overline{\gamma}_{j,k} v_k \frac{\partial^2 p_{j+1,k}}{\partial v_k \partial y_k}, \end{aligned}$$

with boundary condition

$$p_{j+1,k}(z\,; 0, y_k, v_k) = e^{\mathrm{i} z y_k}.$$

This can be verified with the Feynman-Kac formula.

Since $\overline{\gamma}_j$ are constant, the above equation can be solved explicitly. The ansatz

$$\varphi_{j+1,k}(z;T,v_k) = \exp\left(A_{j,k}(z;T) + v_k B_{j,k}(z;T)\right)$$

yields

$$A_{j,k}(z;T) = \frac{\kappa_k^{(j+1)}\theta_k^{(j+1)}}{\sigma_k^2}\left[(a_{j,k}+d_{j,k})T - 2\ln\left(\frac{1-g_{j,k}e^{d_{j,k}T}}{1-g_{j,k}}\right)\right],$$

$$B_{j,k}(z;T) = \frac{(a_{j,k}+d_{j,k})(1-e^{d_{j,k}T})}{\sigma_k^2(1-g_{j,k}e^{d_{j,k}T})},$$

where

$$a_{j,k} = \kappa_k^{(j+1)} - \mathfrak{i}\, r\rho_k\sigma_k\overline{\gamma}_{jk}z,$$

$$d_{j,k} = \sqrt{a_{j,k}^2 + r^2\overline{\gamma}_{jk}^2\sigma_k^2(z^2+\mathfrak{i}\, z)},$$

$$g_{j,k} = \frac{a_{j,k}+d_{j,k}}{a_{j,k}-d_{j,k}}.$$

Note that the first lower index $j+1$ at the characteristic function refers to the measure, whereas the first index j at the introduced coefficients refers to relevant forward LIBOR. The second index refers to the component.
It is again the choice of $\overline{\gamma}$ that enables the product in (138) to be started at j. This essential feature will be beneficial in the calibration part. When $j = n-1$, for example, only the last log-LIBOR will contribute a nontrivial factor to the characteristic function. For all others we have

$$\varphi_{n,k} \equiv 1\,, \quad k = 1,...,n-2\,.$$

5.2 The Conditional Characteristic Function for L_j^*

For $j = 1,...,n-1$, we need to determine the conditional characteristic function of $\ln L_j^*(T) - \ln L_j^*(0)$ under the relevant measure P_{*j+1}. For each component $k = 1,...,2(n-1)$ the Heston CIR-process has the general form

$$dv_k = \kappa_k^{(*j+1)}\left(\theta_k^{(*j+1)} - v_k\right)dt + \sigma_k\rho_k\sqrt{v_k}d\widetilde{W}_k^{(*j+1)} + \sigma_k\sqrt{(1-\rho_k^2)}\sqrt{v_k}d\overline{W}_k^{(*j+1)}.$$

In this case and a forward LIBOR dynamic given by (137), with general $v \in \mathbb{R}^{2(n-1)}$, the solution is of the form

$$\begin{aligned}\varphi_{*j+1}(z\,;T,v) &= E_{*j+1}\left[\exp\left(\mathrm{i}z\ln\left(\frac{L_j^*(T)}{L_j^*(0)}\right)\right)\middle| v_k(0)=v_k, 1\leq k\leq 2n-2\right] \\ &= \varphi_{*j+1,0}\left(z\,;T\right)\prod_{k=j}^{2(n-1)}\varphi_{*j+1,k}\left(z\,;T,v_k\right), \qquad (139)\end{aligned}$$

where

$$\varphi_{*j+1,0}(z\,;T) = \exp\left(-\frac{1}{2}(1-r^2)\eta_j^2(T)\left(z^2+\mathrm{i}\,z\right)\right), \quad \eta_j^2(T) = \int_0^T \left|\gamma_j^*\right|^2 dt.$$

Each $\varphi_{*j+1,k}\left(z\,;T,v_k\right)$ is of the form

$$\varphi_{*j+1,k}\left(z\,;T,v_k\right) = p_{*j+1,k}\left(z\,;T,y_k,v_k\right)|_{y_k=0},$$

with $p_{*j+1,k}$ satisfying the parabolic equations

$$\begin{aligned}\frac{\partial p_{*j+1,k}}{\partial T} &= \kappa_k^{(*j+1)}\left(\theta_k^{(*j+1)}-v_k\right)\frac{\partial p_{*j+1,k}}{\partial v_k} - \frac{1}{2}r^2\left(\overline{\gamma}_{j,k}^*\right)^2 v_k\frac{\partial p_{*j+1,k}}{\partial y_k} \\ &+\frac{1}{2}\sigma_k^2 v_k\frac{\partial^2 p_{*j+1,k}}{\partial v_k^2} + \frac{1}{2}r^2\left(\overline{\gamma}_{j,k}^*\right)^2 v_k\frac{\partial^2 p_{*j+1,k}}{\partial y_k^2} + \sigma_k\rho_k r\overline{\gamma}_{j,k}v_k\frac{\partial^2 p_{*j+1,k}}{\partial v_k\partial y_k},\end{aligned}$$

with boundary condition

$$p_{*j+1,k}(z\,;0,y_k,v_k) = e^{\mathrm{i}\,zy_k}.$$

Again this can be verified with the Feynman-Kac formula.

Since $\overline{\gamma}_j^*$ are constant, the above equation can be solved explicitly. The ansatz

$$\varphi_{*j+1,k}(z;T,l,v_k) = \exp\left(A_{j,k}(z;T)+v_kB_{j,k}(z;T)\right)$$

yields

$$A_{j,k}(z;T) = \frac{\kappa_k^{(*j+1)}\theta_k^{(*j+1)}}{\sigma_k^2}\left[(a_{j,k}+d_{j,k})T - 2\ln\left(\frac{1-g_{j,k}e^{d_{j,k}T}}{1-g_{j,k}}\right)\right],$$

$$B_{j,k}(z;T) = \frac{(a_{j,k}+d_{j,k})(1-e^{d_{j,k}T})}{\sigma_k^2(1-g_{j,k}e^{d_{j,k}T})},$$

where

$$a_{j,k} = \kappa_k^{(*j+1)} - \mathrm{i}\, r\rho_k\sigma_k\overline{\gamma}^*_{jk}z,$$

$$d_{j,k} = \sqrt{a^2_{j,k} + r^2\left(\overline{\gamma}^*_{jk}\right)^2\sigma^2_k(z^2+\mathrm{i}\,z)},$$

$$g_{j,k} = \frac{a_{j,k}+d_{j,k}}{a_{j,k}-d_{j,k}}.$$

Note that the first lower index $j+1$ at the characteristic function refers to the measure, whereas the first index j at the introduced coefficients refers to relevant forward LIBOR. The second index refers to the component. It is again the choice of $\overline{\gamma}^*$ that enables the product in (139) to be started at j.

5.3 Extension of Correlation Matrices

Lemma 61 *Let $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times n}$ be positive definite correlation matrices with Eigenvalues α_i and β_i , for $i = 1,..,n$, respectively. Denote by $M \in \mathbb{R}^{n\times n}$ the matrix of rank one that contains only ones as entries. We can conclude the following:*
If $\rho < \frac{1}{n}\cdot\min\{\alpha_i\,,\,\beta_i\,,\ for\ 1,..,n\}$, then

$$\boldsymbol{C} = \begin{pmatrix} \boldsymbol{A} & | & \rho\boldsymbol{M} \\ ---& & --- \\ \rho\boldsymbol{M} & | & \boldsymbol{B}\end{pmatrix}.$$

is a positive definite correlation matrix.

Proof. Symmetric and positive definite matrices A and B have spectral decompositions $\Phi = \Gamma_A A\,\Gamma'_A$ and $\Psi = \Gamma_B B\,\Gamma'_B$, where Φ and Ψ are diagonal matrices containing the Eigenvalues α_i and β_i , for $i = 1,..,n$. Since similar matrices have the same eigenvalues, we can pre- and post-multiply C with

$$\Gamma = \begin{pmatrix} \Gamma_A & | & \mathbf{0} \\ --- & & --- \\ \mathbf{0} & | & \Gamma_B\end{pmatrix} \quad \text{and} \quad \Gamma' = \begin{pmatrix} \Gamma'_A & | & \mathbf{0} \\ --- & & --- \\ \mathbf{0} & | & \Gamma'_B\end{pmatrix}.$$

respectively, without changing the eigenvalues. Instead of C we can thus analyze

$$\Gamma C\Gamma' = \begin{pmatrix} \Gamma_A A\,\Gamma'_A & | & \rho\,\Gamma_A M\,\Gamma'_B \\ --- & & --- \\ \rho\,\Gamma_B M\,\Gamma'_A & | & \Gamma_B B\,\Gamma'_B\end{pmatrix} = \begin{pmatrix} \Phi & | & \rho\,\Gamma_A M\,\Gamma'_B \\ --- & & --- \\ \rho\,\Gamma_B M\,\Gamma'_A & | & \Psi\end{pmatrix}.$$

Set

$$D := \begin{pmatrix} \Phi & | & \mathbf{0} \\ --- & & --- \\ \mathbf{0} & | & \Psi \end{pmatrix}.$$

and note that D and $\Gamma C\,\Gamma'$ are symmetric. By a minimization theorem (see for example: Mardia, Kent and Bibby, page 479 Theorem A.9.2) we know that

$$\min x'\Gamma C\,\Gamma' x = \lambda_{\min}$$

subject to the constraint $x'Dx = 1$, where $\lambda_{\min}$ denotes the smallest eigenvalue of $D^{-1}\Gamma C\,\Gamma'$. Thus $\lambda_{\min} > 0$ implies positive definiteness of $\Gamma C \Gamma'$ and therefore the one of C.

The Eigenvalues λ of $D^{-1}\Gamma C \Gamma'$ satisfy

$$\det\left(D^{-1}\Gamma C\Gamma' - \lambda\mathbf{1}\right) = \det \begin{pmatrix} (1-\lambda)\mathbf{1} & | & \rho\,\Phi^{-1}\Gamma_A M\,\Gamma_B' \\ --- & & --- \\ \rho\,\Psi^{-1}\Gamma_B M\,\Gamma_A' & | & (1-\lambda)\mathbf{1} \end{pmatrix} = 0.$$

If $\lambda \neq 1$, a trivial case, this is equivalent to

$$\det\left((1-\lambda)\mathbf{1} - \frac{\rho^2}{1-\lambda}\Psi^{-1}\Gamma_B M\,\Gamma_A'\Phi^{-1}\Gamma_A M\,\Gamma_B'\right) = 0. \tag{140}$$

In general we have for the product $MQ' = QM$, with Q some orthogonal matrix. Furthermore, MQ' is symmetric, because $(MQ')' = QM' = QM = MQ'$. Thus any diagonal matrix commutes with MQ'. Setting $Q = \Gamma_A\Gamma_B$, these last facts render (140) equivalent to

$$\begin{aligned} 0 = & \det\left(\mathbf{1} - \frac{\rho^2}{(1-\lambda)^2}\Psi^{-1}\Phi^{-1}MQ'QM\right) \\ = & \det\left(\mathbf{1} - \frac{n\rho^2}{(1-\lambda)^2}\Psi^{-1}\Phi^{-1}M\right) \\ = & \det\left(\mathbf{1} - \frac{n\rho^2}{(1-\lambda)^2}\begin{pmatrix} \frac{1}{\alpha_1\beta_1} \\ \cdot \\ \cdot \\ \cdot \\ \frac{1}{\alpha_n\beta_n} \end{pmatrix}(1,...,1)\right). \end{aligned}$$

By the Matrix Determinant Lemma,

$$\det(\mathbf{X} + \mathbf{u}\mathbf{v}^{\mathrm{T}}) = (1 + \mathbf{v}^{\mathrm{T}}\mathbf{X}^{-1}\mathbf{u})\,\det(\mathbf{X}),$$

we have now

$$
\begin{aligned}
(140) \Leftrightarrow \qquad & 1 - \frac{n\rho^2}{(1-\lambda)^2}(1, ..., 1) \begin{pmatrix} \frac{1}{\alpha_1\beta_1} \\ \cdot \\ \cdot \\ \cdot \\ \frac{1}{\alpha_n\beta_n} \end{pmatrix} = 0 \\
\Leftrightarrow \qquad & 1 = \frac{n\rho^2}{(1-\lambda)^2} \sum_{i=1}^{n} \frac{1}{\alpha_i\beta_i} \\
\Leftrightarrow \qquad & (1-\lambda)^2 = n \sum_{i=1}^{n} \frac{|\rho||\rho|}{\alpha_i\beta_i}
\end{aligned}
$$

By assumption the right-hand-side is smaller one, thus for all Eigenvalues we must have $\lambda > 0$. ■

Remark 62 *A simple example shows that the bound on ρ in the lemma cannot be sharpened. Take $n = 1$ and $A = B = (1)$, then the Eigenvalues are both unity, but the choice $\rho = 1$ turns*

$$
\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.
$$

into a singular matrix. Of course this example does not show the reverse direction, it only demonstrates that something can go wrong if ρ is picked out of bound

Remark 63 *In the beginning of the proof we used the fact that similar matrices have the same eigenvalues. Premultiplying C with*

$$
\begin{pmatrix} \boldsymbol{A}^{-1} & | & \boldsymbol{0} \\ --- & & --- \\ \boldsymbol{0} & | & \boldsymbol{B}^{-1} \end{pmatrix}.
$$

directly and utilizing the Mardia, Kent and Bibby argument, would reach the conclusion earlier.

5.4 An Approximation

The following lemma provides justification for the approximations made in (123) and (124).

Lemma 64 *For $1 \leq i < n$, at $t = T_i$ we have*

$$E\left(\int_0^t |\Gamma_i^*|^2 \, ds\right) = \int_0^t |\gamma_i^*|^2 \, ds.$$

Proof. We need to show that at $t = T_i$

$$E\left((1-r^2)\sum_{j=1}^{2(n-1)}\int_0^t \left(\gamma_{ij}^*\right)^2 ds + r^2 \sum_{j=1}^{2(n-1)}\int_0^t \left(\overline{\gamma}_{ij}^*\right)^2 v_j \, ds\right) = \sum_{j=1}^{2(n-1)}\int_0^t \left(\gamma_{ij}^*\right)^2 ds.$$

Since by construction $Ev_j = 1$, the equation reduces to

$$(1-r^2)\int_0^t |\gamma_i^*|^2 \, ds + r^2 \int_0^t |\overline{\gamma}_i^*|^2 \, ds = \int_0^t |\gamma_i^*|^2 \, ds.$$

But from the definition of $\overline{\gamma}_i^*$ this holds indeed at $t = T_i$. ■

5.5 CIR

Consider a CIR model of the form

$$dv(t) = \kappa(\theta - v(t))dt + \sigma\sqrt{v(t)}dW(t), \quad \kappa, \theta, \sigma > 0.$$

Given $v(u)$, $v(t)$ with $t > u$ is distributed with density

$$\nu \chi_d^2(\nu x, \xi)$$

where $\chi_d^2(x, \xi)$ is the density of a noncentral chi-square random variable with d degrees of freedom and noncentrality parameter ξ and

$$\begin{aligned}
\nu &= \frac{4\kappa}{\sigma^2(1 - e^{-\kappa(t-u)})} \\
\xi &= \frac{4\kappa e^{-\kappa(t-u)}}{\sigma^2(1 - e^{-\kappa(t-u)})} v(u) \\
d &= \frac{4\theta\kappa}{\sigma^2}.
\end{aligned}$$

The conditional mean of $v(t)$ is given by

$$E(v(t)|v(u)) = \nu^{-1}(\xi + d) = (v(u) - \theta)e^{-\kappa(t-u)} + \theta$$

and the conditional second moment is

$$\begin{aligned}
E(v^2(t)|v(u)) &= \frac{(2(d + 2\xi) + (\xi + d)^2)}{\nu^2} \\
&= \left(1 + \frac{2}{d}\right)[E(v(t)|v(u))]^2 - \frac{2}{d} e^{-2\kappa(t-u)} v^2(u).
\end{aligned}$$

References

[AndAnd:01] L. Andersen, J. Andreasen, *Volatility Skews and Extensions of the Libor Market Model*, Applied Mathematical Finance 7, 1-32 (2000)

[AndAnd:02] L. Andersen, J. Andreasen, *Volatile Volatilities*, Risk, vol. 15, no. 12, 163-168 (2002)

[AndBro:01] L. Andersen, R. Brotherton-Ratcliffe, *Extended Libor Market Models with Stochastic Volatility*, Working paper, Gen Re Securities (2001)

[AndPit:07] L. Andersen, V. Piterbarg, *Moment Explosions in Stochastic Volatility Models*, Finance and Stochastics 11, 29-50 (2007)

[Ar:73] L. Arnold, *Stochastische Differentialgleichungen*, Oldenbourg, München (1973)

[BaxRe:96] M. Baxter, A. Rennie, *Financial Calculus*, Cambridge University Press (1996)

[BeMaSch:07] D.Belomestny, S.S.Mathew, J.G.M.Schoenmakers, *A Stochastic Volatility Libor Model and its Robust Calibration*, Preprint No. 1276, WIAS Berlin (2007)

[BelSchoe:06] D. Belomestny, J.G.M. Schoenmakers, *A Jump-Diffusion Libor Model and its Robust Calibration*, Reprint No. 1113, WIAS Berlin (2006)

[BerPle:79] A. Berman, R.J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press (1979)

[Bil:86] P. Billingsley, *Probability and Measure*, second ed., Wiley (1986)

[Bj:98] T. Bjoerk, *Arbitrage Pricing in Continuous Time*, Oxford University Press, Oxford (1998)

[BlaSch:73] F. Black, M. Scholes, *The Pricing of Options and Corporate Liabilities*, Journal of Political Economy (1973)

[Bla:76] F. Black, *The Pricing of Commodity Contracts*, Journal of Financial Economics (1976)

[BrGaMu:97] A.Brace, D.Gatarek, M.Musiela, *The Market Model of Interest Rate Dynamics*, Mathematical Finance 7, 127-147, (1997)

[CheSco:01] R-R. Chen, L. Scott, *Stochastic Volatility and Jumps in Interest Rates: An Empirical Study*, Working Paper, Rutgers University, Morgan Stanley & Co, (2001)

[CoDuGo:02] P. Collin-Dufresne, R.S. Goldstein, *Do Bonds Span the Fixed Income Markets? Theory and Evidence for Unspanned Stochastic Volatility*, Journal of Finance 57, 1685-1729, (2002)

[CoDuGo:01] P. Collin-Dufresne, R.S. Goldstein, *Stochastic Correlation and the Relative Pricing of Caps and Swaptions in a Generalized Affine Framework*, Working paper, Carnegie Mellon University (2001)

[CoTa:03] R. Cont, P. Tankov, *Financial modelling with Jump Processes*, Chapman & Hall (2003)

[CIR:85] J.C. Cox, J.E. Ingersoll, S.A. Ross, *A Theory of the Term Structure of Interest Rates*, Econometrica 53, 385-407, (1985)

[DaiSin:03] Qiang Dai, Kenneth Singleton, *Term Structure Dynamics in Theory and Reality*, Rev Fin, 16:631-678 (2003)

[Dud:89] R.M. Dudley, *Real Analysis and Probability*, Wadsworth& Brooks (1989)

[Duf:96] D. Duffie, *Dynamic Asset Pricing Theory*, second ed., Princeton University Press (1996)

[DPS:00] D. Duffie, J. Pan, K. Singleton, *Transform Analysis and Asset Pricing for Affine Jump-diffusions*, Econometrica, 68(2) (2000)

[Fell:77] W. Feller, *An Introduction to Probability Theory and Its Applications*, Wiley, Vol I, Vol II, New York (1968)

[Gant:77] F.R. Gantmacher, *The Theory of Matrices*, Chelsea Publishing Company, New York (1977)

[Glass:03] P. Glasserman, *Monte Carlo Methods in Financial Engineering*, Springer, New York-Berlin-Heidelberg (2003)

[Harr:79] J.M. Harrison, D. Kreps, *Martingales and Multiperiod Securities Markets*, Journal of Economic Theory (1979)

[Harr:81] J.M. Harrison, S.R. Pliska, *Martingales and stochastic integrals in the theory of continuous trading*, Stoch. Processes& Appl.11,215-260, (1981)

[HeiWu:01] M. Heidari, L. Wu, *Are Interest Rate Derivatives Spanned by the Term Structure of Interest Rates?*, Working paper, Fordham University (2001)

[Hes:93] S.L. Heston, *A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options*, The Review of Financial Studies, Vol.6 (1993)

[Hu:97] J.C. Hull, *Options, Futures, and Other Derivatives*, third ed.,Prentice Hall (1997)

[HuWh:87] J.C. Hull, A. White, *The Pricing of Options on Assets with Stochastic Volatility*, The Journal of Finance, Vol.42, 281-300 (1987)

[HuKe:00] P.J. Hunt, J.E. Kennedy, *Financial Derivatives in Theory and Practice*, Wiley, Chichester (2000)

[IkeWat:92] N. Ikeda, S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North-Holland, Amsterdam (1981)

[JaKaSu:01] R. Jagannathan, A. Kaplin, S. Sun, *An Evaluation of Multifactor CIR Models Using Libor, Swap Rates, and Swaption Prices*, Journal of Econometrics, forthcoming, (2001)

[Jam:97] F. Jamshidian, *Libor and Swap Market Models and Measures*, Journal of Finance and Stochastics, 1, 293-330 (1997)

[Jam:99] F. Jamshidian, *Libor Market Models with Semimartingales* , Workingpaper, NetAnalyticLimited (1999)

[JaLiZh:03] R. Jarrow, H. Li, F. Zhao, *Interest Rate Caps "Smile" Too! But Can the Libor Market Models Capture It?*, working paper (2003)

[KaJae:06] C. Kahl, P. Jaeckel, *Fast Strong Approximation Monte-Carlo Schemes for Stochastic Volatility Models*, Technical Report, ABN AMRO (2006)

[KaSh:91] I. Karatzas, S.E. Shreve, *Brownian Motion and Stochastic Calculus*, second ed., Springer, New York (1991)

[KaSh:98] I. Karatzas, S.E. Shreve, *Methods of Mathematical Finance*, second ed., Springer, New York (1998)

[Ken:94] M. Kendall, *Advanced Theory of Statistics*, 6th ed., Halsted Press, London (1994)

[Kloed:03] P.E. Kloeden, *Numerical Dynamics*, Lecture notes, J. W. Goethe University Frankfurt (2003)

[KloePla:92] P.E. Kloeden, E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin (1992)

[KloePla:94] P.E. Kloeden, E. Platen, *Numerical Solution of SDE Through Computer Experiments*, Springer, Berlin (1994)

[Levy:65] P. Levy, *Processus stochastique et mouvement brownian*, Gauthiers-Villars, Paris (1965)

[LoSaSc:01] F. Longstaff, P. Santa-Clara, E. Schwartz, *The Relative Valuation of Caps and Swaptions: Theory and Evidence*, Journal of Finance 56, 2067-2109, (2001)

[Mar:79] K.V. Mardia, J.T. Kent, J.M. Bibby, *Multivariate Analysis*, third ed., Academic Press, Florida (1982)

[Mer:73] R.C. Merton, *The Theory of Rational Option Pricing*, Bell Journal of Economics and Management Science (1973)

[MiNoe:03] U. Noegel, S. Mikhailov , *Heston's Stochastic Volatility Model. Implementation, Calibration and Some Extensions*, Wilmott Magazin, 74-49, (2003)

[Mikk:02] P. Mikkelsen, *Cross-Currency Libor Market Models*, Working paper, University of Aarhus (2002)

[Mils:95] G.N. Milstein, *Numerical Integration of Stochastic Differential Equations*, Kluwer Academic Publishers (transl. from Russian 1988), (1995)

[MiSaSo:97] K.Miltersen, K.Sandmann, D.Sondermann, *Closed Form Solutions for Term Structure Derivatives with Log-Normal Interest Rates*, Journal of Finance 52, 409-430, (1997) [MuRu:97] M.Musiela, M. Rutkowski, *Martingale Methods in Financial Modelling*, Springer, Berlin (1997)

[Ne:96] S.N. Neftci, *An Introduction to the Mathematics of Financial Derivatives*, Academic Press, San Diego (1996)

[Oks:92] B. Øksendal, *Stochastic Differential Equations*, third ed., Springer, Berlin (1992)

[Pit:05] V.V. Piterbarg, *Stochastic Volatility Model with Time-dependent Skew*, Applied Mathematical Finance 12, no.2, 147-185 (2005)

[Reb:96] R. Rebonato, *Interest-Rate Option Models*, John Wiley & Sons, Chichester (1996)

[RenTou:96] E. Renault, N. Touzi, *Option Hedging and Implied Volatilities in a Stochastic Volatility Model*, Mathematical Finance 6, 279-302 (1996)

[RevYor:94] D. Revuz, M. Yor, *Continuous Martingales and Brownian Motion*, second ed., Springer, Berlin (1994)

[RogWil:00] L.C.G. Rogers, D. Williams, *Diffusions, Markov Processes and Martingales*, third ed., Cambridge University Press (2000)

[Schloe:02] E. Schloegl, *A Multicurrency Extension of the Lognormal Interest Rate Market Models*, Journal of Finance and Stochastics, 6, 173-196 (2002)

[Schoe:05] J.G.M. Schoenmakers, *Robust Libor Modelling and Pricing of Derivative Products*, Chapman & Hall (2005)

[SchoeCo:05] J.G.M. Schoenmakers, B. Coffey, *Libor Rate Models, Related Derivatives and Model Calibration*, Working Paper (1999)

[Shreve:97] S.E. Shreve, *Stochastic Calculus and Finance*, Lecture notes, Carnegie Mellon University (1998)

[Sin:98] C. Sin, *Complications with Stochastic Volatility Models*, Advanced Applied Probability 30, 256-268 (1998)

[Sto:89] J. Stoer, *Numerische Mathematik 1*, Springer, Berlin (1989)

[StoBu:89] J. Stoer, R. Bulirsch, *Numerische Mathematik 2*, Springer, Berlin (1989)

[Vasi:77] O. Vasicek, *An Equilibrium Characterisation of the Term Structure*, Journal of Financial Economics 5,177-188 (1977)

[WuZha:06] L. Wu, F. Zhang; *Libor Market Model with Stochastic Volatility*, Journal of Industrial and Management Optimization, vol 2, 199-207 (2006)

[Zhu:07] J. Zhu; *An Extended Libor Market Model with Nested Stochastic Volatility Dynamics*, Working Paper, Sal. Oppenheim (2007)

Curriculum Vitae

Stanley Sijan Mathew
18. February 1963 in Sofia

60323 Frankfurt am Main
Reuterweg 92
Germany

Schulen und Universitäten :

Apr. 1982 Abitur in Essen (1,9)

1982-1983 Studium der Elektrotechnik an der Ruhruniversität Bochum

1983-1986 Studium der Mathematik an der RWTH Aachen, Nebenfach: Wirtschaft, Vordiplom 1986

1986-1989 Graduate Student in Statistik an der University of California Santa Barbara, UCSB, Fulbright-Stipendium, Masters 1988 (GPA: 3.85)

2004-2005 Studium der Mathematik an der Goethe-Universität in Frankfurt, Diplom Mathematik (1,1), Betreuer: Prof. Dr. Peter E. Kloeden

Wissenschaftliche und Industrielle Arbeitsverhältnisse :

1984-1986 Hilfswissenschaftler am Lehrstuhl A für Mathematik an der RWTH Aachen, Prof. Dr. Paul Leo Butzer

1987-1989 Teaching Assistant am Department of Statistics an der UCSB in Santa Barbara, Prof. Dr. Joseph Gani

1990-1992 Lehrbeauftragter an der Hochschule Fulda, Vorlesungen: Statistik und Wirtschaftsmathematik für Wirtschaftswissenschaftler

1992-2004 Händler für Zinsderivate bei diversen Banken: BHF-Bank, Landesbank Rheinland Pfalz, SEB, Aareal-Bank

2005 - — Wissenschaftlicher Angestellter am Institut für Numerische Analysis der Goethe-Universität, Dissertation mit Prof. Dr. Peter Kloeden und Dr. John Schoenmakers vom WIAS Berlin

Professoren als akademische Lehrer an der RWTH Aachen:

P.L. Butzer, R.J. Nessel, H. Pahlings, B. Rauhut, H.-J. Zimmermann

Professoren als akademische Lehrer an der UCSB Santa Barbara:

J. Gani, P. Todorovic, S. Rachev, R. Jammalamadaka

Professoren als akademische Lehrer an der Goethe-Universität:

P.E. Kloeden, J. Bliedtner, C. Schlag, N. Hofmann,