UNIVERSITÉ PARIS-EST

Thse présentée pour obtenir le grade de

Docteur de l'Université Paris-Est

Spécialité: Mathématiques

par

Ernesto Palidda

Ecole Doctorale : Mathématiques et Sciences et Technologies de l'Information et de la Communication

Modélisation du smile de volatilité pour les produits dérivés de taux d'intérêt.

Thèse soutenue le 29 mai 2015 devant le jury composé de:

Damiano Brigo	*Rapporteur*
Martino Grasselli	*Rapporteur*
Nicole El Karoui	*Examinateur*
Christophe Michel	*Examinateur*
Bernard Lapeyre	*Directeur de thèse*
Aurélien Alfonsi	*Co-Directeur de thèse*

Contents

Acknowledgements

This PhD thesis has been a far longer journey than initially expected: looking back I value every moment of this experience, and I strongly believe it has deeply contributed to change my mindset. Although my PhD thesis officially started in October 2010, and most of the work presented here has been done after that date, I believe this journey has started in July 2008, when I first entered the Groupe de Recherche Opérationnelle as an intern. Since then, I have been exposed to mathematical models used to manage interest rates derivatives. I believe this experience would not be so rich if it wasn't for the people I met along the way. Some of them have deeply contributed to this work and I would like to thank them here.

I am grateful to my PhD supervisor Bernard Lapeyre for being so open and receptive to the format of the PhD thesis, for its curiosity on mechanics of the banking industry, for always being available to meet and discuss the progress of my work, and for its continuous guidance during this work. I am grateful to Aurélien Alfonsi, who co-spervised my PhD thesis, took the time to review some of the most involved computations line by line and allowed me to benefit of his deep understanding of the mathematical tools involved in this work.

I am thankful to Abdelkoddousse Ahdida, who strongly contributed to the advances of this work, and shared with me some unique Sunday's afternoons computing the higher order terms of prices expansions.

I am grateful to Christophe Michel who has "commissioned" the work and internally supervised my work at C.A.-C.I.B. Christophe has always been very open to discuss the progress of my work, and provided me with a valuable business point of view. I am grateful to the member of Christophe's team who took the time to follow the advances during my journey, and in particular to Anas Benabid and Vincent Porte.

I am grateful to a former colleague and friend, Victor Reutenauer, who contributed to the definition of the PhD project, for interesting discussions and his useful comments around this work.

I owe a great deal to Nicole El Karoui, as I believe she gave a great contribution to this work and to my global training in mathematical finance. I am grateful for the long discussions around an extremely high variety of subjects at Crédit Agricole. Not only Nicole shared her extremely vast knowledge of financial mathematics, and her awareness of the latest advances in probability, but also she passed her unique, and out of the box views on the markets and on the derivatives business.

I am grateful to my first mentor Noureddine El Hadj Braiek from whom I learned the basics of interest rates modeling and who shared with me his strong views on best usage practice for managing interest rates derivatives, and in particular on the calibration of interest rates models. I would also like to thank Fabrice Bidaud, who accompanied me in the first part of this long journey and along with Noureddinne helped me learn the basics on interest rates modeling.

I would like to thank the team members of my team at GRO : Karim Mzoughi and Abdelmoula Omar who had the patience and strength to stand me every day for respectively five and three years.

I am grateful to Stephan Krzyzaniak my former line manager and friend who was very understanding and contributed to create a favorable environment where I could continue my PhD work. Stephan also introduced me to the notion of management of interest rate risk in Assets and Liability Management.

I am thankful to Andrée-Lise Remy and Alain Charmant who accepted to recruit me at GRO and allowed me to pursue my PhD thesis while working.

Finally, I would like to thank all the members of GRO team, the interns and in particular Sloane Martin, the most senior member of the team, who was very kind to me since my first day, and made my every day life easy.

Abstract

The first part of this thesis is devoted to the study of an Affine Term Structure Model (ATSM) where we use Wishart-like processes to model the stochastic variance-covariance of interest rates. This work was initially motivated by some thoughts on calibration and model risk in hedging interest rates derivatives. The ambition of our work is to build a model which reduces as much as possible the noise coming from daily re-calibration of the model to the market. It is standard market practice to hedge interest rates derivatives using models with parameters that are calibrated on a daily basis to fit the market prices of a set of well chosen instruments (typically the instrument that will be used to hedge the derivative). The model assumes that the parameters are constant, and the model price is based on this assumption; however since these parameters are re-calibrated, they become in fact stochastic. Therefore, calibration introduces some additional terms in the price dynamics (precisely in the drift term of the dynamics) which can lead to poor P&L explain, and mishedging. The initial idea of our research work is to replace the parameters by factors, and assume a dynamics for these factors, and assume that all the parameters involved in the model are constant. Instead of calibrating the parameters to the market, we fit the value of the factors to the observed market prices.

A large part of this work has been devoted to the development of an efficient numerical framework to implement the model. We study second order discretization schemes for Monte Carlo simulation of the model. We also study efficient methods for pricing vanilla instruments such as swaptions and caplets. In particular, we investigate expansion techniques for prices and volatility of caplets and swaptions. The arguments that we use to obtain the expansion rely on an expansion of the infinitesimal generator with respect to a perturbation factor. Finally we have studied the calibration problem. As mentioned before, the idea of the model we study in this thesis is to keep the parameters of the model constant, and calibrate the values of the factors to fit the market. In particular, we need to calibrate the initial values (or the variations) of the Wishart-like process to fit the market, which introduces a positive semidefinite constraint in the optimization problem. Semidefinite programming (SDP) gives a natural framework to handle this constraint.

The second part of this thesis presents some of the work I have done on the hedging of interest rate risk in ALM. This work was motivated by the business at Crédit Agricole S.A. and in particular by the Financial Division of the bank. The purpose of this part of the dissertation is twofold. First we want to communicate on a field of Finance which is less known by the mathematical finance community, and presents some interesting modeling challenges. Secondly we try to present an original approach to modeling and hedging interest rate risk.

Chapter 6 is an attempt to formalize some of concepts that are used in practice in ALM. We recall some of the key concepts such as the schedule of an asset or a liability and the interest rate gap, and introduce a new concept : the notion of envelope. This concept will look familiar to people used to derivatives pricing, and the hedging of the interest rate risk

of an asset or a liability (closing the gap using the language of ALM) is very similar to the hedging of an option. The remaining chapters present the results of the work we have done in three different projects.

Introduction

The present document presents some of the research work I have done during my PhD thesis. This work was mainly done while I was working at the Groupe de Recherche Opérationnelle in Crédit Agricole. Given the working context in which I have done my PhD thesis, my research work was mainly inspired and sometime directly motivated by concrete demands of the business. The document contains two parts which are independent one to another, and yet describe two different faces of the same business : the fixed income derivative business. The first part of this document is devoted to the study of an Affine Term Structure Model where we use Wishart-like processes to model the stochastic variance-covariance of interest rates. This model aims at providing a robust framework to manage exotic interest rates derivatives. This work has led to a prepublication [AAP14]. The second of this document presents some of the research work I have done for modeling interest rates risk in Asset and Liability Management. The objective of this second part is to try and formalize some of the concepts that are used in practice when managing the interest rates exposure of large banks.

Motivation

The first part of the document presents the study of an Affine Term Structure Model (ATSM) where we use Wishart-like processes to model the stochastic variance-covariance of interest rates. This work was initially motivated by some thoughts on calibration and model risk in hedging interest rates derivatives. The ambition of our work is to build a model which reduces as much as possible the noise coming from daily re-calibration of the model to the market.

It is standard market practice to hedge interest rates derivatives using models with parameters that are calibrated on a daily basis to fit the market prices of a set of well chosen instruments (typically the instrument that will be used to hedge the derivative). The model assumes that the parameters are constant, and the model price is based on this assumption; however since these parameters are re-calibrated, they become in fact stochastic. Therefore, calibration introduces some additional terms in the price dynamics (precisely in the drift term of the dynamics) which can lead to poor P&L explain, and mishedging. The initial idea of our research work is to replace the parameters by factors, and assume a dynamics for these factors, and assume that all the parameters involved in the model are constant.

Instead of calibrating the parameters to the market, we fit the value of the factors to the observed market prices. Considering enough factors to represent the dynamics of the relevant market quantities, we can use the model to produce a robust hedge against these factors. The theoretical framework which has been chosen to achieve this goal is the framework of Affine Term Structure Models [DK96b]. This choice allows us to keep tractability, and also it is very easy to increase the number of factors in the model as long as the additional factor also follows an affine diffusion process. Clearly the choice of affine diffusions is restrictive and limits the ability, and the degrees of freedom of the model to produce a dynamics consistent with the market. We consider this is the lesser of two evils, arguing that even if our dynamical assumption on the factors is wrong, we can always design a conservative pricing of derivatives using the model. For example, we could follow El Karoui et al. [EKJPS98] which, using the monotonicity properties of the price as a function of volatility shows that we can bound the price of a derivative on an underlying which has an unknown dynamics as soon as we can bound the volatility of this underlying. In [ALP95], Avellaneda et al. show that as soon as we can bound the diffusion term of the underlying, we can build a conservative price for any derivative on this underlying by means of non-linear PDE.

The second part of the document presents some of the work I have done on the hedging of interest rate risk in ALM. This work was motivated by the business at Crédit Agricole S.A. and in particular by the Financial Division of the bank. The purpose of this part of the dissertation is twofold. First we want to communicate on a field of Finance which is less known by the mathematical finance community, and presents some interesting modeling challenges. Secondly we try to present an original approach to modeling and hedging interest rate risk.

State of the art

The model we study in this thesis belongs to the class of ATSM. ATSM are an important class of models for interest rates that include the classical and pioneering models of Vasicek [Vas77] and Cox-Ingersoll-Ross [CIR85]. These models have been settled and popularized by the papers of Duffie and Kan [DK96b], Dai and Singleton [DS00] and Duffie, Filipovic and Schachermayer [DFS03]. We refer to Filipovic [Fil09] for a textbook on these term structure models. The Linear Gaussian Model (LGM) is a simple but important subclass of ATSM that assumes that the underlying factors follow a Gaussian process. It has been considered by El Karoui and Lacoste [EKL92] and El Karoui et al. [EKLM+91], and still today widely used in the industry for pricing fixed income derivatives, thanks to its simplicity.

There are several examples of ATSM which allow to reproduce stylized facts such as the volatility smile. Recently Wishart processes have been used to model stochastic volatility of interest rates. The first proposals were formulated in [GMS10], [GS03], [GS07], [Gou06] both in discrete and continuous time, for interest rates and equity underlyings. Applications to multifactor volatility and stochastic correlation can be found in [DFGT08b], [DFGT08a], [DFGI11], [DFGI14], [DFG11], [BPT10], [BL13], [BCT08], both in option pricing and portfolio management. These contributions consider the case of continuous path Wishart processes, [GT08] and [Cuc11] investigate processes lying in the more general symmetric cones state space domain, including the interior of the cone $\mathcal{S}_d^+(\mathbb{R})$. A Wishart ATSM has been defined by Benabid, Bensusan and El Karoui and studied by Bensusan in his Phd thesis [Ben10]. Their model is inspired by the canonical specification of ATSM given by [DS00], and uses Wishart processes to define a rich volatility dynamics. In a series of papers by Gnoatto [Gno13] and Gnoatto and Grasselli, [GG14a], and Gnoatto et al. [BGH13] have considered affine term structure models, in which the yield curve is driven by a Wishart process. A mean reverting Wishart model for the log-prices of commodities has been considered in [CWZ15]. The model we study in this thesis differs in several aspects. The idea of the model is to build a stochastic variance covariance perturbation of the LGM. We want to keep the model structure of the LGM, and in particular the interpretation of some of the factors as the main modes of the yield curve (as explained in 2.4.1) and use the Wishart-like process to model the stochastic variance-covariance of the factors. We chose to consider a general affine variance-covariance dynamics which is more general than the Wishart dynamics which has previously been considered. This allows to account for mean reversion of the variance-covariance, and gives more flexibility to model the term structure of implied volatility. We also considered a correlation structure between what we call the yield curve factors, and the Wishart-like stochastic variance-covariance. The proposed dynamics for the factors of the model is similar to the dynamics of log-prices in Da Fonseca et al. [DFGT08a].

A large part of our work has been devoted to the development of an efficient numerical framework to implement the model. We study second order discretization schemes for Monte Carlo simulation of the model. We merely followed the footsteps of Ahdida and Alfonsi [AA13], who studied the Monte Carlo simulation of Wishart processes and their affine extension. Using a splitting technique they are able to propose exact and high order discretization schemes for Wishart-like processes. These schemes apply directly to our model in the case where there is no correlation between the yield curve factors and the stochastic variance-covariance. The case of a non-zero correlation required some additional work to apply a similar splitting technique.

We also study efficient methods for pricing vanilla instruments such as swaptions and caplets. First, we consider the methods based on the Fourier inversion and the Fast Fourier Transform (FFT) that have been presented by Carr and Madan [CM99] and Lee [Lee04]. These methods can be applied directly for Caplets by working with the forward Caplet price. In order to apply these methods for swaptions we need to perform a classic approximation step. We refer to Schrager and Pelsser [SP06] and Singleton and Umantsev [SU02] for a detailed description of these methods which directly apply to our model. Second, series expansion methods of Gram-Charlier type can also be directly applied for both caplets and swaptions. These methods have previously been applied for pricing swaptions, see for example Collin-Dufresne and Goldstein [CDG02] and Tanaka et al. [TYW05]. We also investigate expansion techniques for prices and volatility of caplets and swaptions. The arguments that we use to obtain the expansion have been developed in the book of Fouque et al. [FPS00]. They rely on an expansion of the infinitesimal generator with respect to a perturbation factor. Recently, this technique was applied by Bergomi and Guyon [BG12] to provide approximation under a multi factor model for the forward variance. We have adapted this approach to interest rates, and in particular of factorial term structure models.

Finally we have studied the calibration problem. There are very few examples where the calibration of interest rates model to the market is discussed formally. See [HKP00] for a notable example. In particular, the calibration of ATSM has been tackled from a statistical estimation point of view, rather than describing how to fit the observed market prices. As mentioned before, the idea of the model we study in this thesis is to keep the parameters of the model constant, and calibrate the values of the factors to fit the market. In particular, we need to calibrate the initial values (or the variations) of the Wishart-like process to fit the market, which introduces a positive semidefinite constraint in the optimization problem. Semidefinite programming (SDP) gives a natural framework to handle this constraint. See Roupin [Rou12] and Boyd [BV03] for an introduction to SDP. SDP has previously been applied to the calibration of the Libor market model by d'Aspremont [d'A03] and Brace et al. [BW00].

The theory of hedging of interest rates risks in ALM has mainly been developed in France. The seminal book of Frachot, Roncalli et al. [DFR03] and the book of Adam [Ada08] are the main references on the subject. These books describe the key notions that are used to manage interest rate risk in ALM[1]. There is a structural reason that explains why the theory of interest rates risk in ALM has mainly been developed in France (and probably also that explains why this subject has not reached a higher level of popularity in the mathematical finance community), and this reason is that France is among the very few countries which has fixed rate retail saving accounts and mortgages. This explains the degree of sophistication required by Financial divisions of large French retail banks. In [DFR03] and [Ada08] the key concepts of ALM are formalized and mathematical definitions for the key concepts are provided. However, the theory of interest rate risk in ALM has not reached the maturity of the theory of pricing derivatives. The theory of hedging interest rate risk in ALM is composed of different layers which are addressed independently and look much more like a practical procedure than a mathematical theory. The concepts of schedule, interest rate gap are often treated separately.

[1] The hedging of interest rates risk in ALM and the modeling of schedules has been studied in a number of internal papers at GRO which provided extensive material for the book [DFR03]. I was lucky to have access to these papers. Some of these papers are available on Thierry Roncalli's webpage thierryroncalli.com

Objective and methodology

As mentioned before the motivation for developing an ATSM using Wishart-like processes was to build a model which is able to reduce the noise coming from re-calibration of the model. The idea is to use only factors for calibration. Markovian term structure models used to price exotic derivative structures are usually calibrated to a suitably chosen set of vanilla interest rates derivatives by adjusting the parameters of the diffusion and the initial values of the factors. Typically the initial value of the factors is set to an arbitrary value (often zero). As the mapping between the underlying factors and the market assets (zero-coupon bonds, interest rates swaps...) is determined by the parameters of the model, calibrating the model parameters is equivalent to calibrate the (parametric) functional mapping between the factors and the market assets. Here, we take the opposite viewpoint. We fix the parameters of the dynamics of the factors, and precisely the parameters determining the drift and diffusion of the factors, and we only calibrate the value of the factors to fit the market prices. We aim to performing a global calibration of the model. The mapping between the value of the factors and the market assets is assumed to be fixed, based on the desired dynamics of the market assets. The model should be able to reproduced the relevant market prices of vanilla interest rates derivatives, and reasonably tractable. In particular the focus has been put on the ability of the model to produce a volatility dynamics for interest rates. The questions we tried to answer were:

- What are the degrees of freedom of Wishart-like processes to reproduce the swaption volatility cube dynamics?

- Is it possible to specify an ATSM using Wishart processes such that we can identify the model factors to the modes of the movements of the swaption volatility cube?

To this day we feel we have only partially answered these questions. The main contribution of our work is to provide some theoretical insights of how to specify the model in such a way that we can indeed identify the factors of the model with the main modes of the yield curve and swaption volatility dynamics, and to provide a number of numerical tools which allow to efficiently implement the model.

The starting point of our work has been the specification of the model. Rather than following the specification proposed by Bensusan et al. [Ben10], we chose to specify the model as a stochastic variance co-variance version of the LGM model, in which we use the Wishart-like process as a stochastic variance-covariance. First it appeared natural to use the Wishart-like processes as a random variance-covariance matrix, secondly we wanted to base our specification on the LGM model. The main reason being that the LGM is widely used in the industry, and it is easy to specify the model in such a way that the factors can be identified with the main modes of the yield curve dynamics. The idea was to keep intact this one to one correspondence between the yield curve modes and the LGM factors in our model, and add a similar identification between the swaptions implied volatilities modes and the variance-covariance factors. We started from a general version of the model which has then been progressively reduced. The reduction has been motivated by tractability, but also by the fact we observed that some of the parameters in the initial model were redundant. The reduction we have made has not reduced the degrees of freedom of the initial model.

The proposed model is a natural generalization of the LGM model using Wishart processes in the spirit of the quadratic gaussian model proposed by Piterbarg [Pit09]. The model allows to generate a volatility smile and a stochastic correlation between rates. The model generates a multi-dimensional dynamics of the volatility. In terms of degrees of freedom of the model, the model is very similar to stochastic volatility specifications of the Dai and Singleton canonical form of ATSM [DS00]. The main difference is that in the model we propose here, the positive definiteness of the stochastic volatility matrix is insured by the positive definiteness of the Wishart process. Other recent affine models use in a similar but different way Wishart processes. We mention here Gnoatto [Gno13] and Gnoatto and Grasselli, [GG14a], and Gnoatto et al. [BGH13]. Let us note that most of the stochastic volatility ATSM presented in the literature and also the most commonly used in the industry consider a one or two dimensional dynamics of the stochastic volatility. See for example [TS10] and [Pit09]. The numerical implementation presented in Chapter 5 considers a model with a volatility driven by a Wishart process of dimension 2 and 4, meaning that the volatility is driven respectively by 3 and 10 factors. It is important to mention that the implementation of the model proposed here is not in line with the standards required to use the model in a production environment. As mentioned before, the underlying motivation for building our model is to avoid re-calibration noise. Therefore we have intentionally studied a fully homogeneous version of the model, as opposed to typical interest rates models used in the industry which often consider a term structure of the diffusion parameters and fit these parameters to exactly reproduce the market prices of the calibration instruments. The price to pay is a less accurate fit. The archetypal example of market model used to manage swaptions is the SABR model [HKLW02]. In fact, SABR is rather a parametric representation of the swaption volatility cube and cannot be used to price exotic interest rates products. Market practice consists in using one SABR model (which means 4 parameters) per swaptions maturity and expiry.

Once the specification of the model was chosen, it was necessary to develop a numerical framework to implement the model. For intellectual honesty, we should mention that the time devoted to the development of such a framework has taken far longer than expected. However, the development of our numerical framework also provided us with a number of tools that helped us having a better theoretical understanding of the model, and confirmed some of our intuitions.

Matrix Riccati differential equations (MRDE) appear as a central tool in the model. The solutions of such equations can blow up in finite time. The study of the MRDE which is presented in 3.1.3 allowed us to derive some sufficient conditions for non explosion. These conditions have been translated in terms of the model parameters so that relevant market quantities such as the zero-coupon bonds and the moments of zero-coupon bonds are well defined. Proposition 14 gives the exact statement for the existence of the general Fourier/Laplace transform of the model state variables. Also, we were able to derive some constraints on the model parameters so that the state variables verify some ergodicity property. Proposition 15 gives the exact statement. Given the affine structure of the model, and once established the domain of existence of the MRDE (or at least a subset of this domain) it was straightforward to apply Fourier inversion and the Fast Fourier Transform (FFT) based method and series expansion methods of Gram-Charlier type for pricing vanilla instruments. These methods are efficient in terms of their computational cost, and provide reasonably accurate results in terms for pricing caplets. We are more reserved in using these methods for pricing of swaptions. Not completely satisfied by transform based and series expansion methods, we have investigated an alternative route for pricing vanilla instruments. The paper of Bergomi and Guyon [BG12] gives a very straightforward way to derive an expansion of the European option prices and of the implied volatility in stochastic volatility models in the order 2 of the volatility of volatility. This approach seemed natural given that we view the model as a perturbation of LGM. This expansion methodology cannot be directly applied to ATSM. One of the key elements to derive the expansion is the fact that the payoff of option does not depend on the volatility of the underlying. However, this is not true in ATSM. Even in the simple Vasicek model, the yield curve depends on the volatility of the spot rate, and therefore the payoff of caplets and swaptions also depend on the volatility of the spot rate. Using a change of variable we where able to eliminate the dependence of the payoff on the volatility. With respect to the transform methods and series expansion methods, getting an expansion on the smile is complementary. On the one hand, it is less accurate to calculate a single price since we only calculate here the expansion up to order 2 in the volatility of volatility. On the other hand, it is more tractable for a first calibration of the model and gives a good approximation for key quantities on the smile. Also, it confirms our intuitions on the main drivers of the smile and on the role of the model parameters and factors in terms of the volatility statics and dynamics.

Finally, we have focused our efforts into extending the work of Ahdida and Alfonsi [AA13] to build an efficient Monte Carlo simulation framework for the model. In 3.2.3, we show that in presence of a non-trivial correlation structure a remarkable splitting property similar to the one verified by the generator of Wishart processes of the generator still holds for the generator of the state variables of the model. This allows to apply the splitting technique already applied in [AA13] to build second order discretization schemes for the model. It is interesting to note that the remarkable splitting property of the generator of the state variables appears again in section 4.3 for the price and volatility expansion of caplets, and shows that the implied volatility structure is similar to the structure of the generator.

Let us note that, except for the price and volatility expansion we develop in section 4.3

(which would need some adaptation), the numerical framework developed for the model (and in particular the Monte Carlo simulation framework) applies to a more general specification which includes the interest rates models of Bensusan et al. [Ben10], and Gnoatto [Gno13] and [BGH13], and the equity model of Da Fonseca et al. [DFGT08a].

Once we had finished developing an efficient numerical framework for the model implementation, we have tested the model. In particular, we have tested the ability of the model to capture the market implied volatility from a static point of view as well as from a dynamic point of view. Based on the results of our numerical investigation we can conclude that in the cases where the view of the model as a perturbation of LGM is valid, and the expansion technique works, the at-the-money smile is very close to the zero order term of the expansion. Not surprisingly, this term is the implied volatility in the LGM model, and the at-the-money cumulated variance is a linear function of the Wishart-like process. This suggested a stepwise calibration of the model. The first step of the calibration being the calibration of the zero order term of the expansion to the at-the-money surface of swaptions volatilities. We define two alternative calibration problems, the first consists in calibrating the values of the volatility factors to fit the implied volatility of swaptions, the second consists in calibrating the variations of the volatility factors to fit the variations of the implied volatility of swaptions. We followed the work of d'Aspremont [d'A03] and Brace et al. [BW00] who have used a similar approach for the calibration of the libor market model, and proposed a solution using SDP. In fact, both our calibration problems can be formulated as the optimization of a linear criteria under a positive semidefinite constraint. SDP provides a natural framework to solve such problems. The results of our calibration experiments are presented in 5.3.

We have only partially explored the potential of the model, and in particular its ability to produce a robust hedging of interest rates products. We leave it for future investigation to put the model into numerical practice and test the hedging strategies produced by the model. An interesting topic to investigate is to consider a constrained calibration of the model. Following the interpretation we have given of the model and based on the calibration results we have obtained, we think that the model provides additional degrees of freedom which have not been exploited. For example the factors driving the correlation between the factors. It is fairly easy to add some constraints to the SDP framework, and this would allow -while being calibrated- to build conservative or aggressive prices for an additional derivative (for example a CMS spread option). A similar study has been made by d'Aspremont in [d'A03], where conservative and aggressive bounds for swaptions are obtained in a calibrated model.

The work on modeling and hedging interest rates risk in ALM was motivated by the demand of the Financial division of Crédit Agricole. Chapter 6 is an attempt to formalize some of concepts that are used in practice in ALM. We recall some of the key concepts defined in Frachot et al. [DFR03] and Adam [Ada08], such as the schedule of an asset or a liability and the interest rate gap, and introduce a new concept : the notion of envelope.

This concept will look familiar to people used to derivatives pricing, and the hedging of the interest rate risk of an asset or a liability (closing the gap using the language of ALM) is very similar to the hedging of an option. The remaining chapters present the results of the work we have done in three different projects. Chapter 7 presents two projects in which we have proposed a modeling approach to hedge the interest rate risk associated to specific retail banking products : the Livret A and the PEL. Both these products include hidden and explicit optionality, and require -to some extent- that we are able to model the behavior of the clients. The livret A is an inflation-indexed, state-regulated retail product which includes optionality, and we have chosen to model the product as a proper hybrid inflation/rates derivative. We propose different modeling approaches and study the corresponding hedging strategy. The PEL is a very complex retail product which contains a number of optional features, and can be converted into a mortgage. A key element for hedging this product is the estimation of the time at which the client decides to close or convert its account. Following the initial work of [BDJ$^+$00] we have proposed a model that aims to model the different behaviors of the clients. The model is based on the (unrealistic) assumption that the client has a rational behavior and follows a optimal decision process. We found that -to some extent- the model is able to model different client behaviors and provides an estimation of the distribution of the time at which the account is closed or converted into a mortgage. We also backtested the model on historical data. Finally chapter 8 compares two alternative ways to hedge options. The difference lies in the choice of the instruments used to hedge the option. These strategies are well known, and can be found in any derivative textbook. The risk-neutral hedging strategy uses the classic hedging portfolio made of the underlying asset and of the bank account. The forward-neutral strategy replaces the bank account by a loan with maturity the expiry of the trade (or equivalently by a zero-coupon of maturity the expiry of the trade). We discuss the implementation of the strategies in practice to hedge options in ALM and present the results of the applications of such strategies in a very simple model. The two approaches give surprisingly different results in terms of the delta (the quantity of underlying asset in the hedging portfolio) and in terms of the variance of the "hedging error" of the strategy.

Document structure

Part I introduces some important notions of the interest rates derivatives market and introduces the key notions of term structure modeling. In particular, we focus on factorial term structure models, and ATSM. In chapter 1, we present the interest rates market, we define the basic products, and we discuss some of the peculiarities and recent evolutions of this market. The last section of the chapter presents the results of a brief empirical analysis of the interest rates swap and swaption US data. In chapter 2 we present the key notions of term structure modeling and present the modeling framework in which we will work in the following. In Section 2.1, we provide a very brief presentation of the Heath Jarrow and Morton (HJM) [HJM92] framework. In Section 2.2 we present the idea of modeling using state variables. Section 2.3 presents the class of ATSM. In Section 2.4, we give notable examples of models which belong to the class of ATSM. In section 2.5, we briefly present the Markov Functional model of Hunt, Kennedy and Pessler [HKP00]. In section 2.6 we discuss the nature of volatility in factorial term structure models, and illustrate the notion of model risk through an example. Finally, Section 2.7 we discuss hedging in factorial term structure models.

Part II is devoted to the definition and study of an ATSM using Wishart-like process to model stochastic variance-covariance. In chapter 3 we define and study the dynamics of the state variables of the model. In Section 3.1, we study the theoretical properties of the process, providing precise statement on the infinitesimal generator, on the Fourier/Laplace transform, and on the ergodicity properties of the process. In Section 3.2 we provide second order discretization schemes for Monte Carlo simulation of the state variables. In chapter 4, we study the stochastic variance-covariance perturbation of the LGM, and provide the numerical framework for the model. In Section 4.1, we define the model, we derive the bond reconstruction formula and the change of measure, and discuss the specification of the model. In Section 4.2, we discuss the application of methods based on Fourier inversion and series expansions for pricing caplets and swaptions. We presents some numerical results of the application of these techniques for pricing caplets. In Section 4.3, we present an expansion method for pricing caplets and swaptions, and present some numerical tests of the method. Finally, Section 4.4 briefly discusses the hedging in the model. In chapter 5 we propose a calibration method and present some numerical results. In particular, we test the ability of the model to fit the volatility of swaptions and to reproduce the dynamics of the implied volatility of swaptions. In Section 5.1, we present SDP framework, which is the numerical tool that will be used to solve the calibration problem. In Section 5.2 we discuss how to chose the model parameters (which are meant to be constant) in such a way that the role of the factors is clearly identified. In Section 5.3, we present the results of our numerical experiments.

Part III is devoted to the modeling and hedging of interest rate risk in ALM. In chapter 6 we present the key notions in ALM, such as the schedule of an asset or a liability and the interest rates gap. In chapter 7 presents two projects in which we have proposed a modeling approach to hedge the interest rate risk associated to specific retail banking products : the Livret A and the PEL. Finally chapter 8 compares two alternative different ways to hedge options in ALM.

Introduction (Version francaise)

Ce mémoire présente une partie du travail de recherche que j'ai effectué dans le cadre de ma thèse. Ce travail a été principalement effectué durant ma permanence au Groupe de Recherche Opérationnelle du Crédit Agricole. Etant donné le contexte dans lequel j'ai conduit mon travail de thèse, celui-ci a été principalement inspiré, et parfois même directement motivé par les besoins concrets des équipes opérationnelles. Le document contient deux parties qui sont indépendantes, et néanmoins représentent deux faces de la même activité: la couverture du risque de taux d'intérêt. Dans la première partie du document on étudie un modèle affine de la dynamique de la courbe des taux, ou un processus affine dans l'espace des matrices semidéfinies positives de type Wishart est utilisé pour modéliser la dynamique de la variance-covariance entre les taux d'intérêt. L'ambition de notre travail est la construction d'un modèle qui fournisse une couverture globale et robuste des risques d'un book de produits exotiques de taux. Ce travail a conduit à une pré-publication [AAP14]. La deuxième partie du document est dédiée à la couverture du risque de taux dans la gestion actif-passif du bilan d'une banque. L'objectif de cette seconde partie est de formaliser les principaux concepts qui sont utilisés en pratique dans ce domaine.

Motivation

La première partie du document étudie un modèle affine de la dynamique de la courbe des taux, ou un processus affine dans l'espace des matrices semidéfinies positives de type Wishart est utilisé pour modéliser la dynamique de la variance-covariance entre les taux d'intérêt. Ce travail a été initialement motivé par une réflexion sur la calibration et le risque de modèle dans la couverture des produits dérivés de taux d'intérêt. L'ambition de notre travail est de construire un modèle qui réduise, voire élimine les biais induits par la re-calibration.

La pratique de marché consiste à valoriser et couvrir les produits dérivés de taux en utilisant des modèles dont les paramètres sont régulièrement calibrés pour reproduire les prix de marché d'un ensemble d'instruments (typiquement les instruments qui seront ensuite utilisés pour la couverture du produit). Le modèle suppose que les paramètres sont constant, hors le procédé de calibration rend ces paramètres stochastiques, et induit un biais dans la dynamique des prix (plus précisément dans le drift de la dynamique du prix). Ce biais peut introduire des erreurs de couverture et violer la condition d'autofinancement de la stratégie de couverture. L'idée du modèle que nous présentons, est de remplacer les paramètres par des facteurs, que l'on suppose stochastiques. Nous proposons alors de calibrer les valeurs des facteurs et de maintenir les paramètres du modèle constants. Si l'on considère suffisament de facteurs de sorte a reproduire la dynamique des principaux facteurs de risques du produit que l'on souhaite valoriser et couvrir, le modèle permet alors de produire une stratégie de couverture autofinancé qui ne comporte pas de biais de re-calibration. Le cadre théorique que nous avons choisi pour notre modèle est celui des modèles affines de taux d'intérêt défini par Duffie et Kan [DK96b]. Ce choix nous permet de limiter la compléxité d'implémentation du modèle. Aussi, il est très simple d'augmenter le nombre de facteurs du modèle tout en restant dans le même cadre. Il suffit de supposer que le facteur supplémentaire est lui-même affine. Le choix d'un processus affine est clairement restrictif, et ne permet certainement pas de produire une dynamique cohérente avec celle observé dans le marché. Nous considérons que c'est le moindre des deux maux. Notre utilisation du modèle repose davantage sur l'exploitation de la forme fonctionelle entre les facteurs du modèle et les prix (que l'on suppose stable), que sur la pertinence des hypothèses dynamiques du modèle. Une fois déduite la forme fonctionelle donnant les prix en fonction des facteurs[2] par le modèle, nous oublions notre hypothèse sur la dynamique des facteurs et implicitons les facteurs des prix de marché. Alors, il est toujours possible de construire un prix conservateur et une stratégie de couverture correpondante. Par example, à l'aide des propriétés de monotonie des prix d'options en temps que fonctions de la volatilité, El Karoui et al. [EKJPS98] montrent qu'il est possible de borner le prix d'une option dès lors que l'on peut borner la volatilité du soujacent. Le modèle de volatilté incertaine d'Avellaneda [ALP95] est un exemple de construction de prix conservateurs dans un modèle ou la dynamique de la volatilité du soujacent est inconnue, mais bornée. De même nous pourrions envisager de construire des prix conservateurs en supposant que les facteurs sont bornés.

La deuxième partie de ce mémoire présente quelques travaux sur la couverture du risque de taux pour la gestion actif-passif du bilan de la banque. Ce travail a été motivé par les besoins de la direction financière de Crédit Agricole S.A. Les principaux objectifs de cette partie du mémoire sont: d'une part de communiquer sur un domaine qui est peux connu de la communauté des mathématiques financière, et qui (selon notre opinion) présente des défis de modèlisation importants; d'autre part nous proposons une approche originale à la modélisation du risque de taux en ALM.

Etat de l'art

Le modèle que nous étudions dans cette thèse s'inscrit dans le cadre des modèles affines de taux d'intéret. Les modèles affines de taux sont une classe importante de modèles, en font partie le modèle de Vasicek [Vas77] et Cox-Ingersoll-Ross [CIR85]. Un cadre formel a été défini et développé par Duffie and Kan [DK96b], Dai et Singleton [DS00] et Duffie, Filipovic et Schachermayer [DFS03]. Nous renvoyons à Filipovic [Fil09] pour une référence sur les modèles affines. Le modèle linéaire Gaussien (LGM) est un exemple important de

[2] qui certe repose sur des hypothèse de la dynamique des facteurs.

modèle affine qui suppose que les facteurs sont gaussiens. Il a été initialement considéré par El Karoui et Lacoste [EKL92] and El Karoui et al. [EKLM+91], et est très répandu dans l'industrie pour la valorisation des dérivés de taux d'intérêt.

Il y a de nombreux modèles affines qui permettent de reproduire le smile de volatilité observé dans le marché. Récemment les processus de type Wishart ont été utilisés pour modéliser la volatilité des taux d'intérêt. Un modèle affine Wishart a été défini par Benabid, Bensusan et El Karoui et a été étudié par Bensusan dans sa thèse [Ben10]. Le modèle est inspiré de la specification canonique des modèles affines de Dai et Singleton [DS00], et le processus de Wishart est utilisé pour définir une riche dynamique de la volatilité. Dans une série de papiers Gnoatto [Gno13] et Gnoatto et Grasselli, [GG14a], et Gnoatto et al. [BGH13] ont étudié un modèle affine, dans lequel le taux court est une fonction du processus de Wishart uniquement. Le modèle que nous étudions dans cette thèse diffère pour plusieurs aspects. L'idée du modèle est celui d'une perturbation à variance-covariance stochastique du LGM. Nous souhaitons maintenir la structure du LGM, et en particulier l'interpretation de certains des facteurs comme des vecteurs propres de la dynamique de la courbe de taux (voire 2.4.1) et utilisé le processus de type Wishart pour modéliser la variance-covariance de ces facteurs. Aussi, nous considérons un processus affine dans l'espace des matrices semidefinies qui est plus général qu'un processus de Wishart. Ceci nous permet de prendre en compte des effets de type retour à la moyenne pour la volatilité et permet de modéliser la term structure de la volatilité implicite. Enfin nous avons considéré une structure de corrélation entre les facteurs de taux et les facteurs de volatilités. La dynamique des facteurs du modèle est proche de la dynamique des log-prix proposé par Da Fonseca et al. [DFGT08a].

Une grande partie de notre travail a été dévoué à l'étude de méthodes numériques permettant une implémentation efficace du modèle. Nous proposons des schémas de discrétisation d'ordre deux pour la simulation Monte Carlo. Nous avons suivi Ahdida et Alfonsi [AA13] qui ont proposé des schémas exact et d'ordre supérieur pour les processus de Wishart et leurs extensions affines. Dans le cas d'une corrélation nulle entre les facteurs de volatilité et les facteurs de courbe, ces schémas s'appliquent directement à notre modèle et permettent de construire un schéma d'ordre 2. Le cas général a necessité une extension du résultat de décomposition du générateur infinitésimal valable pour le Wishart.

On a également étudié des méthodes efficaces pour valoriser les options vanilles, tels que swaptions et caplets. Les méthodes basés sur l'inversion de Fourier présentées par Carr et Madan [CM99] et Lee [Lee04] sont particulièrement adaptées au cadre des modèles affines, du fait de l'existence de formules analytiques ou semi-analytiques pour la transformé de Laplace/Fourier. Ces méthodes s'appliquent directement à notre modèle pour la valorisation des caplets. L'application de ces méthodes pour la valorisation des swaptions nécessite d'une approximation classique. Nous renvoyons à Schrager et Pelsser [SP06] et Singleton and Umantsev [SU02] pour une description détaillé de ces méthodes qui s'applique directement à notre modèle. Une alternative aux méthodes par inversion de Fourier est donné par les méthodes d'expansion en série de la distribution de type Gram-Charlier ou Edgeworth. Ces méthodes s'appliquent directement à notre modèle pour la valorisation de caplets et swaptions. Ces méthodes ont été appliqué à la valorisation de swaptions dans les modèles affines, nous renvoyons à Collin-Dufresne et Goldstein [CDG02] et Tanaka et al. [TYW05]. Nous avons ensuite étudié des techniques d'expansion pour les prix et la volatilité des caplets et des swaptions. On utilise une expansion du générateur infinitésimal par rapport à un facteur de perturbation. Ces techniques sont développées dans le livre de Fouque et al. [FPS00]. Récemment , cette technique a été appliquée par Bergomi et Guyon [BG12] pour déduire une expansion de la volatilité implicite dans un modèle multi-facteur de la variance forward. Nous avons adapté cette approche au contexte des taux d'intérêt, et en particulier au cadre des modèles factoriels de taux.

Nous avons étudié le problème de la calibration du modèle. Il y a très peux d'exemples dans la littérature ou les techniques de calibration du modèle sont décrites et analysées de facon formelle. Une exception remarquable est le modle Markov functional de Hunt et al. [HKP00]. En particulier, la calibration des modèles affine a été davantage étudié d'un point de vue de l'estimation statistique des paramètres du modèle, que de la résolution d'un problème inverse pour reproduire les prix observés dans le marché. La philosophie du modèle étant de calibré les facteurs du modèle, nous devons en particulier calibré les facteurs de type Wishart, ce qui introduit une contrainte semidefinie positive dans le problème de calibration. La programmation semidéfinie (SDP) fourni un cadre formel naturel pour gérer une telle contrainte. Nous renvoyons à Roupin [Rou12] et Boyd [BV03] pour une introduction sur la SDP. L'application de la SDP au problème inverse de calibration d'un Libor market model à été étudié par d'Aspremont [d'A03] et par Brace et al. [BW00].

La théorie de la gestion du risque de taux en ALM a principalement été développée en France. L'oeuvre seminale de Frachot, Roncalli et al. [DFR03] et le livre d'Adam [Ada08] sont les principales références dans le domaine. Ces livres décrivent les notions fondamentales qui sont utilisées pour calculer et gérer le risque de taux dans la gestion actif-passif du bilan d'une banque. Il y a une raison structurelle qui explique pourquoi la théorie du risque de taux en ALM a principalement été développé en France, ainsi que le niveau de sophistication nécessaire dans les direction financières des grandes banques en France. La France est parmi les seuls pays qui propose des livrets d'épargne et des prêts à taux fixe.

Objectif et méthodologie

Nous insistons sur la principale motivation du modèle que nous développons dans cette thèse. Nous souhaitons construire un modèle qui reduise le bruit du à la re-calibration. Le point de vue que nous adoptons consiste à remplacer les paramètres du modèle par des facteurs, et à calibrer les valeures des facteurs ou de leurs variations aux prix de marché. Les contraintes que nous imposons sur notre modèle est qu'il soit capable de reproduire les prix des quantités de marché pertinentes, soit les produits optionnels de taux d'intérêt.

En particulier nous concentrons notre attention sur la capacité du modèle à reproduire une dynamique de la volatilité des taux d'intérêt. Nous essayons de répondre aux questions suivantes:

- Quels sont les degrés de liberté du modèle pour reproduire la dynamique du cube de volailité des swaptions?

- Est-t-il possible de spécifier le modèle de sorte à identifier les facteurs du modèle avec les vecteurs propres du cube de volatilité des swaptions?

Les résultats de nos travaux ne répondent que très partiellement à ces questions. La principale contribution de notre travail est de fournir une étude théorique sur la spécification du modèle. En particulier, nous fournissons des résultats qui permettent de justifier théoriquement les intuitions sur le modèle, et donnent des arguments solides pour spécifier le modèle de telle sorte à identifier les facteurs avec les mouvements principaux de la courbe de taux et du cube de volatilité implicite des swaptions. Aussi, nous fournissons un cadre numérique robuste qui permet une implémentation efficace du modèle.

Le point de départ de notre recherche a été le choix de la spécification du modèle. Comme indiqué précédemment, le modèle que nous proposons diffère des examples de modèle affines Wishart existants dans la littérature. Nous avons choisi de spécifier le modèle comme une perturbation à variance-covariance stochastique de LGM. Il nous parait avant tout naturel d'utiliser les processus de type Wishart pour modéliser la dynamique de la matrice de variance-covariance des facteurs de la courbe de taux. Ensuite, nous souhaitions baser notre specification sur le LGM. La raison principale étant que celui-ci permet une identification claire de ses facteurs avec les mouvements principaux de la courbe de taux (niveau, pente, courbure...). L'idée de notre modèle est de maintenir intacte cette identification et d'obtenir une identification similaire entre les facteurs de volatilité et le cube de volatilité implicite des swaptions.

Une fois choisie la spécification du modèle, il a été nécessaire développer un cadre numérique permettant une implementation efficace du modèle. Par honnêteté intellectuelle nous devons admettre que le développement de techniques numériques pour l'implémentation du modèle a pris plus de temps que prévu. Néanmoins, le développement de ces techniques a enrichi notre rflexion sur le modèle et a parfois fourni des outils théoriques permettant de confirmer nos intuitions sur le modèle.

Les equations différentielles de Riccati matricielles ont un role essentiel dans le développement du modèles. Elles apparaissent dans la formule de reconstruction des zéro-coupon, ainsi que pour les transformés de Fourier/Laplace des facteurs. Les solutions de telles equations peuvent exploser en temps fini. L'étude de l'équation différentielle de Riccati matricielle présenté en 3.1.3 nous a permis de déduire des conditions suffisantes pour que la solution n'explose pas. Nous avons traduit ces conditions en terme des paramètres de modèle, de sorte à ce que les quantités de marché telles que les zéros-coupon et leurs moments soient finis. La proposition 14 donne une formulation précise du résultat. Aussi, nous déduisons de cette étude des conditions sur les paramètre de diffusion des facteurs pour que ceux-ci vérifient une propriété d'ergodicité. La proposition 15 donne une formulation précise du résultat. Etant donnée la structure affine du modèle, et étant établi le domaine d'existance des solutions de l'equation différentielle de Riccati matricielle il était naturel d'appliquer des methods par inversion de Fourier et de méthodes par expansion de la densité (ou de la fonction de répartition) de type Gram-Charlier/Edgewoth pour la valorisation des produits dérivés vanille. Ces méthodes sont efficaces en terme d'implémentation et donne des résultats satisfaisants pour la valorisation des caplets. Nous avons d'avantage de réserves quant à l'application de ces méthodes pour la valorisation des swaptions. Non satisfait par ces méthodes, nous avons exploré une approche alternative pour valoriser les instruments vanille. Le papier de Bergomi et Guyon [BG12] donne une approche directe permettant de développer une expansion des prix et de la volatilité implicite pour les options européennes à l'ordre 2 en la volatilité de volatilité. Etant donné la spécification du modèle comme perturbation de LGM, il nous a paru naturel d'appliquer cette approche. Néanmoins, l'expansion ne peux s'appliquer telle quelle aux contexte des modèles affines. Un des éléments clé pour déduire l'expansion est le fait que le payoff de l'option ne dépend pas de la volatilité du soujacent. Hors ceci n'est pas le cas dans le cadre des modèles affines de taux d'intérêt. Même dans le modèle de Vasicek, la courbe de taux dépend de la volatilité du taux court, et donc le payoff des caplets et swaptions, dépendent aussi de la volatilité. En utilisant un changement de variable nous éliminons cette dépendance. L'expansion que nous développons fourni d'une part une méthode numérique pour valoriser les options vanilles, complémentaire aux méthodes par inversion de Fourier et par expansion de la densité en série; et d'autre part donne un outils théorique permettant d'analyser l'impact des facteurs sur la forme et la dynamique du cube de volatilité.

Enfin nous avons étendu l'excellent travail de Ahdida et Alfonsi [AA13] pour fournir un cadre numérique permettant de simuler le modèle de facon efficace. Dans 3.2.3 nous montrons que en présence d'une structure de corrélation non triviale, le générateur des facteurs vérifie une propriété de décomposition similaire à celle utilisé dans [AA13]. Il est intéressant de remarquer que la propriété de décomposition du générateur des facteurs se retrouve dans l'expansion des prix et de la volatilité des caplets et swaptions 4.3.

A l'expection de l'expansion des caplets et swaptions dévelopés en 4.3, les méthodes numériques proposées dans cette thèse, et en particulier les schémas de discrétisation pour la simulation Monte Carlo s'appliquent à une spécification bien plus générale que celle que vous forunissons ici, et qui inclu les modèles de Bensusan et al. [Ben10], de Gnoatto [Gno13] et [BGH13], et le modèle actions Da Fonseca et al. [DFGT08a].

Une fois pourvu des outils numériques permettant d'implémenter le modèle, nous avons testé le modèle. En particulier, nous avons testé la capacité du modèle à reproduire la volatilité implicite observée dans le marché, d'un point de vue statique et dynamique. Nos expériences numériques montrent que lorsque le point de vue du modèle comme perturbation de LGM est valide, et l'expansion donne des résultat satisfaisants, la volatilité à la monnaie des swaptions et caplets est proche du terme d'ordre zero de l'expansion. Sans surprise, ce terme correspond exactement à la volatilité implicite que donnerait le LGM avec une volatilité deterministe. Aussi, la variance cumulée à la monnaie est une fonction linéaire du processus de type Wishart. Ce résultat suggère une méthode de calibration séquentielle, dont la première étape consiste à calibrer le terme d'ordre zero de l'expansion sur la volatilité à la monnaie des caplets et swaptions. Nous définissons deux problèmes de calibration alternatifs : le premier consite à calibrer la valeur des facteurs de volatilité pour reproduire les prix des swaptions; le second consiste à calibrer les variations des facteurs de volatilités pour reproduire les variations des prix des swaptions. Nous avons suivi le travail de d'Aspremont [d'A03] et Brace et al. [BW00] qui ont formulé le problème de calibration d'un modèle de type BGM de facon similaire, et ont proposé une solution en utilisant la SDP. Nos deux problèmes de calibration peuvent être formulés dans un cadre de programmation SDP. La SDP fourni un cadre théorique naturel permettant de résoudre de telles problèmes. Les résultats numériques sont présentés en 5.3.

Nous n'avons que partiellemnt exploré le potentiel du modèle, et en particulier sa capacité à produire une couverture robuste pour les dérivés de taux d'intérê. Nous laissons à des travaux futurs l'analyse numérique des stratégies de couverture produites par le modèle. Un thème de recherche qui nous parait porteur est d'ajouter des contraintes au problème de calibration. Les résultats de nos expériences de calibration et l'interprétation du modèle suggèrent que le modèle comporte des degrés de liberté que nous n'avons pas encore exploré. Par exemple, les facteurs qui gouverne la corrélation. Il est relativement simple d'ajouter des contraintes dans un problème SDP, et ceci permettrai, tout en étant calibré au prix de marché, de construire un prix conservateur ou aggressif pour un produit dérivé supplémentaire (par example un CMS spread). Une étude similaire a été faite par d'Aspremont dans [d'A03], ou sont obtenus des bornes aggressives et conservatrices pour les swaptions dans un modèle calibré.

Le travail de recherche pour la modélisation du risque de taux dans la gestion actif-passif a été motivé par les demandes de la direction financière de Crédit Agricole. Le chapitre 6 est une tentative de formalisation de certains des concepts qui sont utilisés dans la pratique de marché. Nous rapellons les concepts essentiels définis par Frachot et al. [DFR03] et Adam [Ada08], tels que l'écoulement d'un poste du bilan et le gap de taux, et introduisons un nouveau concept: le concept d'enveloppe. Ce concept permet d'assimiler la couverture d'un poste du bilan, à la couverture d'un produit dérivé optionnel. Le chapitre 7 présente les travaux que nous avons menés dans le cadre de deux études, dans lesquelles nous avons proposé une modélisation permettant de couvrir l'exposition de la banque à des produits d'épargne spécifiques: le livret A et le PEL. Ces produits contienne une dimension optionnelle implicite et explicite, and nécessite une modélisation du comportement des clients. Le livret A est un produit indéxé sur l'inflation régulé par l'Etat qui comporte de l'optionnalité. Nous avons choisi de le modéliser comme un produit hybride taux/inflation. Nous étudions différentes modélisations, et analysons les stratégies de couverture qui en résultent. Le PEL est un produit d'épargne très complexe qui contient un certain nombre de clauses optionnelles et peut être converti en crédit. Un élément clé pour la couverture d'un tel produit est l'estimation du temps de sortie (défini comme le moment ou le client décide soit de convertir, soit de clôturer sont contrat). Nous avons suivi les travaux de [BDJ+00], et proposons un modèle qui décrit différent comportements possibles du client. Se modèle se base sur l'hypothèse discutable que le client agit suivant un comportement rationel et suis une stratégie de sortie optimale. Nous montrons que le modèle permet (en une certaine mesure) de modéliser le comportement des différents clients et fourni une estimation de la distribution des temps de sortie pour ceux-ci. Nous avons aussi backtesté le modèle sur des données historiques. Enfin le chapitre 8 compare deux stratégies de couverture alternatives pour les options. La différence réside dans le choix des instruments utilisés pour la couverture. Ces stratégies sont classiques et elles sont dcrites dans n'importe quel livre sur les dérivés de taux. La stratégie que nous appellons risque-neutre utilise le portefeuille de couverture classique constitué de l'actif risqué et du compte cash. La stratégie que nous appelons forward-neutre remplace le compte cash par un zéro-coupon ou de facon équivalente par un pret de maturité la maturité de l'option. Nous discutons l'implémentation de ces stratégies en pratiquepour couvrir les options dans le cadre de la gestion actif-passif. Les deux approches sont très différentes en termes du delta qu'elles produisent et en termes de la variance de l'erreur de couverture de la stratégie.

Part I

Models for managing interest rates derivatives

Chapter 1

The interest rates market

In this chapter we provide a brief description of the interest rates market. We will mainly focus on the over the counter (OTC) interest rates derivatives market and on the products for which we will develop pricing models in the following. We describe the products and the main market practices. We also mention the latest evolutions in the market and in particular the different adjustments of the price that are necessary in order to take into account of funding, liquidity and credit risks. The main contribution of this thesis is to propose an innovative TS model to manage IR derivatives. It is important to describe the modeling challenges of today's market.

1.1 Reference rates and products

Interest rates products trade OTC and typically via inter-dealer brokers. These act as intermediaries, they facilitate price discovery and transparency by communicating dealer interests and transactions, enhance liquidity, and allow financial institutions anonymity in terms of their trading activities. A clear distinction is to be made between the vanilla interest rates market and the exotic interest rates market. Sometime the line separating the two is quite thin, and this is particularly true considering the latest developments in the market. Vanilla products are supposed to be liquidly traded in the market and though they are not traded on an organized market there is enough market consensus to assume that their price results from the equilibrium of offer and demand. The management of an individual vanilla trade doesn't really require a model, of course this is not true if we consider a portfolio of vanilla trades. Exotic products are highly non liquid and their prices is highly dependent on the model one chooses for pricing. Their management requires an accurate analysis of the underlying risks and an accurate design of an hedging strategy.

1.1.1 Reference rates

IBOR and overnight rates are the main reference rates that underly the most liquid interest rates derivative products. IBOR stands for interbank offered rate and is the average interest rate estimated by leading banks in the main market places (the main example is LIBOR which is the rate in the London market) that they would be charged if borrowing from other banks, standard maturities for IBOR rates are 1, 3 and 6 months. This rate is purely declarative and it has recently been exposed to manipulations, and it should (in theory at least) be as close as possible to the rate at which banks effectively borrow. IBOR rates are used as the reference for the floating legs of interest rates swaps (IRS), and they are also often used as an underlying of structured coupon for more sophisticated products.

The overnight rate OIS is the day to day rate that is charged for ongoing liquidity needs of banks. This rate plays a key role in presence of collateral agreements between two counterparts. It is often used as the rate to which collateral is remunerated.

1.1.2 Linear products

The main examples of linear interest rates derivatives are IRS and forward rates agreements (FRA). The flows of these products are linear functions of the reference rates and "theoretically" their price is determined by arbitrage arguments and is model independent. In the following we will discuss how the latest developments in the market have broken this reality.

A **FRA** is a contract involving three time instants: The current time t, the expiry time $T > t$, and the maturity time $S > T$. The contract gives its holder an interest-rate payment for the period between T and S. At the maturity S, a fixed payment based on a fixed rate K is exchanged against a floating payment based on the spot rate $L_T(S-T)$ resetting in T for the duration $S-T$. Basically, this contract allows one to lock-in the interest rate between times T and S at a desired value K, with the rates in the contract that are simply compounded. Formally, at time S one receives $(S-T)KN$ units of currency and pays the amount $(S-T)L_T(S-T)N$, where N is the contract nominal value. The value of the contract at date S is therefore

$$N(S-T)(K - L_T(S-T)), \tag{1.1}$$

where we have assumed that both legs have the same daycount convention. At date t the forward rate of maturity T and tenor τ, which we denote by $L_t(T,\tau)$, is the rate K that makes the FRA contract expiring at T and maturing at $T+\tau$ have value 0 at the contract initiation. We denote by $\mathrm{FRA}_t(T,\tau,K)$ the time t value of a FRA contract of nominal 1, expiry T and maturity $T+\tau$. FRA contracts are typically issued at value 0 at initiation and prices are quoted through forward rates.

An **IRS** is a contract involving a set of prespecified dates $T_0, T_1, ..., T_n$ that exchanges payments between two differently indexed legs. For a payer swap, at every instant $T_i, i = 1...n$ the fixed leg pays out the amount

$$N(T_i - T_{i-1})K \tag{1.2}$$

corresponding to a fixed interest rate K, a nominal value N, whereas the floating leg pays the amount

$$N(T_i - T_{i-1})L_{T_{i-1}}(T_i - T_{i-1}), \tag{1.3}$$

corresponding to the interest rate $L_{T_{i-1}}(T_i - T_{i-1})$ resetting at the previous instant T_{i-1} for the maturity given by the current payment instant T_i, with T_0 a given date. Again we have assumed here that both legs have the same daycount convention and pay on the same set of dates. We denote by $\mathrm{IRS}_t((T_i)_{i=0}^n, K)$ the time t price of the IRS of payment dates $(T_i)_{i=0}^n$ and fixed rate K. At date T_0, the swap rate for the set of dates $(T_i)_{i=0}^n$, which we denote by $S_{T_0}((T_i)_{i=0}^n)$ is the fixed rate K that makes the IRS contract have value 0 at contract initiation. Standard IRS contracts are initiated at date T_0, but one can also chose to initiate the contract at a date $t < T_0$ in which case the contract is called a forward swap contract (FS) and we denote by $S_t((T_i)_{i=0}^n)$ the date t forward swap rate for the set of dates $(T_i)_{i=0}^n$. Typical IRS and FS payment dates are separated by a fixed tenor $\tau = T_i - T_{i-1}, i = 1, ..., n$, usually 6 months for the Euro IRS market and 3 months for the USD IRS market, we will denote by $S_t(T_0, n, \tau)$ the associated forward swap rate.

1.1.3 Non linearity episode 1: optionality

Non linearity arises when ever the the payoff of a product is a non-linear function of the reference rates. The most common examples of non-linear products are optional products such as swaptions and caplets.

The caplets and cap market

A caplet is an option to enter into a FRA of fixed expiry, maturity and rate. The option can be exercised at expiry of the FRA, precisely the payoff of a caplet at expiry is given by

$$N\left(\mathrm{FRA}_T(T, \tau, K)\right)^+, \tag{1.4}$$

corresponding to a fixed rate K and a nominal N. The fixed rate K is called the strike price of the caplet. We denote by $C_t(T, \tau, K)$ the time t value of the caplet price of maturity T, tenor τ and strike K. Caplets are typically quoted for different strike prices. When the strike price equals the value of the forward rate $L_t(T, \tau)$ we say the caplet is at-the-money. Brokers typically give market quotes for caplets of different maturities, tenors and strike prices. Typically quotes will be available for maturities ranging from 1 to 30 years, and tenors from 3 months to 1 year and a set of strike prices around the forward rate. The closer the strike price to the forward rate, and the more the product will be liquid.

A cap is a collection of caplets of same tenor and strike price of different maturities. Precisely at the maturity dates of the cap $T_1, ..., T_M$ we have the option to exercise a caplet of tenor τ and strike K, and thus receive the payoff

$$N\left(\mathrm{FRA}_{T_i}(T_i, \tau, K)\right)^+ . \tag{1.5}$$

The swaption market

A standard European swaption is an option to enter into a fixed versus floating forward starting interest rate swap at a predetermined rate on the fixed leg. A receiver swaption gives the right to enter a swap, receiving the fixed leg and paying the floating leg, while a payer swaption gives the right to enter a swap, paying the fixed leg and receiving the floating leg[1].

At maturity the payoff of the swaption giving the right to enter in the IRS of payment dates $(T_i)_{i=0}^n$ at the fixed rate K is given by

$$N(\mathrm{IRS}_T((T_i)_{i=0}^n, K))^+. \tag{1.6}$$

corresponding to a nominal N. We will denote its price by $SP_t(T, (T_i)_{i=0}^n, K)$. Typical interest rate swaptions give the right to enter in a swap with payment dates separated by a fixed tenor $\tau = T_i - T_{i-1}, i = 1, ..., n$, we will denote by $SP_t(T, T_0, n, \tau, K)$ the time t value of the swaption. The contract can also allow to enter into a forward start swap, in which case we speak of forward starting swaption. The difference $T_n - T_0$ is called the tenor of the swaption. The maturity T is usually equal to the IRS starting date T_0, in which case we will simply denote the swaption value by $SP_t(T_0, n, \tau, K)$ or $SP_t((T_i)_{i=0}^n, K)$. When the strike price is equal to the forward swap rate $S_t((T_i)_{i=0}^n)$ or $S_t(T_0, n, \tau)$ the swaption is said to be at-the-money. Brokers typically give market quotes for swaptions of different maturities, tenors and strike prices. Typically quotes will be available for maturities ranging from 1 to 30 years, tenors from 2 to 30 years and a set of strike prices around the froward rate.

The swaption contract we have just defined gives the right to enter in a swap at expiry. The contract is similar to call and put options in the equity markets, which give the right to buy or sell the underlying asset. This type of contract is called a physical settlement contract, because, at maturity, it gives the right to enter into a (physical) contract (an IRS or a number of shares). In today's markets options are often used for speculation purposes and the buyers are not necessarily interested in the physical underlying, but rather in the payoff of the option. For this reason, in the equity market, options are often cash settled, meaning that, if the option is exercised at maturity the holder receives an amount of cash which corresponds to the gain of the option. In the interest rates market the situation is

[1]USD swaps exchange a fixed rate for a floating 3-month LIBOR rate, with fixed-leg payments made semi-annually and floating-leg payments made quarterly. EUR swaps exchange a fixed rate for a floating 6-month EURIBOR rate, with fixed-leg payments made annually and floating-leg payments made semi-annually. In both currencies, the daycount convention is 30/360 on the fixed leg and Actual/360 on the floating leg.

similar, and very often swaptions are not physically settled. In the US market, the physical settlement is the market standard. In the EU market, the cash yield settlement is the market standard. We will introduce the different types of settlement later in the following.

1.2 Recent developments

The interbank market has known recent developments since the credit crunch. The treasury desk acquired a crucial role in the business. Credit and counterparty risks have arisen as the major risks in the interbank market. Funding costs which had previously been neglected in the pricing and management of derivatives are now essential.

1.2.1 The IBOR-OIS and other basis

One of the main facts of the liquidity crisis that succeeded the credit crunch of 2008 is the significant basis between the IBOR rate and OIS rate. The number of papers trying to provide an explanation for such a basis in recent years is impressive, see [Mor09], [Mor11] to cite just a few. Two are the main reasons for such a basis: on one hand the IBOR rate is now pricing a credit risk of the counterparts belonging to the IBOR panel, on the other hand it incorporates a liquidity risk component coming from the funding needs of banks. Most of the papers have been focusing on the credit risk component of the basis arguing it is its main component. Recently [DC13] has argued the exact opposite, claiming that the liquidity risk is today the predominant component of this spread. The basis would mainly be explained by the funding constraints that banks have been facing since the credit crunch of 2008. The underlying idea is that it is easier to fund an overnight flow than to fund an in fine flow. Let us note that a basis between these two rates has always existed, but it was negligible in the pre-crisis period.

Simultaneously we have observed the widening of another basis between the FRA and its static replication strategy. For a detailed discussion on the occurrence of these two basis we refer to [Mor09]. The most widely accepted explanation for such a basis is that the FRA is a collateralized while the static replication is not. As a consequence the FRA price is lower since the existence of a collateral agreement allows on one hand to mitigate the credit risk on the counterparty of a deal, on the other hand to deal with an overnight funding of the deal.

In fact the IBOR-OIS basis is only one example of basis which is varying much more than it used to do before the 2008 credit crunch. The after crisis dynamics of the 3M-6M IBOR basis, the different cross-currency basis, has forced banks to take all these basis into account when pricing derivative contracts. The first step consists in using different discount curves for different contracts, the second step consists in building a stochastic model for these basis. A derivative that could be priced with a simple one factor model before the crisis would require today a multiple factor model to account for all the different basis involved in the pricing of the trade.

1.2.2 The impact of collateralization in the derivative pricing

Since the credit crunch, collateral agreements (Credit Support Annexes, CSA hereafter) are regarded with much more interest by the quantitative finance community. The impact of collateralization is now taken into account when pricing derivatives. CSA are used to mitigate the credit risk of a derivative contract. Roughly speaking, collateral posting occurs whenever the movements in mark-to-market are higher than predetermined triggers (usually this will occur at a high frequency: daily or weekly). The collateral is remunerated with a predetermined collateral rate which depends on the counterparty and on the quality of the collaterlal. For cash-like collaterals this rate is often very close to the OIS rate. The main consequence of this is that the hedging strategy of a collateralized derivative contract is in fact funded by the collateral rate (which is the OIS rate) and thus the discounting of the payoff is to be made with the collateral rate rather than with the IBOR rate. Things get much more complicated when the CSA allows to post collateral in different currencies, in this case the collateral posting has a time dependent optional behavior which depends on the trajectory of the cheaper currency to deliver as collateral. Investment banks have now dedicated desks that optimize the collateral posting for deals, and the pricing of collateralized derivatives incorporates the impact of collateral and funding. Through a simplistic example, we will try to give a flavor of how the mechanics of collateralization impact the pricing of derivatives, we refer to [Pit10], [FT11] and [Mor11] for a detailed description on the subject. Let us consider a generic derivative contract written on a generic underlying S_t, and denote by V_t the premium of the derivative contract at time t, let us denote by c_t the instantaneous interest rate at which the collateral is remunerated, f_t the instantaneous interest rate at which the treasury desk has agreed to fund this trade and m_t instantaneous interest rate of a repo contract on S_t. Let us take the perspective of the derivative seller. Assuming that the contract is fully collateralized in cash with a bilateral collateral agreement, the cash flows at time of the transaction are:

1. We receive the derivative premium V_t.

2. We post the amount V_t in cash to the counterparty.

3. We borrow $\Delta_t S_t$ from our treasury desk.

4. We buy a quantity Δ_t of the underlying asset S_t.

5. We enter a repo contract, where we lend the asset to a third counterparty.

The flows 3 to 5 could be replaced by opposite flows assuming we enter in a reverse repo contract, where we would receive the underlying asset and pay the interests for this. At time $t + dt$, the flows of the contract are:

1. We receive the interests on the posted collateral $V_t c_t dt$.

2. We pay the interests on the delta position $\Delta_t S_t f_t dt$.

3. We receive interests on the repo contract $\Delta_t S_t m_t dt$.

Assuming the market is complete and the hedging strategy is self financed, the modified self-financing equation reads

$$V_{t+dt} - V_t = c_t V_t dt - \Delta_t S_t \left(f_t - m_t\right) dt + \Delta_t \left(S_{t+dt} - S_t\right). \tag{1.7}$$

This equation should be compared with the classic self-financing equation below

$$V_{t+dt} - V_t = r_t \left(V_t - \Delta_t S_t\right) dt + \Delta_t \left(S_{t+dt} - S_t\right), \tag{1.8}$$

where r_t denotes the instantaneous "risk-free" rate at which we can borrow and lend money. Equation (1.8) is crucial to be able to represent prices as expectation of discounted cash flows, with an instantaneous discount rate equal to r_t. One of the problem when including the different costs of collateral remuneration and funding is that we break the assumption that the same rate applies for borrowing and lending money, and the cash part of the portfolio evolves at different rates. If we assume that $c_t = f_t - m_t$, then the self-financing equation including collateral costs (1.7) is equivalent to the classic self-financing equation (1.8) with $r_t = c_t$, which leads to represent the derivatives prices as follows:

$$V_t = \mathbb{E}\left[e^{-\int_t^T dsc_s} V_T | \mathcal{F}_t\right], \tag{1.9}$$

if the option is European, the payoff at maturity M is known and we have $V_M = f(S_M)$. Things get much more complicated if we incorporate triggers on the variation of the derivative premium which determine wether or not collateral is posted, partially collateralized trades, unilateral collateral agreement, or the possibility of posting collateral in a different currency. All these features of the CSA create non-linearity in the differential pricing operator and are hard to include in the pricing of a derivative contract. See for example [BP14] and [PB13].

1.2.3 Which discounting curve?

There are two ways to compute (1.9). The first is to build a model for both the underlying asset S_t and the instantaneous collateral rate c. These models should be calibrated in such a way that they reproduce the prices of the relevant vanilla instruments, and in particular the instruments which have the same CSA. The second is to decompose the expectation by making a change of measure,

$$V_t = P^c_{t,M} \mathbb{E}^{c,M} \left[f(S_M) | \mathcal{F}_t\right], \tag{1.10}$$

where $P^c_{t,M}$ is the discount factor set at time t for maturity M associated to the instantaneous collateral rate c, and $\mathbb{E}^{c,M}$ is the measure associated with the numeraire $(P^c_{t,M})_t$. The main question is then the following: "How do we build the discount curve $(P^c_{t,T}, t \leq T)$?". As mentioned before, the collateral rate to apply in the pricing of a derivative contract will depend on the CSA. In practice it is necessary to have multiple curves which will be applied to discount derivatives with different CSA agreements. Also, let us note that the choice of the curve is not solely driven by the CSA of the derivative, another element which must be taken into account is the frequency of the cash flows of the derivative[2]. The construction of the different discount curves would be straightforward if there was a market of liquidly traded IRS for every CSA agreement and every payment frequency. Strictly speaking discount curves used for contracts with a CSA that allows to post collateral in different currencies contain optionality and should be deduced from an optimal collateral posting strategy. Unfortunately, this information is not available in the market, and the instruments which have sufficient liquidity and observability to be used for curve construction are standard IRS[3], which are collateralized, and basis swaps between the different references: OIS-IBOR basis, 3M-6M IBOR basis, Cross-Currency basis...Typically the construction of the curve consists in choosing a reference discount curve (for example the OIS discount curve or the 3 month IBOR curve) and then sequentially build all the curves based on the information available on the standard IRS and on the basis swaps. Even using the (simplistic) formulas (1.9) and (1.10), the price of a future discounted cash flow indexed on a IBOR reference rate would require some convexity and in particular depends on the volatility. In fact the building of the curves cannot be considered as model independent anymore, and is based on some hidden assumptions (typically some assumption on the form of the convexity adjustment).

Typical interest rates derivatives have a payoff indexed on a IBOR rate of maturity δ, and are fully collateralized with cash equivalent collateral. As a consequence the price of collateralized trades depends on two variables, the OIS rate and the δ-IBOR rate. Precisely we need to model the dynamics of two different yield curves, the δ-IBOR curve used to determine the forward rates and the forwards of the payoffs, and the OIS curve to discount the payoff. The simplest approach consists in assuming that the IBOR curve and the basis are independent, in this case one can compute the price easily. Let us consider the price of a European option with payoff $f(y^{IBOR}_{T,\tau}, \tau \geq 0)$ at maturity, where $(y^{IBOR}_{t,\tau}, \tau \geq 0)$ is the IBOR yield curve at date t. We have

$$\begin{aligned} P_t &= \mathbb{E}\left[e^{-\int_t^T dsr_s^{OIS}} f\left(y^{IBOR}_{T,\tau}, \tau \geq 0\right)\right] \\ &= \mathbb{E}\left[e^{-\int_t^T dsr_s^{OIS} - r^{IBOR_s}} e^{-\int_t^T dsr_s^{IBOR}} f(y^{IBOR}_{T,\tau}, \tau \geq 0)\right] \\ &= \mathbb{E}\left[e^{-\int_t^T dsr_s^{OIS} - r^{IBOR_s}}\right] \mathbb{E}\left[e^{-\int_t^T dsr_s^{IBOR}} f\left(y^{IBOR}_{T,\tau}, \tau \geq 0\right)\right]. \end{aligned} \tag{1.11}$$

[2]This is because the basis between the multiple discount curves are mainly driven by credit and liquidity risk.

[3]The characteristics of standard IRS depend on the market currency.

The term $\mathbb{E}\left[e^{-\int_t^T dsr_s^{IBOR}} f(y_{T,\tau}^{IBOR}, \tau \geq 0)\right]$ is the standard price of the option in a mono-curve framework. The term $\mathbb{E}\left[e^{-\int_t^T dsr_s^{OIS} - r^{IBOR_s}}\right]$ is an adjustment that accounts for the spread between the IBOR and OIS rate, in practice this quantity is represented by a spread curve $(s_{t,\tau}, \tau \geq 0)$ defined by

$$e^{-\tau s_{t,\tau}} = \mathbb{E}\left[e^{-\int_t^{t+\tau} dsr_s^{OIS} - r^{IBOR_s}}\right].$$

To the best of our knowledge there is not a market consensus on the way the multi-curve setting is handled. Exploiting the orthogonality in the expression of the price (1.11), some banks have structured their trading desks in such a way that one desk continues to hedge in a mono-curve setting (only focusing on the expression $\mathbb{E}\left[e^{-\int_t^T dsr_s^{IBOR}} f(y_{T,\tau}^{IBOR}, \tau \geq 0)\right]$), and one desk manages the collateral posting and the funding basis. Of course the activities of these two desks are interconnected. Our external and very personal viewpoint on the subject is that the separation of the price (1.11) is not fully justified. A unified view on the price is recommended even if we have a separated desk managing collateral and funding. The difference between the price in a mono-curve framework and the price in a bi-curve framework is called the funding value adjustment (FVA).

1.2.4 Price adjustments

Under the impulse of recent market evolutions and regulation, the pricing of derivatives requires a series of adjustments. As explained before the FVA is the adjustment of the price one has make to take into account the actual funding cost of the hedging strategy of the derivative. Other major adjustments are the credit value adjustment (CVA) and the liquidity value adjustment (LVA). Basel 3 regulation requires that banks are able to compute and exhibit these 3 adjustments for their derivatives contracts.

CVA is the adjustment of the derivative's price to take into account the market cost of hedging the credit risk of the counterparty of this derivatives. The computation of CVA requires the computation of the expected positive exposure of the derivative and the market quote of the default probability of the counterparty. Common market practice consists in making the simplifying assumption that the default of the probability and the mark-to-market (MtM) of the derivative are independent which allows to compute the CVA in a separated form of type $\int_t^T dsEPE_s \times PD_s$, where EPE_s is the expected positive exposure at date s and PD_s is the default probability of the counterparty at date s. The computation of CVA is made on an aggregate level for each counterparty. The adjustment for an individual derivative contract thus depends on the whole portfolio of transactions with the counterparty of the contract.

LVA is the adjustment of the derivative's price to take into account the liquidity costs of the derivative.

As for FVA, the practices change from banks to banks for CVA and LVA. A common trend is however to have separated desks managing CVA, FVA and LVA on an aggregate basis. Working on an aggregate basis has the obvious virtu of optimizing the costs that would be generated by managing the adjustments individually. However this requires that we are able to go from an aggregate measure (at a counterparty level for instance) of the adjustments to an individual (at a transaction level for instance) measure of the adjustment and the other way around. Of course the difficulties come from the fact the adjustments are highly nonlinear and aggregation is not immediate. The pricing of an individual deal cannot be made separately as it used to be in the past, one needs to measure its impact on the whole portfolio of the counterparty and eventually compute the marginal cost of capital of this particular deal for the bank to determine wether it is interesting or not.

1.3 Valuation principles and quotation conventions: pre and post 2008 crisis

Let us now discuss the valuation of vanilla interest rates vanilla products. The milestone for the valuation of vanilla product is the notion of zero-coupon bond. The time t value of a zero-coupon bond of maturity T denoted by $P_{t,T}$ is the value of a payoff of 1 at maturity T. Zero-coupon bonds are the primary assets of the interest rates market. In a (pre-crisis) mono-curve setting there was no ambiguity, one could read a zero-coupon bond price on the unique discount curve available. In a multi-curve setting, one must choose the curve to use, depending on the CSA of the derivative and on the reference IBOR rate in the derivative payoff.

In a mono-curve setting, the prices of linear products can be entirely expressed in terms of zero-coupon bonds of a unique curve

$$\mathrm{FRA}_t(T,\tau,K) = N\tau P_{t,T+\tau}\left(K - L_t(T,\tau)\right) \tag{1.12}$$

$$\mathrm{IRS}_t((T_i)_{i=0}^n, K) = N\left(P_{t,T_0} - P_{t,T_n} - K\sum_{i=1}^n (T_i - T_{i-1})P_{t,T_i}\right). \tag{1.13}$$

This allows us to express the forward libor and swap rate

$$L_t(T,\tau) = \frac{1}{\tau}\left(\frac{P_{t,T}}{P_{t,T+\tau}} - 1\right) \tag{1.14}$$

$$S_t((T_i)_{i=0}^n) = \frac{P_{t,T_0} - P_{t,T_n}}{\sum_{i=1}^n (T_i - T_{i-1})P_{t,T_i}}. \tag{1.15}$$

In a multi-curve setting, the price of linear products is given by

$$\text{FRA}_t(T,\tau,K) = N\tau P^c_{t,T+\tau}\left(K - \mathbb{E}^{c,T+\tau}_t\left[L_T(T,\tau)\right]\right) \quad (1.16)$$

$$\text{IRS}_t((T_i)^n_{i=0},K) = N\left(\sum_{i=1}^{n}(T_i - T_{i-1})P^c_{t,T_i}\left(\mathbb{E}^{c,T_i}_t\left[L_{T_{i-1}}(T_{i-1},T_i - T_{i-1})\right] - K\right)\right) \quad (1.17)$$

where $\mathbb{E}^{c,M}_t$ denotes the expectation under the numeraire $(P^c_{t,M})_t$. The forward libor rate is then defined by

$$L_t(T,\tau) = \mathbb{E}^{c,T+\tau}_t\left[L_T(T,\tau)\right]. \quad (1.18)$$

Let us note that by definition the forward libor rate is a martingale. By construction, the curve associated with the IBOR rate of maturity τ will be such that the forward libor rate can be expressed as a function of the zero-coupon bonds of that curve:

$$L_t(T,\tau) = \frac{1}{\tau}\left(\frac{P^{IBOR(\tau)}_{t,T}}{P^{IBOR(\tau)}_{t,T+\tau}} - 1\right). \quad (1.19)$$

An alternative way to represent forward libor rate is to define the curve of forward libor spreads between the reference curve and the curve associated with the IBOR of maturity τ. Let us assume that $P^c_{t,T}$ is the reference curve, then the forward libor rate is defined as the classic forward libor rate (1.16) plus a spread.

$$L_t(T,\tau) = \frac{1}{\tau}\left(\frac{P^c_{t,T}}{P^c_{t,T+\tau}} - 1\right) + sp_t(T,\tau). \quad (1.20)$$

where the curve $(S_t(T,\tau), t \leq T)$ is the curve of forward spreads between the reference collateral curve and the τ-libor rate curve. Similarly, we can derive the expression of the swap rate. Assuming that the swap has fixed payment frequency $T_i - T_{i-1} = \tau$, we have

$$S_t((T_i)^n_{i=0}) = \frac{P^{IBOR(\tau)}_{t,T_0} - P^{IBOR(\tau)}_{t,T_n}}{\tau\sum_{i=1}^{n} P^c_{t,T_i}}. \quad (1.21)$$

Again, this could be expressed as the classic swap rate plus a spread. The crucial point is whether or not the spread between the reference curve and the base curve is modeled as a stochastic variable. While the market evidence suggests that the different basis are stochastic and their variations in the after crisis environment are sizeable, it is very rare among practitioners to include a stochastic dynamics for these different basis. One of the great difficulties of modeling the spreads as stochastic variables is the lack of instruments to which we can calibrate. In particular, there are no liquidly traded options on the different basis.

The fundamental idea of arbitrage free pricing is that there exists a martingale measure, which is equivalent to the objective or historical measure under which the discounted (by whichever collateral rate) asset prices are martingales. This gives a tool to compute the prices of european contigent claims as expectation of the discounted payoff at maturity. The second fundamental concept is the concept of numeraire [GREK95], which is extensively used in the interest rates market. For a specific detailed description of the valuation of contingent claims in the interest rates market we refer to [BM06], [PA10]. The mathematical consequences of the fundamental results of financial contingent claims valuation theory will be developed in 2.1.

The way interest rates vanilla options prices are quoted is inherited from the conventions of the equity market. In particular caplets and swaptions are often viewed as call options respectively on the forward libor rate and on the forward swap rate. From the arbitrage pricing theory we have

$$\begin{aligned} C_t(T,\tau,K) &= \mathbb{E}\left[e^{-\int_t^T dsc_s}\left(\mathrm{FRA}_T(T,\tau,K)\right)^+|\mathcal{F}_t\right] \\ SP_t(T,(T_i)_{i=0}^n,K) &= \mathbb{E}\left[e^{-\int_t^{T_0} dsc_s}\left(\mathrm{IRS}_T((T_i)_{i=0}^n,K)\right)^+|\mathcal{F}_t\right]. \end{aligned}$$

The market convention is to apply a standard change of numeraire technique [GREK95] and rewrite the above expressions as

$$C_t(T,\tau,K) = P^c_{t,T+\tau}\mathbb{E}^{T+\tau}\left[(L_T(T,\tau)-K)^+|\mathcal{F}_t\right] \tag{1.22}$$

$$SP_t(T,(T_i)_{i=0}^n,K) = \sum_{i=1}^n (T_i-T_{i-1})P^c_{t,T_i}\mathbb{E}^{A((T_i)_{i=0}^n)}\left[(S_{T_0}((T_i)_{i=0}^n)-K)^+|\mathcal{F}_t\right]. \tag{1.23}$$

where $\mathbb{E}^{T+\delta}$ and $\mathbb{E}^{A((T_i)_{i=0}^n)}$ denote respectively the expectation taken w.r.t. the measure associated with the numeraire $P^c_{t,T+\tau}$ and $\sum_{i=1}^n(T_i-T_{i-1})P^c_{t,T_i}$. The quantity $\sum_{i=1}^n(T_i-T_{i-1})P^c_{t,T_i}$ is called the annuity associated to the $\mathrm{IRS}_t((T_i)_{i=0}^n,K)$, and will be denoted by $A_t((T_i)_{i=0}^n)$ in the following. Brokers typically quote the prices of caplets and swaptions in terms of the implied volatility respectively of the forward libor and swap rate. In practice we observe a volatility cube of implied volatilities of caplets and swaptions for different maturities, tenors and strikes.

1.3.1 Non-linearity episode 2: hidden optionality

Sometimes non-linearity is hidden in an apparently linear product, this is the case of libor in arrears and constant maturity swap (CMS in the following). A libor-in-arrears swap is a contract similar to a standard swap rate involving the exchange of a fixed rate against a floating rate, but the rate is payed at the same date of its fixing, precisely for a set of prespecified dates $T_0,T_1,...,T_n$, at every instant $T_i, i=1...n$ the fixed leg pays out the amount

$$N(T_i-T_{i-1})K \tag{1.24}$$

corresponding to a fixed interest rate K, a nominal value N, whereas the floating leg pays the amount

$$N(T_i - T_{i-1})L_{T_i}(T_{i+1} - T_i), \tag{1.25}$$

corresponding to the interest rate $L_{T_i}(T_{i+1} - T_i)$ resetting at the instant T_i for the maturity given by the next payment instant T_{i+1}, with T_0 a given date. While this seems a linear product it in fact hides optionality. The time t value of the floating leg is given by

$$N(T_i - T_{i-1})\mathbb{E}\left[e^{-\int_t^{T_i} ds c_s} L_{T_i}(T_{i+1} - T_i)|\mathcal{F}_t\right].$$

By standard change of measure arguments we have

$$N(T_i - T_{i-1})P^c_{t,T_{i+1}}\mathbb{E}^{T_{i+1}}\left[\frac{1}{P^c_{T_i,T_{i+1}}} L_{T_i}(T_{i+1} - T_i)|\mathcal{F}_t\right],$$

where $\mathbb{E}^{T_{i+1}}$ denotes the expectation under the martingale measure associated to the numeraire $P_{t,T_{i+1}}$. Using the expression of the forward rate we get

$$\mathbb{E}^{T_{i+1}}\left[\frac{1}{P^c_{T_i,T_{i+1}}} L_{T_i}(T_{i+1} - T_i)|\mathcal{F}_t\right] =$$

$$\begin{aligned}
&= \mathbb{E}^{T_{i+1}}\left[(1 + (T_{i+1} - T_i)(L_{T_i}(T_{i+1} - T_i) - sp_{T_i}(T_i, T_{i+1} - T_i)))L_{T_i}(T_{i+1} - T_i)|\mathcal{F}_t\right] \\
&= L_t(T_i, T_{i+1} - T_i) + (T_{i+1} - T_i)\mathbb{E}^{T_{i+1}}\left[L_{T_i}(T_{i+1} - T_i)((L_{T_i}(T_{i+1} - T_i) - sp_{T_i}(T_i, T_{i+1} - T_i)))|\mathcal{F}_t\right].
\end{aligned}$$

The last term involves a payoff which is a quadratic function of the libor rate, and a cross term between the libor and the basis spread, and thus is clearly a non linear product. By application of the static replication formula [CW02], the value of the libor-in-arrears swap can be obtained as a linear combination of caplets prices of different strikes.

1.3.2 Viewing swaption as basket options

It is standard market practice to view swaptions as basket options either on a set of forward libor rates, either on a set of ZC bonds. We present these two approaches since these will be key in building some of our pricing methods.

Swaptions as options on coupon bonds

We can rewrite the payoff of a swaption (1.6) as follows:

$$N(\mathrm{IRS}_T((T_i)_{i=0}^n, K))^+ = N\left(1 - P^{IBOR}_{T,T_n} - K\sum_{i=1}^n (T_i - T_{i-1})P^c_{T,T_i}\right)^+,$$

deducing that the price of a swaption can be represented by

$$SP_t(T,(T_i)_{i=0}^n,K) = \mathbb{E}\left[e^{-\int_t^{T_0} dsc_s}\left(1 - P^{IBOR}_{T_0,T_n} - K\sum_{i=1}^n (T_i - T_{i-1})P^c_{T_0,T_i}\right)^+ |\mathcal{F}_t\right] \quad (1.26)$$

$$= P^c_{t,T_0}\mathbb{E}^{T_0}\left[\left(1 - P^{IBOR}_{T_0,T_n} - K\sum_{i=1}^n (T_i - T_{i-1})P^c_{T_0,T_i}\right)^+ |\mathcal{F}_t\right] \quad (1.27)$$

which can be viewed as a put option of strike 1 on the coupon bond $\sum_{i=1}^n C_i P^c_{T,T_i} + P^{IBOR}_{T_0,T_n}$, defined by $C_i = (T_i - T_{i-1})K, i = 1..., n$. The swaption price is thus equivalent to the price of a basket option on the basket of ZC bonds $(P^c_{t,T_1}, ..., P^c_{t,T_n}, P^{IBOR}_{t,T_n})$. Let us note that this vision of the option is probably more relevant for someone who is used to the equity option market since it links the price of the option with the actual market underlyings, which are ZC bonds. As opposed the price vision given by (1.23) is artificial since the forward swap rate is not a proper underlying.

We will discuss in details in section 4.2 how this particular view of the swaption price can be used to derive efficient pricing methods in the context of affine term structure models.

Swaptions as basket options on forward IBOR rates

Let us note that the forward swap rate (1.15) can be written as a basket of forward IBOR rates:

$$S_t((T_i)_{i=0}^n) = \sum_{i=1}^n \omega_t^i L_t(T_{i-1}, T_i - T_{i-1}) \quad (1.28)$$

$$\omega_t^i = \frac{(T_i - T_{i-1})P^c_{t,T_i}}{\sum_{k=1}^n (T_k - T_{k-1})P^c_{t,T_k}}. \quad (1.29)$$

It is important to note that the weights (4.31) are stochastic and admit a complicated dynamics. It is standard practice among practitioners and academics to consider that the contribution of these weights to the diffusion term of the forward swap rate can be neglected[4]. The weights are assumed to be constant and set to their value at the pricing date (see [d'A03], [Pit09] for some examples). This procedure is often referred to in the literature as drift freezing and we shall do the same in this document. Starting from the swaption price expression (1.23) and assuming that the wights ω^i are constant and equal to their value at date t, we can rewrite the swaption price as

[4] To the best of our knowledge there have been very few attempts to quantify either theoretically or numerically this statement. In [d'A03] the author investigates the accuracy of the approximation for pricing swaptions in the log-normal BGM model, he shows that the approximation is less efficient for long maturities and long tenors.

$$SP_t(T,(T_i)_{i=0}^n,K) = A_t((T_i)_{i=0}^n)\mathbb{E}^{A((T_i)_{i=0}^n)}\left[\left(\sum_{i=1}^n \omega_t^i L_{T_0}(T_{i-1},T_i-T_{i-1}) - K\right)^+ |\mathcal{F}_t\right]. \tag{1.30}$$

Which is a basket option on the basket of forward libor rates $L_t(T_{i-1},T_i-T_{i-1})$. We will do an extensive use of the drift freezing method to derive efficient approximations of the price and the volatility of swaptions in the context of affine term structure models.

1.3.3 Different settlements

As mentioned before the market standard for swaptions is not necessarily the physical settlement. In particular, in the EU swaptions market, the standard is a cash yield settlement. A settlement type different from physical translates into a so called convexity adjustment.

The payoff of a cash yield swaption is given by

$$AM((T_i)_{i=0}^n, S_t((T_i)_{i=0}^n))\,(S_t((T_i)_{i=0}^n) - K)^+, \tag{1.31}$$

where

$$AM((T_i)_{i=0}^n, s) = \sum_{i=1}^n (T_i - T_{i-1})\frac{1}{\prod_{k=0}^i (1+(T_k-T_{k-1})s)}.$$

The time t price is thus given by

$$SP_t^{CY}(T,(T_i)_{i=0}^n,K) = \mathbb{E}\left[e^{-\int_t^T ds c_s} AM((T_i)_{i=0}^n, S_T((T_i)_{i=0}^n))\,(S_T((T_i)_{i=0}^n) - K)^+ |\mathcal{F}_t\right]. \tag{1.32}$$

The forward swap rate is not a martingale under the risk-neutral measure, and, in common term structure models, the distribution of the forward swap rate is not known under the risk-neutral measure. The forward swap rate is a martingale under the annuity measure, i.e. the measure associated to the numeraire $A((T_i)_{i=0}^n)$, and usually its distribution is known under this measure. Typically the pricing of cash yield swaption is made under this measure, and the change of measure expressed as a function of the forward swap rate. Precisely we have,

$$SP_t^{CY}(T,(T_i)_{i=0}^n,K) = A_t((T_i)_{i=0}^n)\mathbb{E}^A\left[\alpha(T,S_T((T_i)_{i=0}^n))\,(S_T((T_i)_{i=0}^n) - K)^+ |\mathcal{F}_t\right], \tag{1.33}$$

where α is defined by

$$\alpha(t,s) = \mathbb{E}^A\left[\frac{1}{A_t((T_i)_{i=0}^n)\prod_{k=0}^i(1+(T_k-T_{k-1})S_t((T_i)_{i=0}^n))}|S_t((T_i)_{i=0}^n) = s\right].$$

In order to compute the above conditional expectation, we need a modeling assumption on the joint distribution of the annuity and the forward swap rate. Typically, this will be the case in a term structure model. Let us furthermore note that, in several term structure models, the above expectation cannot be computed explicitly. See [PA10] for an exhaustive discussion on how to approximate the function α.

The computation of the different convexity adjustment coming from either CMS related features or different settlement types is outside the scope of this work. However, it is worth noting that in order to value and (more importantly) risk manage non-physical settlement swaptions it is not sufficient to assume a dynamics of the underlying forward swap rate, and we need a full term structure model, or alternatively, an assumption on the form of the function α.

1.4 Peculiarities

Compared to the equity market, the fixed income market has some important peculiarities. The market underlying is (at least in theory) infinite dimensional, it is the whole yield curve. In practice, we build the yield curve from the market quotes of FRA and IRS, which are available for a finite but rather large set of maturities and tenors. Naturally the market underlying is high dimensional.

1.4.1 Market representation of option prices

As mentioned before the way the market quotes and views the prices of vanilla options such as swaptions and caplets is inherited from the equity market. In particular options prices are quoted in terms of the implied volatility of forward libor and swap rates, these volatilities are determined by inverting the pricing formulas (1.22) and (1.23) w.r.t. either the Black-Scholes, either the Bachelier formula. Let us note that given that the forward rates are assimilated to bond returns it seems more consistent to invert the pricing formulas w.r.t. the Normal model than w.r.t. the log-normal model. Typical values for the implied Normal volatility are between 0.5 and 2 per cent, while the log-normal volatility varies between 20 and 200 per cent.

It follows directly from the pricing formulas (1.22) and (1.23) that each implied volatility is quoted w.r.t. a given different martingale measure, which is different for each couple of maturities and tenors. As a consequence one must pay attention when speaking about the maturity and tenor term structure of the swaption volatility cube, since these volatilities are not computed under the same martingale measure.

1.4.2 Sliding underlyings

Another fundamental difference with equities is that observed smiles cannot be directly translated into the actual hedging instruments that will be used to hedge the options. Neither the forward swap rate or the forward libor rate are traded assets and the smile must be reinterpreted in terms of the assets that are actually traded. The natural instruments to hedge a caplet (resp. a swaption) is a FRA (resp. a forward IRS) with the fixed leg paying the strike of the option. It is immediate to translate the behavior of the smile w.r.t. this contract. However, in practice such a contract can be non liquid and involve high liquidity margins, it is standard market practice to hedge options with standard IRS and FRA at par. In practice hedging is performed at an aggregate level on a book of derivatives[5], the hedging instruments change continuously and it is not immediate to translate the behavior of the smile in terms of the products contained in the hedging portfolio. The notion of backbone, which is key to understand dynamic hedging of an equity option, is not so clear in the interest rates option market. We will discuss hedging in details in section 2.7, and try to formalize the hedging market practice.

1.5 Some empirical facts about the interest rates market

Before we describe the theory of interest rates term structure modeling we give some empirical results describing the phenomenology of the interest rates market. For a comprehensive and accurate description of the phenomenology of the interest rates curve we refer to [BSC$^+$99]. In this work, we focus on the interest rates derivatives market and in particular on the IRS and swaption market. In [TS10] the authors conduct a detailed analysis of the swaption volatility cube and propose a two factors model to capture the dynamics of the volatility cube. The objective of our work is to propose an affine term structure model which is able to capture both the IRS and swaption volatility dynamics.

As mentioned before IRS and swaption are among the most liquidly traded instruments in the fixed income market. Nonetheless one should never forget that these products are traded in an OTC market and not all the prices are completely transparent. For instance, non par IRS and deep out of the money swaptions are not liquidly traded and their price incorporate high liquidity premiums. There is sufficient market consensus to rely on the market quotes of at par IRS and ATM swaptions.

Some of the most commonly used modeling frameworks for interest rates will be described in chapter 2. Theoretically, the underlying of the interest rates market is the whole yield curve (which is infinite dimensional). In practice, the market is perceived as discrete both in the expiry and in the maturity dimension. It is common to reason in terms of pillars of

[5] This is clearly the most cost efficient choice in order to reduce transaction cost and liquidity margins.

both the yield curve and the volatility. This chapter is dedicated to the empirical analysis of both the IRS and swaption market.

An interest rates model is judged for its ability to reproduce statically and dynamically what is observed in the market. On one hand, it is important that the model recovers (to a degree which is considered satisfactory) the market prices of liquidly traded instruments. On the other hand it is important that the model reproduces the dynamics of the interest prices as observed in the market. The present section focuses on the latter, i.e. the description of the dynamical properties of IRS and swaption prices.

1.5.1 PCA of the swaption volatility cube

In the interest rates market, the most liquidly traded instruments are IRS and swaptions. As mentioned before interest rates derivatives trade OTC and their prices are not always transparent. For IRS swaps, it is common market practice to trade systematically at par. For swaptions, out of the money options can be highly illiquid. For these reasons, we only focus on ATM swaptions and ATM IRS, since these are the only instruments for which there is sufficient market consensus to analyze their price dynamics.

The data set

We analyze a set of data in the period going from January 2006 to December 2012. We are grateful to the Fixed Income Quantitative Research team of C.A. CIB who kindly provided us with the data set. The data set comprises the implied normal ATM volatility and ATM forward swap rate of USD swaptions for expiries ranging from 1 month to 30 years and maturities ranging from 1 year to 30 years.

In theory the swaptions' ATM forward swaps should be entirely determined by the yield curve extracted from swap and IBOR rates. In practice, this is not the case due to a number of factors such as the interpolation/extrapolation of the curve, the basis between collateral discount curves and forward IBOR curves... As mentioned before the number of factors impacting the pricing of (even simple) derivative contracts has exponentially increased after the credit crunch, and therefore, the swaption's ATM swap rate can be significantly different from the value one would imply from the forward IBOR yield curve. Assuming that the market quotes for ATM swaptions are reliable, and represent a transparent market, our data set contains the most relevant information to analyze the dynamics of the USD vanilla interest rates market. As mentioned before the market conventions of the EU swaptions market make it more difficult to analyze the prices of swaptions directly. The convexity adjustments introduced in 1.3.3 are model dependent, and therefore the implied volatilities of swaptions cannot be extracted in a model-free way.

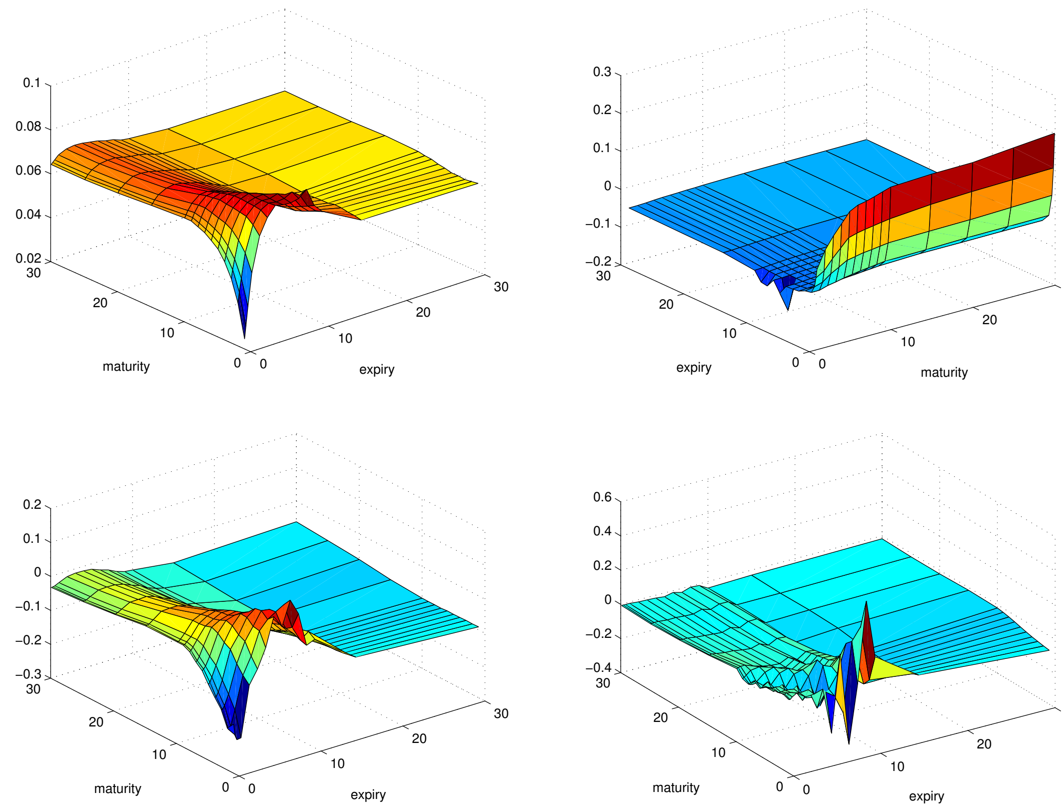

Figure 1.1: First four eigenmodes of the swaption ATM forward swap rates surface.

Results of the PCA

It is common to use PCA to analyze the dynamics of the yield curve. See [LS91]. Empirical results indicate that three factors lead to explain around 98% of the variance of the yield curve. We have performed a PCA on the weekly variations[6] of the swaptions' ATM forward swap rate surface. Figure 1.1 shows the first four factors of the PCA of the swaptions' ATM forward swap rates. The first factor explains 85 % of the variance, the first four factors explain more than 98% of the variance. The results are consistent with the literature. The first factor is similar to a parallel shift of the forward swap curve, the second factor to a change in the slope of the forward swap curve and the third to a change in the convexity of the curve. Let us note that we clearly observe different behaviors between long and short expiries and maturities. Figure 1.2 shows the factors of the PCA for a given expiry. The factors for short expiries are very similar to the one observed when performing a PCA on the yield curve itself. The factors for long expiries show different patterns. These seems to indicate that the IRS and swaptions' market is divided into sub-markets which have very different dynamics. This is consistent with the results of sparse PCA [dBEG07] to the IRS

[6]From wednesday to wednesday to avoid effects coming from the Friday close or the Monday opening.

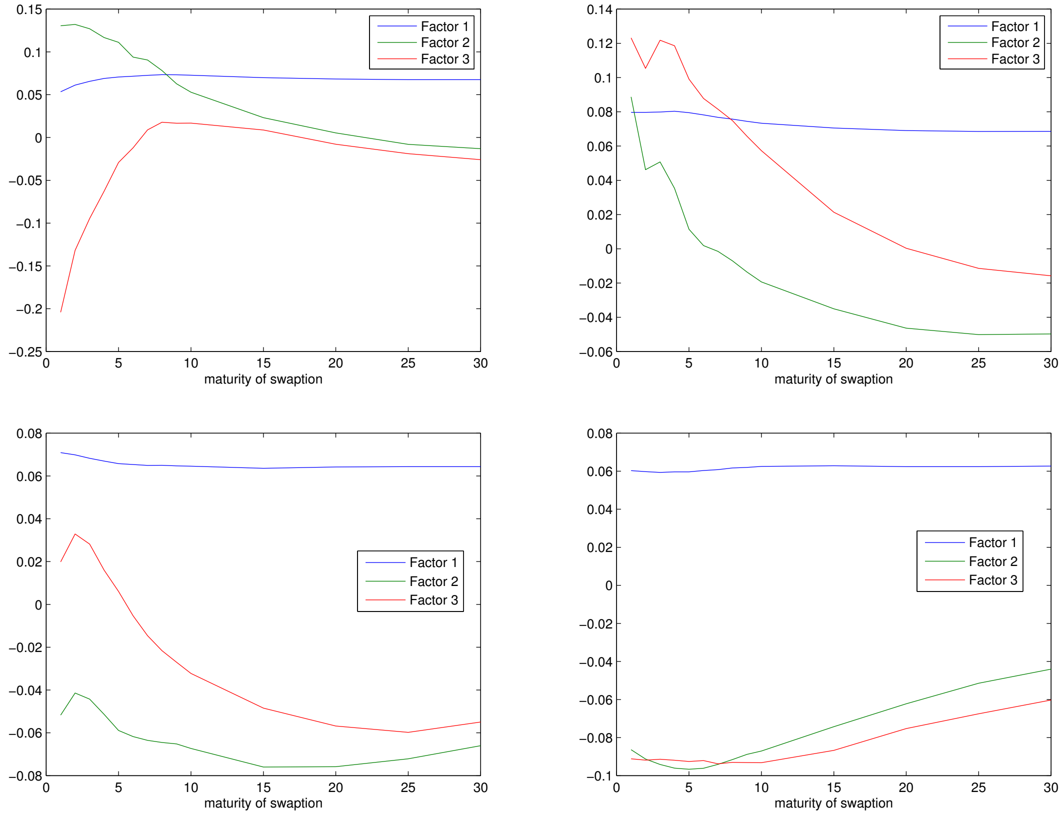

Figure 1.2: First three eigenmodes of the swaption's ATM forward swap maturity term structure.

market.

We have performed a PCA on the weekly variations[7] of the swaptions' ATM normal implied volatility surface. Figure 1.1 shows the first four factors of the PCA. The first factor explains 70% of the variance, the first two factors explain 82%, and the first four factors explain more than 93% of the variance. In [TS10] a PCA of the surface of swaptions' extracted *conditional variances* is performed, the data set ranges from 2001 to 2010, and the first factor explains more than 85% of the variations of the surface. The authors then suggest that the swaption cube can be modeled by a 2-factors stochastic volatility model. Let us mention that the concept of *conditional variance* in [TS10] is different from the ATM implied normal volatility and incorporates the information of the whole swaption smile.

The first and second factors of the PCA (see Figure 1.1) show symmetrical patterns. The shape of the factors indicate that the movements on the short expiries implied volatilities are

[7]From wednesday to wednesday to avoid effects coming from the Friday close or the Monday opening.

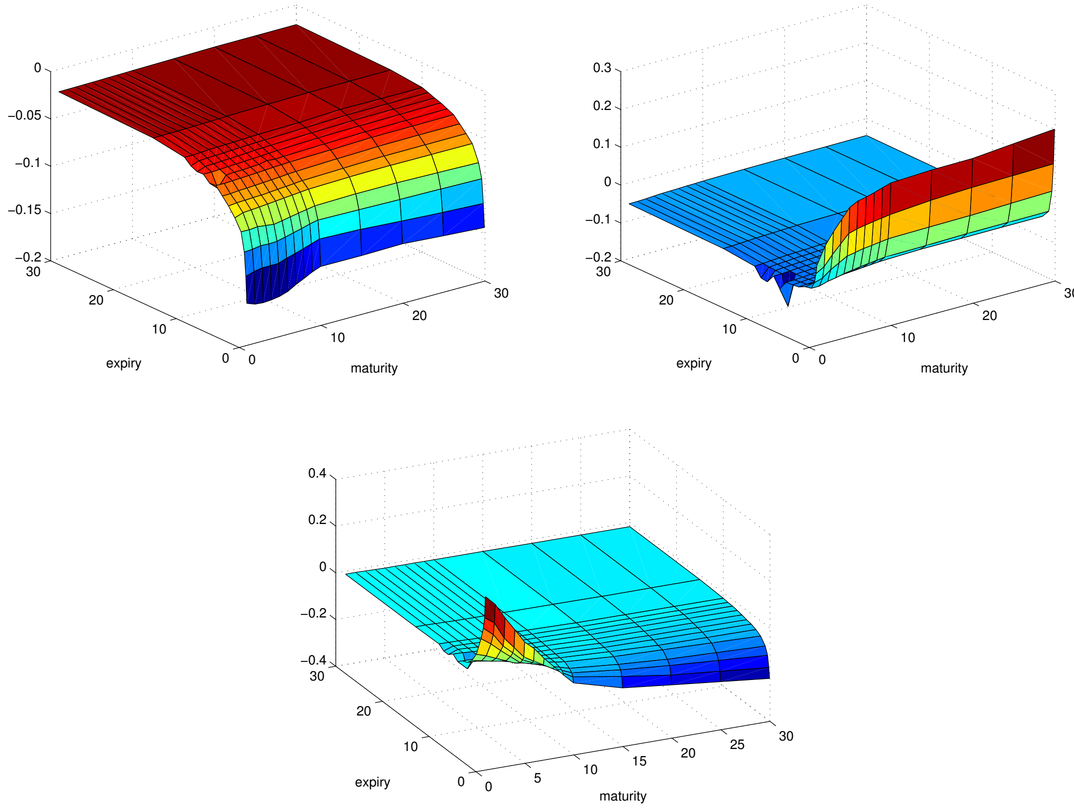

Figure 1.3: First three eigenmodes of the swaption's ATM normal volatility surface.

much more significant than the movements on the long term volatilities. The third factor is similar to a change in the slope of the maturity term structure of volatility for short expiry swaptions. Again, the dynamics of the long term expiry swaptions is very different from the dynamics of the short term expiry swaptions. The results of the PCA indicate that the volatility of swaptions of expiry longer than 10Y seem to have a simple dynamics, and move mainly by parallel shifts. This is confirmed when observing the time series of implied volatilities of different expiries, see Figure 1.4.

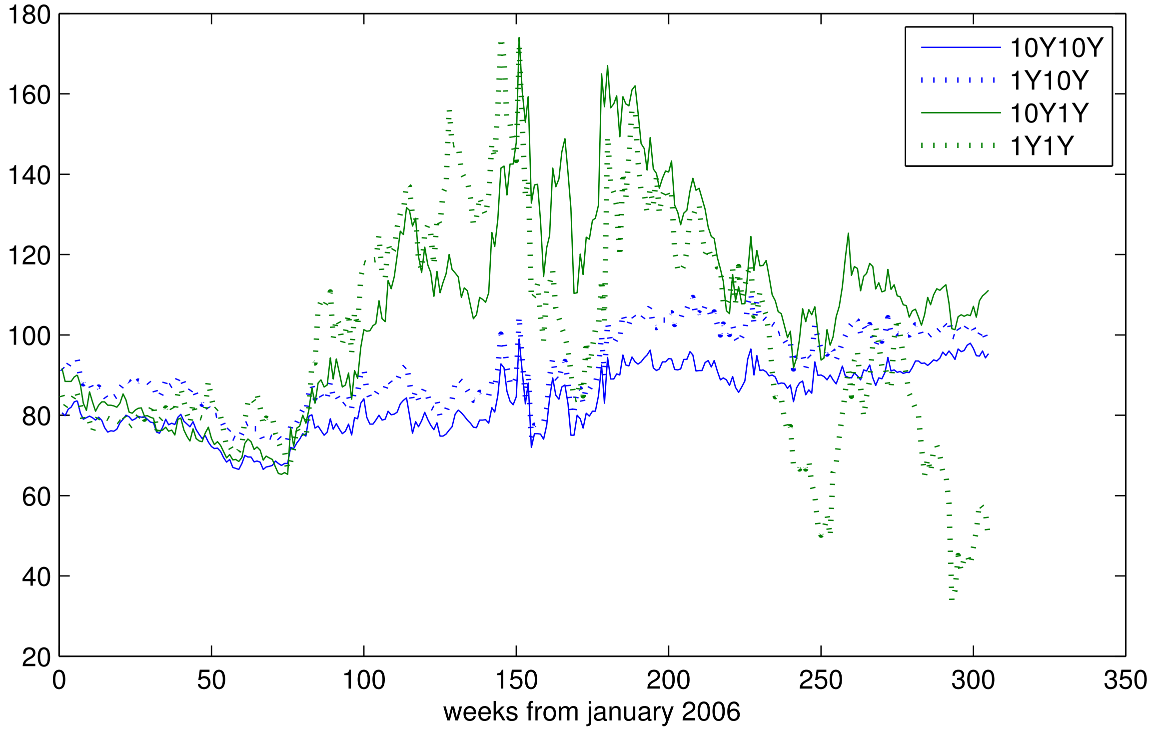

Figure 1.4: Time series of normal implied volatilities.

Chapter 2

Term structure models

In this chapter we proceed to a brief survey of interest rates models used to manage Interest Rates derivatives. We focus on markovian term structure models and in particular on the class of affine term structure models (ATSM). Rather than be exhaustive we focus on the results and properties of the models that will be useful in the following to build our modeling framework, for a much more complete survey of the term structure models used in the interest rates market we refer to [PA10], [BM06] and [DK96a]. We will present the models in a mono-curve setting. The class of ATSM allows to increase the number of factors easily, as soon as the dynamics of the additional factor is also affine. We conduce an analysis of the yield curve and implied volatility dynamics in markovian models through some examples. This allows us to highlight some of the concepts that will be developed in the following. First we show how the models can be specified in such a way that the role of the underlying state variables is clearly identified in terms of yield curve and implied volatility dynamics. Secondly we introduce the notion of "stochastic volatility" perturbation of the linear gaussian model. Finally, through a concrete example we discuss the performance of markovian models in terms of hedging.

Heath-Jarrow-Morton [HJM92] (denoted HJM hereafter) have developed a general framework that allows to build arbitrage-free models of the term structure of interests rates. The appeal of this framework lies in the fact it gives a very straightforward way of specifying an arbitrage-free term structure model and that the resulting model automatically fits the initial yield curve. Two are the main drawbacks of the general version of the model: first HJM model the dynamics of instantaneous forward rates which are not directly observable in the market, second the yield curve is not *markovian* in a finite number of state variables. As a consequence, in its most general version the HJM modeling framework lacks of tractability. Many of the modeling approaches proposed in the years 90's try to overcome these limits. Two are the main alternative modeling approaches: markovian TS models and market models.

Markovian TS models assume that the whole yield curve is driven by a finite (usually small) number of *state variables*, the dynamics of which is assumed to be a Markov process. This approach was pioneered by [EKLM+91], [EKL92] and [Che92]. Duffie and Kan [DK96b] defined a broad class of factorial TS models, [DS00] studied admissible specifications of the model. In all these models the functional mapping between the underlying variables and the yield curve is derived by combining the dynamical assumptions with the arbitrage free constraints. While this modeling approach overcomes the tractability problems of the HJM framework by explaining the dynamics of the curve by a finite set of variables, these variables are not directly observable in the market and hard to identify in general.

Market models radically differ from factorial models. Instead of using convenient but unobservable underlying variables, the approach pioneered by Brace, Gatarek and Musiela [BGM97] and extended by Jamshidian [Jam97], consists in taking rates that are actually traded (and thus quoted) in the market such as IBOR or Swap rates as underlying variables. While this approach solves the observability issues by modeling variables that are directly observable in the market, it conserves the tractability issues because the yield curve is not *markovian* in a finite number of state variables.

Let us note that the difference between these two approaches can be explained by the difference between the markets the authors wanted to model. On one hand Vasicek, Hull and White, Duffie, El Karoui et al. referred to markets which were characterized by a very liquid government bond market (such as the European and U.S. market). The government yield curve or equivalently the forward curve has been identified as the variable to be modeled. On the other hand BGM referred to markets which were characterized by the fact the government debt was of very "bad quality" and illiquid (such as the Australian market), while the swaps and IBOR market was liquid, and thus in this case the relevant market information was made of swap and IBOR rates.

Ahead of these models is the Markov-functional model [HKP00] proposed by Hunt and Kennedy (HK hereafter), which tries to solve the two main limits of the HJM framework simultaneously. The model can be viewed as the analogue of Dupire's local volatility model [Dup94] for interest rates. The curve is assumed to be *markovian* in a finite number of state variables and the mapping between the variables and the curve is directly inferred by the market quotes of vanilla instruments.

The chapter is organized as follows, section 2.1 recalls the general arbitrage-free framework for interest rates defined by Heath, Jarrow and Morton in [HJM92]. Section 2.2 presents factorial interest rates models. Section 2.3 presents the class of ATSM. Section 2.4 presents some examples of popular ATSM, for each model we discuss the yield curve dynamics and implied volatility cube. Section 2.5 presents MF models and formalizes the calibration procedure of Hunt and Kennedy [HKP00]. Finally section 2.6 discusses the nature of volatility in markovian TS models. Through a detailed pricing example using MF model we show the consequence in terms of hedging forward volatility products and introduce the notion of model risk.

2.1 Arbitrage free framework

We introduce the definitions and notations that will be used through this chapter. In all the following we consider a continuous trading economy with a trading interval $[0, T^*]$ for a fixed $T^* > 0$. The uncertainty in the economy is characterized by a probability space $(\Omega, \mathcal{F}, P)$. Information evolves according to the augmented filtration $\{\mathcal{F}_t\}_{t\in[0,T^*]}$ generated by a d-dimensional standard Brownian motion Z. We assume that a continuum of market zero-coupon bonds (ZCB) trade with different maturities, one for each trading date $T \in [0, T^*]$, and denote by $P_{t,T}$ the time t price of the T maturity bond for all $T \in [t, T^*]$ and $t \in [0, T^*]$. We assume that $P_{t,T} > 0$ and that $\partial_T \log P_{t,T}$ exists for all $t \in [0, T^*]$ and $T \in [t, T^*]$. The instantaneous forward rate at time t for date T, $f_{t,T}$ is defined by

$$f_{t,T} = -\partial_T \log P_{t,T}, \quad \text{for} \quad t \in [0, T^*], \quad T \in [t, T^*].$$

The spot rate or short rate r_t is defined as the limit

$$\lim_{T\to t} f_{t,T} = r_t.$$

The key results of Harrison and Kreps [HK79] and Harrison and Pliska [HP81] are the established connection between the absence of arbitrage opportunities and the existence of an equivalent martingale measure under which the discounted asset prices are martingales. Let us introduce the money market account by

$$S_t^0 = \exp\left(\int_0^t ds r_s\right).$$

Assume a general dynamics under the objective probability measure for the instantaneous forward rates

$$f_{t,T} = f(0,T) + \int_0^t ds\alpha_{s,T} + \int_0^t \sigma_{s,T}^T dZ_s, \tag{2.1}$$

where $\alpha_{t,s}$ and $\sigma_{t,T}$ are jointly measurable and adapted and verify

$$\begin{aligned} \int_0^T ds|\alpha_{s,T}| &< \infty \quad P-a.e. \\ \int_0^T ds|\sigma_{s,T}|^2 &< \infty \quad P-a.e. \end{aligned}$$

2.1.1 Arbitrage-free constraints

The following well known result defines the HJM arbitrage free framework [HJM92].

Proposition 1 —*(HJM)* ***Uniqueness of the martingale measure across all bonds*** *- Under some regularity conditions (see [HJM92] for the details) there exists a unique equivalent martingale measure Q, called the risk-neutral measure under which the discounted ZCB $P_{t,T}/S^0_t$ is a martingale for all $T \in [0, T^*]$ and $t \in [0, T]$ if and only if*

$$\alpha_{t,T} = -\sigma^T_{t,T}\left(\phi_t - \int_t^T ds\sigma_{t,s}\right), \quad \text{for} \quad \text{all} \quad T \in [0, T^*], t \in [0, T]. \tag{2.2}$$

In particular, for any vector of ZCB maturities $T_1 < \cdots < T_d$ the unique martingale measure $Q^{T_1,\ldots,T_d}$ under which the discounted bond prices $P_{t,T_i}/S^0_t$ are $Q^{T_1,\ldots,T_d}$-martingales is Q.

Among the regularity conditions necessary for the above proposition to be verified is that for a set of maturities $T_1 < \cdots < T_d$ we have

$$\begin{pmatrix} \sigma^T_{t,T_1} \\ \cdot \\ \sigma^T_{t,T_d} \end{pmatrix} \quad \textit{is} \quad \textit{nonsingular} \quad Q \times \lambda(dt) - a.e.. \tag{2.3}$$

Remark 2 — *The uniqueness of the equivalent martingale measure implies that the market is complete. Thus for any $\mathcal{F}_{T_1}$-measurable random variable π there exists a self-financing strategy H of the ZCBs $P_{t,T_1}, ..., P_{t,T_d}$ such that*

$$H^0_{T_1} S^0_{T_1} + \sum_{i=1}^{d} H^i_{T_1} P_{T_1,T_i} = \pi.$$

Hedging of derivatives instruments and completeness of the market will be studied in details in section 2.7.

2.1.2 Change of probabilities and forward neutral measures

As mentioned before the interest rates market quotes vanilla instruments such as swaptions (resp. caplets) in terms of the normal or log-normal implied volatility of the underlying forward swap rate (resp. forward libor rate). In order to do so the pricing is made under a suitably chosen probability measure under which the the forward swap rate (resp. the forward libor rate) is a martingale.

Definition 3 — *Consider a complete market. Let N denote the price process of a generic numeraire i.e. an asset with strictly positive value. The martingale measure Q^N associated to the numeraire N is defined as the martingale measure under which the asset prices expressed in the numeraire unit S_t/N_t are martingales. From [GREK95] for any $\mathcal{F}_T$-measurable random variable X by*

$$\mathbb{E}^{Q^N}[X] = \frac{S^0_t}{N_t}\mathbb{E}^Q\left[X\frac{N_T}{S^0_T}|\mathcal{F}_t\right]. \tag{2.4}$$

Definition 4 — *The U-forward measure Q^U is defined as the measure associated to the numeraire $P_{t,U}$. The annuity measure Q^A is the measure associated to the numeraire $A_t = \sum_{i=1}^n (T_i - T_{i-1}) P_{t,T_i}$ with $T_0 < T_1 < ... < T_n$.*

We define the forward libor rate $L_t(T,U)$ and the forward swap rate $S_t((T_i)_{i=0}^n)$ by

$$L_t(T,U) = \frac{1}{U-T}\left(\frac{P_{t,T}}{P_{t,U}} - 1\right) \tag{2.5}$$

$$S_t((T_i)_{i=0}^n) = \frac{P_{t,T_0} - P_{t,T_n}}{\sum_{i=1}^n (T_i - T_{i-1}) P_{t,T_i}}. \tag{2.6}$$

By definition, the forward libor rate $L_t(T,U)$ is a martingale under Q^U and the forward swap rate $S_t((T_i)_{i=0}^n)$ is a martingale under Q^A. We will denote respectively by $S_t(T,n,\tau)$ and $A_t(T,n,\tau)$ the forward swap rate associated to the IRS with fixed payment lag of τ, namely $S_t((T+k\tau)_{k=1}^n)$ and the associated annuity. Following the results of [GREK95] we can characterize the dynamics of the bonds under the forward and annuity measures. Let W be a Brownian motion under the risk-neutral measure (defined by proposition 1), defining

$$W_t^U = W_t + \int_0^t ds \sigma_{s,U} \tag{2.7}$$

$$W_t^A = W_t + \sum_{i=1}^n \int_0^t ds \omega_t^i \sigma_{s,U}, \tag{2.8}$$

where $\omega_t^i = \delta_i \frac{P_{t,T_i}}{A_t}$. We have that W^U and W^A are Brownian motions respectively under Q^U and Q^A.

Let us stress that the change of measure (2.8) implies that under the annuity measure Q^A the drift term will involve the stochastic weights ω^i. In the context of markovian TS models (which we present in the following) this often results in complicated dynamics of the underlying risk factors under the measure Q^A.

2.2 Modeling with state variables

The arbitrage-free modeling framework defined by proposition 1 is general and includes a very broad class of models. However its general version lacks of tractability. In particular, a *markovian* structure is desirable to allow an efficient implementation of the model. We follow [HKP00] and give the definition of the class of markovian TS model. Note that at an abstract level the class of models defined in [HKP00] includes a very broad range of models. The main contribution of HK lies in their innovative calibration procedure which will be described later in the chapter.

Definition 5 — *An IR model is said to be markovian or Markov functional if*

1. *There exist a martingale measure Q associated with a numeraire N*
2. *There exist an $\mathbb{R}^d$ valued stochastic process F that is a Markov process under Q*
3. *N is a function of F i.e. $N_t = N(t, F_t)$*
4. *The ZCBs are functions of F i.e. $P_{t,T} = P(t, T, F_t)$ for $0 \le t \le T \le T^*$.*

The components of the process F are often called the state variables.

In a markovian TS model, the spot rate can be defined by continuity as a functional of the underlying state variables F. We will assume that F solves a SDE under Q

$$F_t = F_0 + \int_0^t ds\mu(s, F_s) + \int_0^t \sigma(s, F_s)dZ_s. \tag{2.9}$$

Where μ and σ verify the usual regularity conditions for the above SDE to admit a unique strong solution. From the change of measure formula (2.4) we deduce the expression of the change of measure between Q and Q^N,

$$\frac{dQ^N}{dQ}|_{\mathcal{F}_t} = e^{-\int_t^T dsr_s}\frac{N(T, F_T)}{N(t, F_t)}.$$

By application of Girsanov theorem (assuming the functions N and σ are smooth enough), we deduce the change of measure. Let W^N be a standard Brownian motion under the measure Q^N, then we have that W defined by

$$W_t^Q = W_t^N - \int_0^t ds\frac{\partial_F N(s, F_s)}{N(s, F_s)}\sigma(s, F_s),$$

is a Brownian motion under the risk neutral measure. Then it is immediate that F solves an SDE under the risk-neutral measure (with a modified drift function), precisely

$$F_t = F_0 + \int_0^t ds\mu(s, F_s) - \frac{\partial_F N(s, F_s)}{N(s, F_s)}\sigma(s, F_s) + \int_0^t \sigma(s, F_s)dW_s^Q.$$

Conversely if the spot rate is a function of a stochastic process F, and F is Markov under the risk-neutral measure, then any U-forward measure -the measure associated with the ZCB numeraire $P_{t,U}$- fulfills the conditions of definition 5. The following proposition provides an alternative characterization of markovian TS models.

Proposition 6 — *An IR model is markovian or Markov functional or a short rate model if and only if there exist a stochastic process F such that*

1. *F is a diffusion process solution of (2.9) under the risk-neutral measure*
2. *the spot rate r is a function of F i.e. $r_t = r(t, F_t)$*

2.3 Affine term structure models

Affine term structure models (ATSM) are a very broad class of markovian models. Most markovian models used by market practitioners belong to this class. Affine processes are appreciated for their tractability, in particular Laplace and Fourier transforms of the process admit analytical or semi-analytical expressions. In the IR context the affine structure of the model allows for a explicit or semi-explicit representation of market underlying assets such as bonds, swaps and FRAs.

The fundamental result of Duffie and Kan [DK96b] is the following theorem

Theorem 7 — *In a factorial IR model $\sigma(t,\cdot)\sigma(t,\cdot)^T$, $\mu(t,\cdot)$ and $r(t,\cdot)$ are affine functions where we have defined*

$$\begin{aligned} \mu(t,f) &= a^\mu(t) + b^\mu(t)f \\ (\sigma(t,f)\sigma(t,f)^T)_{ij} &= a^\sigma_{ij}(t) + (b^\sigma_{ij}(t))^T f \\ r(t,f) &= a^r(t) + (b^r(t))^T f. \end{aligned}$$

if and only if the ZCBs are exponential affine functions of the state variables i.e.

$$P_{t,T} = \exp(A(t,T) + B(t,T)^T F_t). \tag{2.10}$$

Furthermore the functions A and B are solutions (if a solution exists) of the following system of o.d.e.

$$\begin{cases} \frac{\partial B}{\partial t}(t,T) = -(b^\mu(t))^T B(t,T) - \sum_{i,j=1}^d B_i(t,T)B_j(t,T)b^\sigma_{ij}(t) + b^r(t) \\ \\ \frac{\partial A}{\partial t}(t,T) = -B(t,T)^T a^\mu(t) - B(t,T)^T a^\sigma(t)B(t,T) + a^r(t) \\ A(T,T) = 0, \quad B(T,T) = 0. \end{cases} \tag{2.11}$$

A consequence of the above theorem is that in order to specify an ATSM one only needs to define an affine diffusion under the risk-neutral probability and an affine expression for the spot rate. [DK96b] shows that under some mild non-degeneracy conditions the diffusion term of any affine diffusion process can be written as

$$\sigma(t,f) = \Sigma\sqrt{\operatorname{diag}(\alpha_i + \beta_i^T f, 1 \le i \le d)}, \tag{2.12}$$

where $\Sigma \in \mathcal{M}_d(\mathbb{R})$, $\alpha_i \in \mathbb{R}$ and $\beta_i \in \mathbb{R}^d$. Clearly, the process defined by the SDE

$$F_t = F_0 + \int_0^t ds(a^\mu + b^\mu F_s) + \int_0^t \Sigma\sqrt{\operatorname{diag}(\alpha_i + \beta_i^T F_s, 1 \le i \le d)}dZ_s, \tag{2.13}$$

is not well defined for any choice of the parameters α_i, β_i, a^μ and b^μ. In particular one needs to impose that the matrix $\operatorname{diag}(\alpha_i + \beta_i^T F_s, 1 \le i \le d)$ is positive defined. Dai and Singleton

[DS00] have studied the admissible conditions for the solutions in the canonical state space $\mathbb{R}_+^p \times \mathbb{R}^{d-p}$. They show that by means of a linear transformation we can always assume that an admissible process F in the canonical state space is solution of the SDE

$$F_t = F_0 + \int_0^t ds(a(t) + b(t)F_s) + \int_0^t \Sigma \begin{pmatrix} \mathrm{diag}(\sqrt{F_s^1}, ..., \sqrt{F_s^p}, 0, ..., 0) & 0 \\ 0 & \sigma_{JJ}(t, F_I) \end{pmatrix} dZ_s, \tag{2.14}$$

where $I = \{1...p\}$, $J = \{p+1...d\}$, $F_I^T = (F_1, ..., F_p)$ and $\sigma_{JJ}(t, f)$ is in $\mathcal{M}_{d-p}(\mathbb{R})$. This proves that by means of a linear transformation any process F in the canonical state space $\mathbb{R}^p \times \mathbb{R}^{d-p}$ can be decomposed in p CIR processes and $p - d$ stochastic volatility processes, the volatility of which is given by a convex combination of the p CIR processes.

The choice of the state space $\mathbb{R}_+^p \times \mathbb{R}^{d-p}$ is probably the most natural, but certainly not the more general. It is in fact sufficient to impose the positivity of the instantaneous variance-covariance matrix $\mathrm{diag}(\alpha_i + \beta_i^T F_s, 1 \leq i \leq d)$. Cuchiero et al. [CFMT11] have studied admissible constraints for affine diffusion in the space of positive semidefinite matrixes.

2.4 Examples of affine term structure models

In this section we provide examples of ATSM commonly used for IR management. For each model, we write the ZCBs dynamics and discuss the implied volatility cube generated by the model.

2.4.1 The multi-dimensional linear Gaussian model

The linear Gaussian model (LGM hereafter) pioneered by El Karoui et al. [EKLM+91] and [EKL92] is a reference for managing IR derivatives. The model is appreciated for its simplicity, and generates an intuitive yield curve dynamics. Without loss of generality we give the following definition of the model. The underlying state variables Y are assumed to be an vectorial Ornstein-Uhlenbeck process under the risk-neutral measure

$$dY = K(\theta - Y)dt + \sigma dZ \tag{2.15}$$

where Z is a d-dimensional Brownian motion, $K \in \mathcal{M}_p(\mathbb{R})$, $\quad \theta \in \mathbb{R}^p$ and $\sigma \in \mathcal{M}_{p,d}(\mathbb{R})$. We assume that the spot rate is given by

$$r_t = \delta + \sum_{i=1}^{p} Y_t^i. \tag{2.16}$$

The model is a natural generalization of the Vasicek model. It also has similar properties, first the model generates negative interest rates. While this is regarded as a limitation of the model, it does not prevents practitioners from using the model. In particular, in a low interest rates environment, the Gaussian model allows to keep a "high" variance, while in alternative standard models such as the log-normal model or the CIR model the variance of rates depends on their level. Another property is the ergodicity property of rates. If the matrix K is similar to a positive matrix, then the model is ergodic, in the sense that the OU process Y converges to its stationary distribution,and similarly the yield curve also converges to a limit distribution. It is generally assumed distinct speed of mean-reversion for each factors. This leads to assume that

$$K = \text{diag}(k_1, \dots, k_p) \text{ with } 0 < k_1 < \cdots < k_p,$$

and we work under this assumption in the sequel.

Yield curve dynamics

The yield curve admits an explicit expression

$$P_{t,t+\tau} = \exp\left(A(\tau) + B(\tau)^T Y_t\right), \tag{2.17}$$

where

$$\begin{aligned} B(\tau) &= -(K^T)^{-1}\left(I_p - e^{-K^T\tau}\right) 1_{\mathbb{R}^p} \\ A(\tau) &= -\delta\tau + \int_0^\tau ds B(s)^T K\theta + \frac{B(s)^T\sigma\sigma^T B(s)}{2}. \end{aligned}$$

In the homogeneous version of the model the yield curve movements are entirely generated by the movements of the underlying state variables Y, in particular between two dates t and $t+\Delta t$, the yield curve variations are given by the support functions $\frac{1-e^{-k\tau}}{k\tau}$ and the variations of the underlying state variables $\Delta Y = Y_{t+\Delta t} - Y_t$. As illustrated by Figure 2.1 the mean reversion parameters K play the role of scale factors for the yield curve movements through the support functions. The factors Y^i associated with the larger parameters k_i impact the whole yield curve term structure and play the role of a homothety of the curve. The factors Y^i associated with the smaller parameters k_i will drive more the long term curve than the short term curve. As illustrate by Figure 2.2, setting different scales for the parameters K it is possible to identify the factors to different movements of the yield curve such as the level and the slope. The methodology used to extract the factors will be detailed in section 5.3.

Implied volatility

Since ZC bonds are exponential affine functions of the underlying state variables (2.17) and the volatility of the state variables is constant, the forward ZC bond $\frac{P_{t,T}}{P_{t,U}}$ is log-normal under the U-forward measure. Also, as a consequence of the change of measure (2.7) the state variables are Gaussian under any U-forward measure. Forward IBOR and swap rates are not exactly Gaussian under the corresponding pricing measure, however they happen to be "quasi" Gaussian, implying that the model does not exhibits a pronounced smile. Let us

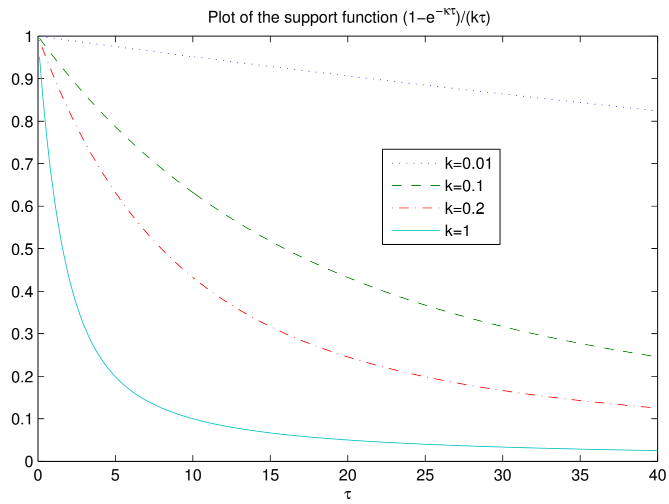

Figure 2.1: Plot of the support function $\frac{1-e^{-k\tau}}{k\tau}$ for different values of k.

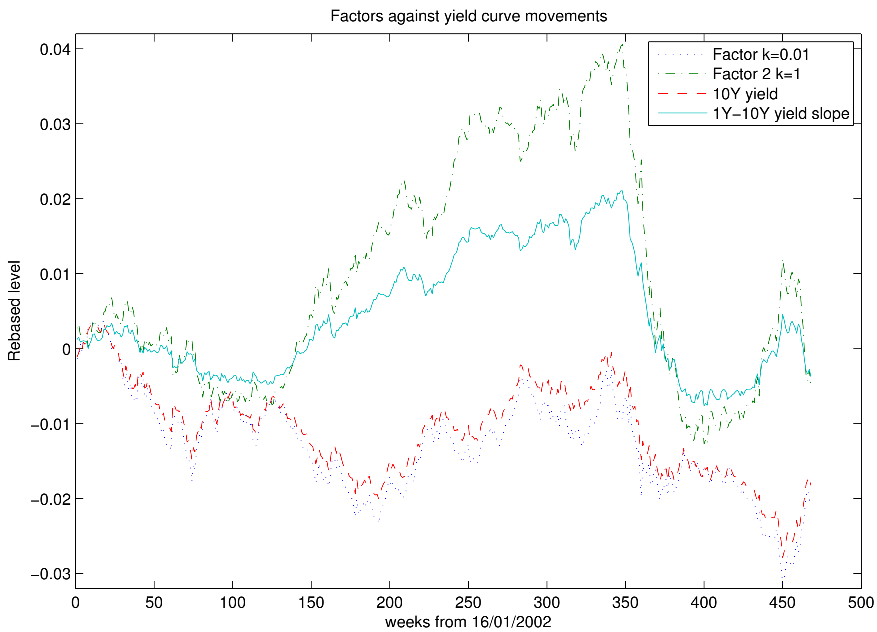

Figure 2.2: Plot of the factors extracted from the weekly variations of the euro yield curve extracted from swaps between 2002 and 2011 against the movements of the curve. The extraction procedure will be detailed in a dedicated chapter.

denote by $\Sigma(\tau,\delta)$ the implied volatility of an option on the forward zero-coupon bond $\frac{P_{t,T}}{P_{t,T+\delta}}$ with expiry τ. We observe that this volatility is close to the volatility of a caplet of expiry τ and maturity δ, and it admits an explicit expression

$$\Sigma(\tau,\delta)^2\tau = \int_0^\tau ds \Delta B(s,\delta)^T V \Delta B(s,\delta), \tag{2.18}$$

where $\Delta B(s,\delta) = B(s+\delta) - B(s)$ and $V = \sigma\sigma^T$.

As illustrated by Figure 2.3 the parameters k_i also play the role of time scales for the maturity and expiry term structure of volatility. For a diagonal volatility matrix σ the behavior is similar to the one we have described for the yield curve and entirely determined by the support functions $(\frac{1-e^{-k_i\delta}}{k_i})^2\frac{1-e^{-2k_i\tau}}{2k_i\tau}$. The effect of off-diagonal elements of the variance-covariance matrix $\sigma\sigma^T$ are determined by the support functions $\frac{1-e^{-k_i\delta}}{k_i}\frac{1-e^{-k_j\delta}}{k_j}\frac{1-e^{-(k_i+k_j)\tau}}{(k_i+k_j)\tau}$. Let us note that for a positive defined matrix K the implied volatility decreases exponentially with expiry. This is not consistent with what we observe in the market, where the implied volatility expiry term structure has a humped shape for maturities up to 2 years. Like forward libor rates, forward swap rates are not linear functions of the underlying state variables Y. We have described the drift feezing procedure in paragraph 1.3.2. Since forward libor rates are "quasi-linear" functions of the factors Y, the forward swap rate is also a "quasi-linear" function of the factors Y, and thus "quasi-normally" distributed. In practice we observe that the model does not exhibits a swaption volatility smile. A reasonable approximation of the implied normal volatility of swaptions can be obtained combining the drift freezing technique approximation on one hand and approaching the swap by an affine function of the underlying state variables Y.

The classic approximation (see [PA10], [BM06]) consists in freezing the weights ω_t^k to their value at 0, this leads to an approximated SDE of the forward swap rate under the annuity measure, the solution of which is a Gaussian process. The forward swap rate is a martingale under the annuity measure Q^A, therefore we must have

$$dS_t((T_i)_{i=0}^m) = \left(\omega_t^0 B(T-t) - \omega_t^m B(T_m - t) - S_t((T_i)_{i=0}^m)\delta\sum_{k=1}^m \omega_t^k B(T_k - t)\right)^T \sqrt{V} dB^A. \tag{2.19}$$

We freeze the weights ω_t^k and the value of the swap rate $S_t((T_i)_{i=0}^m)$ in the diffusion to their value at the pricing date. After this approximation, the swap rate is Gaussian under the annuity measure with a time-dependent normal volatility $\sigma^S(u)$ given by

$$\sigma^S(u) = \sqrt{B^S(u)^T V B^S(u)},$$

where we have defined

$$B^S(u) = \omega_t^0 B(T-u) - \omega_t^m B(T_m - u) - S_t((T_i)_{i=0}^m)\delta\sum_{k=1}^m \omega_t^k B(T_k - u).$$

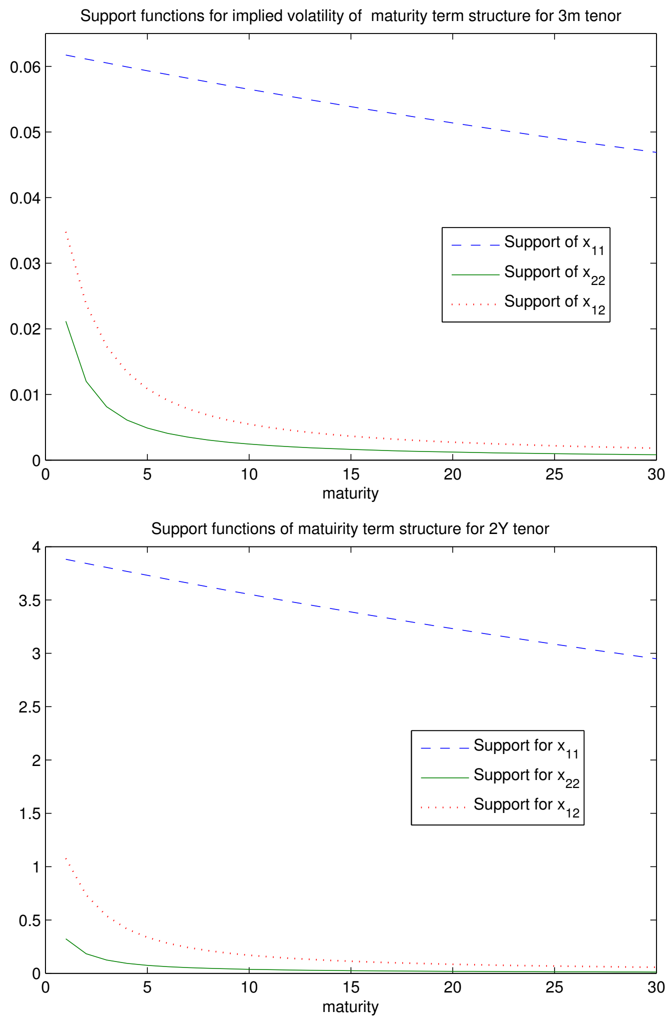

Figure 2.3: Support functions for the volatility term structure. $K = \text{diag}(0.01, 1)$. Maturity 3m up, expiry 2y down.

Let $\Sigma_t^S((T_i)_{i=0}^m), K)$ denote the time t implied normal volatility of a swaption with payment dates $(T_i)_{i=0}^m)$ and strike K, we have the following approximation in the LGM model

$$\Sigma^S((T_i)_{i=0}^m), K)^2(T_0 - t) = \text{Tr}\left(M_t^S((T_i)_{i=0}^m)V\right), \tag{2.20}$$

where we have defined

$$M_t^S((T_i)_{i=0}^m) = \int_t^{T_0} du B^S(u)^T V B^S(u).$$

Let us stress one major difference between the expression of the implied volatilities (2.18) and (2.20). The first is completely time-homogeneous, when the time-to-expiry is constant the volatility of caplets will not change. As opposed to this, the swaption implied volatility (or more precisely its approximation) is not time-homogeneous, since the coefficients of the matrix M^S will depend on the actual values of the weights and forward swap rate on the pricing date. The important fact here is that while the model is time-homogeneous (as it appears clearly from the yield curve dynamics (2.17)), the swaption volatility cube generated by the model is not.

2.4.2 The quadratic Gaussian model

The Quadratic Gaussian Model (QGM hereafter) has been defined by El Karoui et al. in [EKD98]. Piterbarg [Pit09] specified the model as a perturbation of the multi-dimensional linear Gaussian model. While keeping tractability, the model overcomes some of the limits of the LGM generating a smile.

Again we assume that the state variables follow an OU process defined by the SDE (2.15) under the risk neutral measure. The spot rate is given by

$$r_t = \delta + \sum_{i=1}^{n} Y_t^i + Y_t^T \gamma Y_t, \tag{2.21}$$

where $n \leq p$. This class of models belongs to the general class of ATSM, to see this it is sufficient to define a vector of state variables of dimension $p + p(p+1)/2$ by

$$F = (Y^1, ..., Y^p, (Y^1)^2, Y^1Y^2, ..., Y^1Y^p, ..., (Y^p)^2),$$

and observe that F is an affine diffusion and the spot rate is an affine function of F.

Yield curve dynamics

The yield curve curve admits a semi analytical expression

$$P_{t,t+\tau} = \exp\left(A(\tau) + B(\tau)^T Y_t + Y_t^T G(\tau) Y_t\right), \tag{2.22}$$

where G, B and A solves the following system of o.d.e.

$$\begin{aligned}
\partial_\tau G(\tau) &= 2G\sigma\sigma^T G - GK - K^T G - \gamma, \quad G(0) = 0 \\
\partial_\tau B(\tau) &= \left(-K^T + 2G\sigma\sigma^T\right) B + 2GK\theta - \begin{pmatrix} 1_{\mathbb{R}^n} \\ 0 \end{pmatrix}, \quad B(0) = 0 \\
\partial_\tau A(\tau) &= B^T K\theta + \frac{1}{2} B\sigma\sigma^T - \delta, \quad A(0) = 0.
\end{aligned}$$

It is possible to impose constraints on the model parameters to prevent the model to generate negative interest rates (see [EKD98]). Similarly we can impose constraints on the model parameters so that the model is ergodic.

The support function G admits a semi-analytical expression by solving a matrix Riccati differential equation (MRDE) (we will discuss the resolution of MRDE in details in 3.1.2). Let us mention that the solution of a Riccati o.d.e. may explode in finite time. We refer to [Sha86] for a detailed analysis of the explosion of Riccati o.d.e.. A sufficient condition for non explosion of G is that γ is positive semidefinite, in which case the matrix function G is negative defined and converges. The support function B admits an explicit expression

$$B(\tau) = \int_0^\tau du \exp\left(-K(\tau-u) + 2\int_u^\tau ds G(s)\sigma\sigma^T\right)\left(2G(u)K\theta - 1_{\mathbb{R}^p}\right).$$

When $G(\tau)\sigma\sigma^T$ and $G(\tau)K\theta$ are small the support function B is a perturbation of the support function of the LGM model. This is the case in the specification of the QGM of Piterbarg [Pit09]. This is illustrated by Figure 2.5 which represents the coefficients of the vector function B obtained in a 4 factors QGM assuming that $K = \text{diag}(0.01, 0.1, 0.2, 1)$, $\sigma = 0.01 I_d$, $\theta = 0$, $n = p$ and $\gamma = 5 I_d$. The higher the mean-reversion, the lower the impact of the quadratic perturbation on the support function. The behavior of the support function G is illustrated by Figure 2.4. The graphics are obtained with the specification of the model of [Pit09], we have considered the toy 2-factors model defined by $K = \text{diag}(0.1, 0)$, $\theta = 0$, $\sigma = \text{diag}(0.01, 0.005)$, $n = 1$ and $\gamma = \epsilon \begin{pmatrix} \rho & \bar{\rho}/2 \\ \bar{\rho}/2 & 0 \end{pmatrix}$.

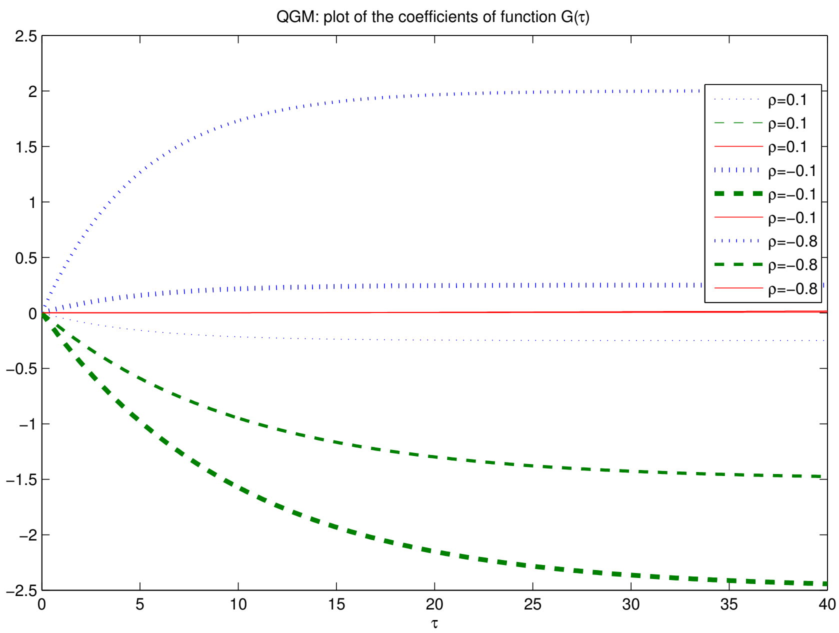

Figure 2.4: QGM: plot of the coefficients of the matrix G for different values of the parameter ρ in a 2-dimensional QGM. The blue, green and red curve are respectively the coefficients G_{11}, G_{12} and G_{22}.

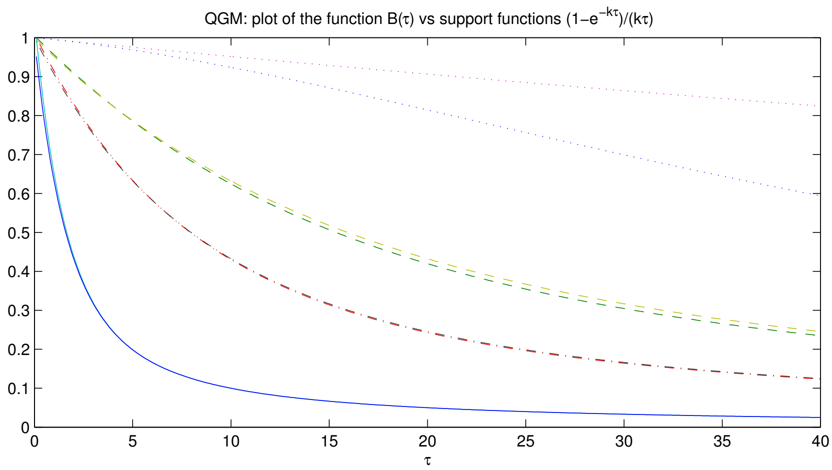

Figure 2.5: QGM: plot of the coefficients of the support function $-B$ against the support functions of the LGM model $\frac{1-e^{-k\tau}}{k\tau}$ for different values of k.

This specification is specially designed to limit the dependence of the spot rate in the second factor of the model, which plays the role of a pseudo stochastic volatility. One can get an intuition of this by analyzing the expression of the spot rate

$$r_t = \delta + Y_t^1 + \epsilon(\rho(Y_t^1)^2 + \bar{\rho} Y_t^1 Y_t^2). \tag{2.23}$$

As illustrated by Figure 2.4, the coefficient $G_{22}(\tau)$ is negligible compared to the other coefficient of the matrix, as a consequence the second factor Y_2 will only appear in the expression of the zero-coupon bonds through the term $G_{12}(\tau)Y^1Y^2$.

Implied volatility

The distribution of the yield curve is defined by a Gaussian variable (the linear term) plus a "chi-square" like variable (the quadratic term). Let us note that the volatility of ZCBs is stochastic,

$$dP_{t,T} = P_{t,T}\left(r_t dt + (B(T-t) + 2G(T-t)Y_t)^T \sigma dW\right),$$

it is a linear function of the factors Y, and thus the prices of swaptions and caplets generated by the QGM exhibit a volatility smile. We will discuss the characteristics of the volatility generated by Markovian models later in the section, for now let us just stress that both the volatility and the yield curve are linear functions of the factors Y. The model parameters, and in particular ϵ and ρ determine the shape of the smile. The parameter ϵ determines the level of the coefficients of the matrix G and thus the magnitude of the quadratic term $Y^T G(\tau) Y$ which determines the departure from normality of the rates. It thus determines the convexity of the volatility smile. The parameter ρ determines the sign of the coefficient G_{11}. As mentioned before the coefficient G_{22} is small w.r.t. to the coefficients G_{11} and G_{12}, so that the zero-coupon bond is approximatively

$$\log P_{t,T} = A(T-t) + B_1(T-t)Y_t^1 + G_{11}(T-t)(Y_t^1)^2 + 2G_{12}Y_t^1Y_t^2.$$

We can separate the above expression into a linear term $B_1(T-t)Y_t^1$ which is the main driver of the yield curve and a quadratic term $G_{11}(T-t)(Y_t^1)^2 + 2G_{12}(T-t)Y_t^1Y_t^2$ which drives the stochastic volatility. As illustrated by Figure 2.4, the parameter ρ determines the magnitude and sign of G_{11}, the higher the absolute value of the parameter ρ, the more the factor Y^1 will drive the quadratic term in (2.22), the more the model will be a local volatility model. For ease of notations we will drop the time dependence and note B_1, G_{11} and G_{12}. The correlation between these two terms gives an indication on the skew generated by the model, we have

$$cov(G_{11}(Y_t^1)^2 + 2G_{12}Y_t^1Y_t^2, B_1Y_t^1) = B_1G_{11}cov(Y_t^1, (Y_t^1)^2) + 2G_{12}B_1cov(Y_t^1Y_t^2, Y_t^1). \quad (2.24)$$

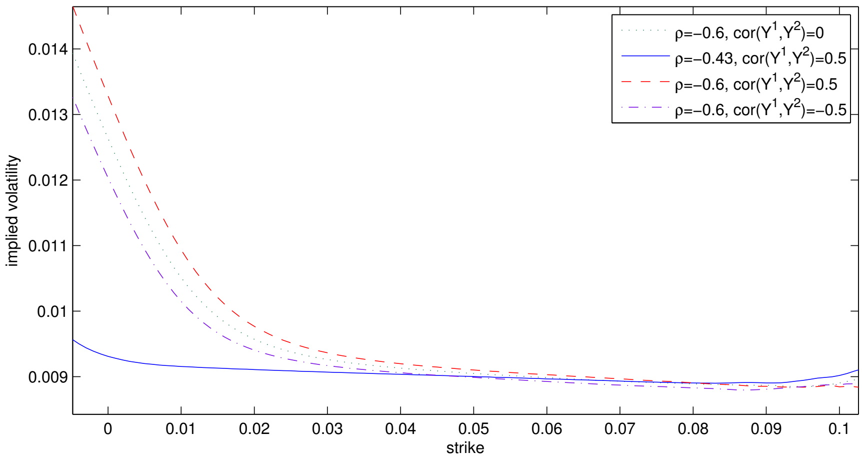

Figure 2.6: QGM: Plot of the implied volatility smile of the 2Y expiry 1Y maturity caplet. The price is obtained by Monte Carlo simulation with 10^7 samples. The ATM rate is around 3.7%. The graphics are obtained with $\epsilon = 1$.

Let us now analyze the expression (2.24). Assuming that the parameter K are set following the reasoning we have presented for the LGM, the drivers of the value (and in particular of the sign) of the correlation are B_1G_{11}, G_{12} and $cov(Y^1Y^2, Y^1)$. From Figure 2.4 we know that the sign of G_{12} doesn't changes with ρ, so that sign of the second term of the right hand side of (2.24) only depends on the correlation between Y^1 and Y^2. The sign of G_{11} depends on the sign of ρ. The parameters ρ and $cov(Y^1, Y^2)$ are thus the determinants of both the magnitude and the sign of the covariance (2.24). Figures 2.6 shows how these parameters can be used to generate different shapes of the smile. In this particular example positive values of the parameter ρ tend to generate a flat smile.

Let us now discuss the volatility smile dynamics. In the homogeneous version of the model the parameters are assumed to be constant over time. The dynamics of the yield curve and of the volatility smile is assumed to be entirely captured by the underlying state variables Y. In the previous paragraph we have illustrated how in the LGM the underlying state variables capture the main movements of the yield curve. The specification of the QGM analyzed here is designed to keep a LGM-like dynamics of the yield curve and add a factor for the volatility dynamics. In the two factors model we expect the first factor to drive the yield curve and the second factor to drive the smile. Figure 2.7 shows that the movements of the factor Y^2 imply a change in the convexity of the smile (the higher the value of Y^2, the higher the convexity of the smile). The ATM smile is fairly unsensitive to the changes of this factor. The movements of the smile generated by the movements of the factor Y^1 depend on the choice of the parameter ρ. As mentioned before, the parameter ρ is a slider which determines the characteristic of the volatility of the model, $\rho = 0$ corresponds to a "stochastic volatility" version of the model and $\rho = 1$ corresponds to "local volatility" version of the model. By construction the movements of Y^1 are assimilated to the movements of the yield curve, and in particular of the ATM forward rate. The movements of the smile generated by a movement of Y depend upon the value of the parameter ρ. For $\rho = -0.6$ (Figure 2.8) we observe the well known undesirable behavior of local volatility models, the smile moves in the same direction of the ATM forward rate (meaning it moves in the opposite direction of the forward ZCB, which is the underlying asset) and that the ATM value of volatility is sensitive to the movements of the underlying forward rate. For $\rho = 0$ we observe that a similar behavior of the smile but the ATM volatility value is less sensitive to the movements of the ATM forward rate, the smile slides as the ATM forward rate slides, which is close to be the behavior of stochastic volatility models.

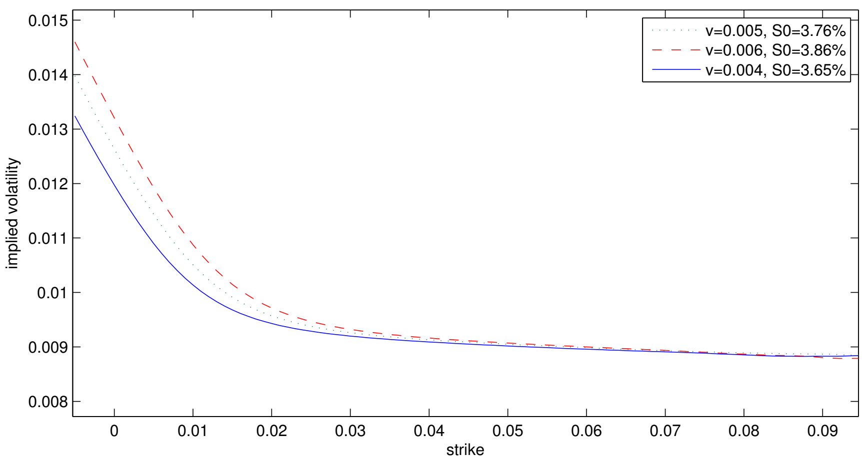

Figure 2.7: QGM: Plot of the implied volatility smile of the 2Y expiry 1Y maturity caplet for different initial value v of the factor Y^2. The price is obtained by Monte Carlo simulation with 10^7 samples. The graphics are obtained with $\epsilon = 1$.

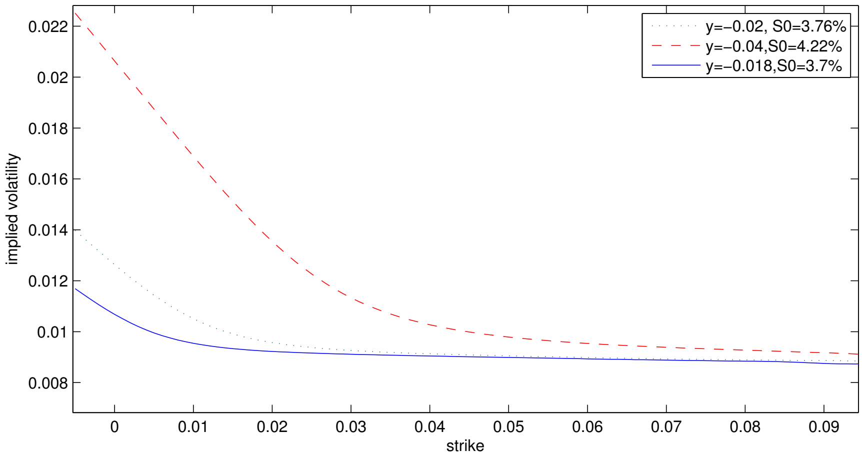

Figure 2.8: QGM: Plot of the implied volatility smile of the 2Y expiry 1Y maturity caplet for different initial value y of the factor of Y^1. The price is obtained by Monte Carlo simulation with 10^7 samples. The graphics are obtained with $\epsilon = 1., \rho = 0.6$ and $v = 0.005$.

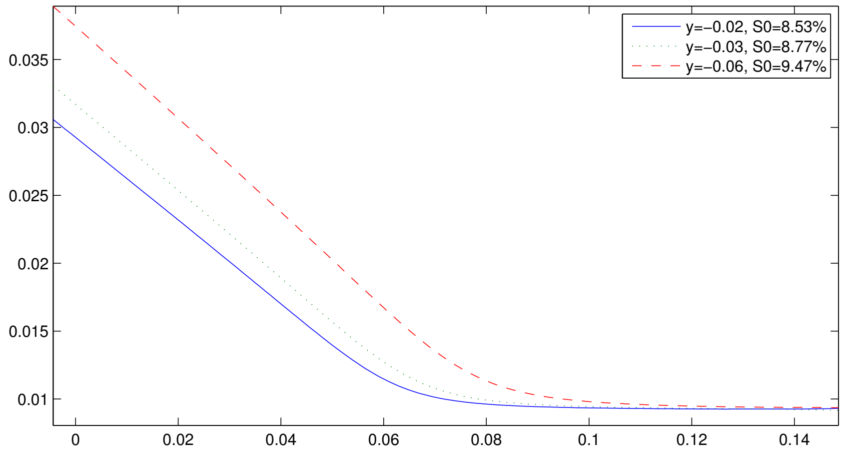

Figure 2.9: QGM: Plot of the implied volatility smile of the 2Y expiry 1Y maturity caplet for different initial value y of the factor of Y^1. The price is obtained by Monte Carlo simulation with 10^7 samples. The graphics are obtained with $\epsilon = 0.5, \rho = -0.6$ and $v = 0.05$.

2.5 Markov functional models

Ahead of the class of ATSM is the MF model defined in [HKP00]. Up to now we have not tackled the calibration problem. The Markov functional models of HK are particularly interesting for their calibration procedure. In this section we briefly describe the main characteristics of the model, and formalize the calibration procedure. Let us note that the practical implementation of the calibration procedure poses some serious technical challenges[1], those aspects are not treated here since it is not the purpose of our work. Instead we take a formal approach to the model and presents some crucial theoretical facts on it. At an abstract level the class of MF models as defined by definition 5 includes a very broad range of models. The main contribution of HK lies in their innovative calibration procedure. When we refer to MF models, we will refer to the models calibrated following the HK algorithm. MF models can be viewed as the analogue of Dupire's implied diffusion for interest rates. In order to handle the supplementary dimension of the fixed income market (i.e. the tenor of the underlying) HK have designed a calibration procedure which is backward and discrete as opposed to Dupire's calibration which is global and continuous.

2.5.1 Calibration procedure

In this paragraph we give a qualitative description of the calibration procedure, a detailed description is given in appendix A. Market quotes of vanilla products (caplets, digital caplets and swaptions) give the knowledge of the implied distributions of swap and IBOR rates under the pricing probability. The calibration can be seen as the choice of the model parameters so that the distributions (of swap and IBOR rates) generated by the model matches the implied distributions. In MF models the distributions of the rates generated by the model depend on the distribution of the underlying markov process F and on the functional. The model philosophy is to (arbitrarily) choose the distribution of F, and to adjust the functional[2].

In the IR world the pricing is made under the forward-neutral probabilities, therefore the market quotes give an information on the implied distribution of swap and IBOR rates under the forward-neutral probabilities. The power of HK's calibration is that it allows a stepwise reconstruction of the distributions of zero-coupon bonds or equivalently of instantaneous forward rates under the corresponding forward-neutral probabilities. This is done through a discrete backward algorithm: we first extrapolate the market information given by quotes with long maturities and then proceed backward using the arbitrage-free martingale property. Suppose we want to price an exotic product which depends on the value of a generic underlying S (we think to S as a swap rate or a CMS rate) at dates $T_1, ..., T_n$. In a MF model the process S is a function of the underlying Markov process F, precisely there exist a function $S(t, f)$ such that

$$S_t = S(t, F_t), \forall t.$$

In order to price such a product we need to determine the functions $S(T_i, \cdot)$ and $N(T_i, \cdot)$ for $i = 1, ..., n$. We assume we know the functional $N(T_n, \cdot)$ (in practice this is usually done by taking the terminal zero-coupon bond P_{t,T_n} as numeraire). Given the knowledge of a

[1]those familiar with the implementation of Dupire's local volatility model know how hard it is to go from theory to practice

[2]This has the obvious virtu that the resulting model is extremely efficient in terms of implementation. In particular the model is easy to simulate.

continuum of market prices $(C(T_i, K))_K$ of vanillas (we usually take digital caplets) with maturity $T_1, ..., T_n$ on the underlying S, we first use the prices $(C(T_n, K))$ to extrapolate the functional $S(T_n, \cdot)$. We then proceed by induction using the relationships

$$\frac{P(T_i, T_{i+1}, f)}{N(T_i, f)} = \mathbb{E}^{\mathbb{Q}}\left[\frac{1}{N(T_{i+1}, F_{T_{i+1}})}|F_{T_i} = f\right], \quad i = 1, ..., n-1$$

which are a direct consequence of the martingale property.

Let us insist on the characteristics of this calibration procedure. For obvious reasons we refer to it as a backward algorithm. On the other hand we can observe that we only use the martingale property between the dates T_{i+1} and T_i. This explains why we refer to it as a discrete algorithm. Finally we emphasize that the algorithm strongly relies on the following assumptions.

Hypothesis 1 — *We assume that the forward IBOR and swap rates are increasing monotonic function of the markov process X.*

Hypothesis 2 — *We assume we know the distribution of the Markov process F under the martingale measure $\mathbb{Q}$.*

2.5.2 Yield curve dynamics and implied volatility

By construction the model fits the initial yield curve and the whole implied volatility smile of a column or a co-terminal ant-diagonal of the swaption volatility cube. The functional mapping between the underlying state variables and the yield curve is defined for a finite set of tenors and for a finite set of dates only. In fact, the model fails to produce a continuous dynamics for the yield curve. Similarly there is no notion of the volatility outside the set of dates that have been used to calibrate the model. The model gives a static representation of the market and the dynamical features of the model are entirely determined by the distributional assumptions made on the underlying driving process (which is completely arbitrary). In the next section we provide some evidence of the limits of MF models.

2.6 Volatility and hedging in factorial models

In the preceding we have given some examples of the factorial models, some of which exhibit a volatility smile. In this section we discuss the nature of the volatility in these models. A clear distinction has to be made between local volatility models and stochastic volatility models. Strictly speaking in both cases the volatility of the underlying asset is a stochastic process, however in local volatility models the instantaneous volatility process is a function of the underlying asset. The nature of the volatility has some strong implications on the dynamics of the volatility smile generated by the model. Through a detailed pricing example using MF model we show the consequence in terms of hedging forward volatility products.

2.6.1 The nature of volatility

This subsection presents some views on the nature of volatility in factorial models and tries to introduce the notion of model risk. The results presented here are the synthesis of a more general work on the consequences of calibration of interest rates models which was jointly conducted with Nourreddine El Hadj Braiek and has benefited of the supervision of Nicole El Karoui.

A Markovian model cannot be considered as a "pure" stochastic volatility model. The yields are functions of the underlying state variables F. Given a set of rates $R_t^1, ..., R_t^p$ (the rates R^i can be viewed either as yields, IBOR rates or swap rates), the vector $R^T = (R^1, ..., R^p)$ is a function of the underlying state variables, $R = \psi(F)$. If the functional relationship is invertible then we can write the diffusion process of R

$$dR = (\cdots)dt + \sigma(t, R)dZ,$$

for some function $\sigma(t, r)$. Strictly speaking, the model is thus a local volatility model (as any Markovian TS model) as long as functional relationship between R and F is invertible. Sufficient conditions for a Markovian model to exhibit an unspanned volatility are discussed in [CD02]. We have seen that the specification of the QGM given in [Pit09] is specifically designed to limit the dependence of the yield curve on a chosen factor which is therefore difficult to express in terms of the yield curve and acts as a "pseudo" stochastic volatility. We will follow this path in building a multi-dimensional stochastic variance-covariance LGM.

The nature of the volatility has major implications in terms of hedging. In the equity market local volatility model assume that any option can be hedged by trading the underlying asset only, whereas in stochastic volatility model even the simple call option cannot be perfectly hedged/replicated by trading in the underlying asset only. Similarly in interest rates market Markovian model will allow to hedge by trading in a finite (usually small depending on the dimension of the underlying state variables) set of assets (such as IRS or FRAs). The capability or not to hedge a derivative by trading in the underlying asset only is linked to the notion on complete market which will be developed later.

2.6.2 The implications of local volatility: an introduction to hedging and model risk

It is well known that local volatility has some undesirable features in terms of dynamics of the volatility smile, we refer to [HKLW02] and [Gat06] for a detailed discussion on the subject. As mentioned before MF models are the archetype of local volatility model for interest rates, the diffusion is directly inferred by market prices of vanilla options. While the calibration procedure insures that the model fits a continuum of market prices for a chosen set of tenors and maturities, the model performs poorly for hedging forward volatility products. The MF model gives us the occasion to give an introductory example of the consequences of the structure of the model and the calibration procedure in terms of hedging.

Let us analyze a "pure" forward volatility product, and show that the forward volatility vision in a MF model is not satisfactory. In particular the resulting hedging strategy contradicts the natural hedging strategy obtained by static replication. Also the product we analyze here is not actually traded in the market, there are very similar derivatives in the interest rates exotic market such as vol bondlets, volatility swaps (see [PA10] p. 221 for the payoff description), and the analysis we conduce here extends to these products. The choice has been made because the smoothness of the payoff allows to apply a static replication formula suggesting a natural hedging strategy which is model independent.

Product description

The product we call a forward start variance swap (FSV hereafter) pays at every instant $T_2, T_2+\delta, ..., T_2+n\delta$ the fixed amount

$$\delta\left(S_{T_2}\left(T_2,n,\tau\right)-S_{T_1}\left(T_2,n,\tau\right)\right)^2.$$

Then the payoff is given by

$$A_{T_2}\left(T_2,n,\delta\right)\left(S_{T_2}\left(T_2,n,\tau\right)-S_{T_1}\left(T_2,n,\tau\right)\right)^2.$$

Which allows us to compute the price as follows

$$\mathrm{FSV}\left(0\right)=A\left(0,T_2,n,\tau\right)\mathbb{E}^{A(T_2,n,\tau)}\left[\left(S_{T_2}\left(T_2,n,\tau\right)-S_{T_1}\left(T_2,n,\tau\right)\right)^2\right]$$

An easy computation leads to

$$\mathrm{FSV}\left(0\right)=A\left(0,T_2,n,\delta\right)\mathbb{E}^{A(T_2,n,\tau)}\left[S_{T_2}\left(T_2,n,\tau\right)^2\right]-A\left(0,T_2,n,\tau\right)\mathbb{E}^{A(T_2,n,\tau)}\left[S_{T_1}\left(T_2,n,\tau\right)^2\right].$$

Static replication strategy

It comes directly from the static replication formula [CW02] that

$$\mathrm{FSV}\left(0\right)=2\left(\int_0^{S(0,T_2,n,\tau)}SP\left(0,T_2,n,\tau,K\right)dK+\int_{S(0,T_2,n,\tau)}^{+\infty}\mathrm{SR}\left(0,T_2,n,\tau,K\right)dK\right)$$

$$-2A\left(0,T_2,n,\tau\right)\left(\int_0^{S(0,T_2,n,\tau)}\mathbb{E}^{A}\left[\left(K-S_{T_1}\left(T_2,n,\tau\right)\right)^+\right]dK+\int_{S(0,T_2,n,\tau)}^{+\infty}\mathbb{E}^{A}\left[\left(S_{T_1}\left(T_2,n,\tau\right)-K\right)^+\right]dK\right) \quad (2.25)$$

The above formula suggests a natural hedging strategy for the FSV, which is made of a long position in a portfolio of swaptions of maturity T_2 and a short position in a portfolio of forward start swaptions of maturity T_1. Forward start swaptions, also called mid-curves are exotic products actively traded in the market. We can write the forward swap rate as

$$S_{T_1}(T_2,n,\tau) = w^1_{T_1} S_{T_1}\left(T_1, \left[\frac{T_2-T_1}{\tau}\right]+n,\tau\right) - w^2_{T_1} S_{T_1}\left(T_1, \left[\frac{T_2-T_1}{\tau}\right],\tau\right),$$

where

$$\begin{aligned} w^1_t &= \frac{\sum_{k=1}^{\left[\frac{T_2-T_1}{\tau}\right]+n} P_{t,T_1+k\tau}}{\sum_{k=1}^{n} P_{t,T_2+k\tau}} \\ w^2_t &= \frac{\sum_{k=1}^{\left[\frac{T_2-T_1}{\tau}\right]} P_{t,T_1+k\tau}}{\sum_{k=1}^{n} P_{t,T_2+k\tau}}. \end{aligned}$$

It is common market practice to freeze the value of the weights w^1 and w^2 at zero and price the forward start swaption as an options on the weighted spread $w^1(0)\,S_{T_1}\left(T_1,\left[\frac{T_2-T_1}{\tau}\right]+n,\tau\right) - w^2(0)\,S_{T_1}\left(T_1,\left[\frac{T_2-T_1}{\tau}\right],\tau\right)$. The pricing is then performed by using a model which exploits the distribution of $S_{T_1}\left(T_1,\left[\frac{T_2-T_1}{\tau}\right]+n,\tau\right)$ embedded by the market prices of swaptions of maturity T_1 and tenor $T_2+n\delta-T_1$, the distribution of $S_{T_1}\left(T_1,\left[\frac{T_2-T_1}{\tau}\right],\tau\right)$ embedded by the market prices of swaptions of maturity T_1 and tenor T_2-T_1 and by adding some correlation assumptions between these two swap rates[3]. The important fact here is that the price of the forward starting swaption strongly depends on the market prices of vanilla swaptions of maturity T_1. Let us note that when the ratio $\frac{T_2-T_1}{n\delta}$ is close to zero (meaning the tenor of the IRS is much larger then the difference between the maturity and the fixing date), the forward start swaptions prices are close to the corresponding swaptions prices of maturity T_1, and the short position can be approximated by a short position in a portfolio of swaptions of maturity T_1.

Model price

We can write

$$\begin{aligned} A(0,T_2,n,\tau)\,\mathbb{E}^{A(T_2,n,\tau)}[S_{T_1}(T_2,n,\tau)^2] &= \mathbb{E}^{\mathbb{N}}[\frac{1}{N_{T_1}} A_{T_1}(T_2,n,\tau)\,S_{T_1}(T_2,n,\tau)^2] \\ &= \mathbb{E}^{\mathbb{N}}[\frac{N_{T_1}}{A_{T_1}(T_2,n,\tau)}\left(\frac{P_{T_1,T_2}-P_{T_1,T_2+n\tau}}{N_{T_1}}\right)^2] \end{aligned}$$

In a MF model $A_t(T,n,\tau)$, N_t and $P_{t,T}$ are functions of the underlying Markov process F, respectively $A(t,T,n,\tau,F_t)$, $N(t,F_t)$ and $P(t,T,F_t)$, thus we have:

[3]One common technique consists in using a copula to model the joint distribution

$$\begin{aligned}
\frac{A\left(T_1,T_2,n,\tau,F_{T_1}\right)}{N\left(T_1,F_{T_1}\right)} &= \mathbb{E}^{\mathbb{N}}\left[\frac{A\left(T_2,T_2,n,\tau,F_{T_2}\right)}{N\left(T_2,F_{T_2}\right)}|F_{T_1}\right] = K_{T_2,T_1}\frac{A\left(T_2,T_2,n,\tau,\cdot\right)}{N\left(T_2,\cdot\right)}\left(F_{T_1}\right)\\
\frac{P\left(T_1,T_2,F_{T_1}\right)}{N\left(T_1,F_{T_1}\right)} &= \mathbb{E}^{\mathbb{N}}\left[\frac{1}{N\left(T_2,X_{T_2}\right)}|F_{T_1}\right] = K_{T_2,T_1}\frac{1}{N\left(T_2,\cdot\right)}\left(F_{T_1}\right)\\
\frac{P\left(T_1,T_2+n\tau,F_{T_1}\right)}{N\left(T_1,F_{T_1}\right)} &= \mathbb{E}^{\mathbb{N}}\left[\frac{P\left(T_2,T_2+n\tau,F_{T_2}\right)}{N\left(T_2,F_{T_2}\right)}|F_{T_1}\right] = K_{T_2,T_1}\frac{P\left(T_2,T_2+n\tau,\cdot\right)}{N\left(T_2,\cdot\right)}\left(F_{T_1}\right).
\end{aligned}$$

Using the above expressions we can give the model price of the preceding product

$$\mathrm{FSV}^{model}\left(0\right) = A\left(0,T_2,D\right)\mathbb{E}^{A(T_2,D)}\left[S\left(T_2,T_2,n,\tau,F_{T_2}\right)^2\right]$$

$$-N\left(0,f_0\right)\mathbb{E}^{\mathbb{N}}\left[\frac{1}{K_{T_2,T_1}\frac{A(T_2,T_2,n,\tau,\cdot)}{N(T_2,\cdot)}\left(F_{T_1}\right)}\left(K_{T_2,T_1}\frac{1}{N\left(T_2,\cdot\right)}\left(F_{T_1}\right) - K_{T_2,T_1}\frac{P\left(T_2,T_2+n\tau,\cdot\right)}{N\left(T_2,\cdot\right)}\left(F_{T_1}\right)\right)^2\right]. \tag{2.26}$$

Now we can observe that the model price (2.26) only depends on the distribution of the process F, and on the functionals of swap rates, zero coupon bonds and numraire at date T_2. It follows directly from the calibration procedure described in section A that the functionals $S(T_2,T_2,n,\tau,\cdot), N(T_2,\cdot), P(T_2,T_2+n\tau,\cdot)$ only depend upon the market prices of vanillas of maturities larger then T_2. Therefore if we build our hedging portfolio using a MF model (by computing the sensitivities of the product price w.r.t. the calibration set for instance) we will obtain an hedging portfolio which only contains vanilla products with maturities larger then T_2. Another way to say this is that if we compute the Vegas of the product with a MF model, we will only have Vegas on the swaptions and/or digital caplets with maturities higher then T_2, whereas intuitively we should also have a Vegas sensitivity to the products with maturity T_1. We call this feature of MF models "the forward variance paradox" because this result clearly contradicts the natural hedging portfolio suggested by the static replication given in the preceding paragraph. This feature shows how the distribution of the process F plays an important role in terms of forward volatility risk vision and management. As shown by (2.26), the only way for the model price of a FSV to depend on the market prices of maturity T_1 (and therefore to hope a certain consistency with (2.25)) is that the transition kernel between T_1 and T_2 K_{T_2,T_1} depends on the market prices of maturity T_1. This is not possible in the traditional vision of the MF model, in which the distribution of X is arbitrarily chosen.

2.7 Hedging Interest rates derivatives: theory

The present section is devoted to the hedging of IR options. We define the notion of completeness of the market and try to build natural hedging portfolios for hedging simple derivatives. We describe how the modeling assumptions -and in particular a factorial framework- impacts the hedging portfolios computation. While the market seems to be naturally incomplete (when considering the bond prices curve as an infinite-dimensional underlying), assuming that the dynamics is driven by a finite dimensional underlying driving process usually results in multiple "equivalent" strategies for hedging. The choice of the best hedging strategy then often results of practical considerations such as the cost of implementation of a strategy.

2.7.1 Hedging in a general AOA framework

Let us consider a general arbitrage-free framework defined in the section 2.1. We assume that in the economy p assets are traded continuously from time 0 until T^*. Their prices are modeled as semimartingale $S^i, i = 1...p$. In this chapter we will assume that the the spot rate cannot be negative i.e. $r_s \geq 0, s \geq 0$[4]. Let us denote by S^0 the money market account defined by

$$S_t^0 = e^{\int_0^t dsr_s}.$$

Definition 8 — *A trading strategy is a $p+1$-dimensional predictable process*

$$H = (H_t^0, H_t^1, ..., H_t^p, t \leq T),$$

its value is given by

$$V(H)_t = H_t^0 S_t^0 + \sum_{i=1}^p H_t^i S_t^i.$$

We say the portfolio is self-financed if

$$dV(H)_t = H_t^0 dS_t^0 + \sum_{i=1}^p H_t^i dS_t^i.$$

The notion of complete market is relative to a set of chosen hedging instruments.

Definition 9 — *A financial market is complete w.r.t. the securities $S^1, ..., S^p$ if any contingent claim is attainable i.e. if for any $\mathcal{F}_T$-measurable random variable π there exists a self-financing strategy H such that*

$$V(H)_T = \pi \quad P - a.e.$$

A fundamental result of Harrison and Pliska states that if there exists an unique equivalent martingale measure such that the discounted asset prices $\tilde{S}^i, i = 1...p$ are martingales, then the market is complete w.r.t. $S^1, ..., S^p$.

[4]Let us note that this assumption is not verified in most of the models we have presented in 2.4, nor in the model we will define in 4. We make this assumption to avoid technical difficulties which can occur when defining the prices as expectations of discounted payoffs.

2.7.2 Hedging in factorial models

In factorial models the market is driven by an underlying Markov process of finite dimension F. Let us assume that F follows an homogeneous diffusion process on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_t, P)$

$$F_t = F_0 + \int_0^t ds\mu(s, F_s) + \int_0^t \sigma(s, F_s)dZ_s,$$

where $\sigma : \mathbb{R}^d \to \mathcal{M}_d(\mathbb{R})$ is smooth and a.e. non singular[5]. In particular the market prices are functions of this process. Let us denote by $P(t, T, F_t)$ (resp. $f(t, T, F_t)$) the function giving the price of ZCB $P_{t,T}$ (resp. the value of the instantaneous forward rate $f_{t,T}$). Given that σ is a.e. non singular, condition (2.3) translates into the following condition

$$\begin{pmatrix} \partial_f P(t, T_1, f)^T \\ \cdot \\ \partial_f P(t, T_d, f)^T \end{pmatrix} \quad is \quad nonsingular \quad \lambda_d(df) \times \lambda(dt) - a.e.. \tag{2.27}$$

Continuous factorial models we have presented in sections 2.3 and 2.4 structurally verify (2.2), and are therefore arbitrage-free and complete in the sense of proposition 1. The uniqueness of the equivalent martingale measure implies that the market is complete. Thus for any $\mathcal{F}_{T_1}$-measurable random variable π there exists a self-financing strategy H of the ZCBs $P_{t,T_1}, ..., P_{t,T_d}$ such that

$$H^0_{T_1} S^0_{T_1} + \sum_{i=1}^{d} H^i_{T_1} P_{T_1,T_i} = \pi.$$

2.7.3 Hedging with market securities

In practice there exist no such security as a default free ZCB, the hedging of contingent claims is thus made through market traded contracts such as Interest Rates Swaps (IRS), Forward Rates Agreements (FRA), Swaptions, Caps and Floors. The existence of a self-financing replicating portfolio made of ZCBs is not relevant for practical purposes. As mentioned before the notion of complete market depends on the securities we use to hedge contingent claims. In the following we investigate the theoretical results on hedging with securities different from ZCBs.

General HJM framework

Consider now an European option with payoff given by an $\mathcal{F}_t$-measurable variable h such that $\mathbb{E}\left[|h|^2\right] < \infty$ and denote by V its price at time t. The price is given by

$$V_t = \mathbb{E}^Q \left[e^{-\int_t^T ds r_s} h | \mathcal{F}_t \right],$$

[5]Note that this is the minimal assumption to incorporate the models of practical interest. In particular we want this assumption to be verified for the CIR model, in which the volatility is given by $\sqrt{F}$.

The discounted price $\tilde{V}_t = V_t/S_t^0$ is a martingale under Q, by the martingale representation theorem there exists an adapted vector process K such that

$$\tilde{V}_t = V_0 + \int_0^t K_s^T dZ_s.$$

Given a set of hedging securities $S^1, ..., S^p$, the martingale representation theorem also implies that

$$\tilde{S}_t^i = S_0^i + \int_0^t (K_s^i)^T dZ_s. \tag{2.28}$$

An hedging portfolio $\phi_t = (H^0, H^1, ..., H^p)$, where H^0 denotes the cash and H^i denotes the quantity of asset S^i in the portfolio. Then any self-financing portfolio solving the linear equation

$$\sum_{i=1}^p H_t^i K_t^i = K_t \tag{2.29}$$

is a replicating portfolio for the derivative V. Suppose $H^1, ..., H^p$ solves the above equation then defining

$$H_t^0 = \tilde{V}_t - \sum_{i=1}^p H_t^i \tilde{S}_t^i,$$

the portfolio H replicates the option V and it is easy to prove that it is self-financed. There is no existence result for such a portfolio in the general setting, meaning that the model is not complete in general.

Hedging in factorial models

The market asset prices are functionals of the state variables X, we have

$$\begin{aligned} V_t &= V(t, F_t) \\ S_t^i &= S^i(t, F_t) \end{aligned}$$

The martingale representation (2.28) of the prices are explicitly given by these functionals:

$$\begin{aligned} K_t &= e^{-\int_0^t ds r_s} \sigma(t, F_t)^T \partial_X V(t, F_t) \\ K_t^i &= e^{-\int_0^t ds r_s} \sigma(t, F_t)^T \partial_X S^i(t, F_t). \end{aligned}$$

We can rewrite the hedging equation (2.29)

$$\sigma(t, F_t)^T \sum_{i=1}^{p} \partial_X S^i(t, F_t) K_t^i = \sigma(t, F_t)^T \partial_X V(t, F_t) \tag{2.30}$$

Since $\sigma(t, f)$ is assumed to be non singular for any $t \geq 0, f \in \mathbb{R}^d$, the self-financing hedging portfolio is entirely defined by the vector $H^{1:p} \in \mathbb{R}^p$ solution of the following linear equation

$$J(t, F) H^{1:p} = \partial_X V. \tag{2.31}$$

where

$$J(X) = \begin{pmatrix} \partial_X S^1 & \cdots & \partial_X S^p \end{pmatrix}$$

If a solution of the above linear equation exists the Hedging portfolio defined by

$$\phi_t = (\tilde{V}_t - H^{1:p} \cdot \tilde{S}_t^i, H^{1:p})$$

is self financed and replicates V. Let us denote by r the rank of the matrix $J(t, F)$ and r_v the rank of $(J(t, F) \partial_X V)$. One of the following assertions holds:

- if $r_v = r < p$ the set of solutions of the linear system is an infinite set and has the form $\{u_0, u_1, ..., u_{p-r} : k_i \in \mathbb{R}, 1 \leq i \leq p - r\}$, where $u_0, u_1, ..., u_{n-r}$ are vectors of $\mathbb{R}^p$ satisfying $J(t, F) u_0 = \partial_X V$, $J(t, F) u_i = 0$ for $1 \leq i \leq p - r$.
- if $r_v = r = p$, the solution is unique
- if $r < r_v$ the system has no solutions

2.7.4 Being consistent with market practice

In the preceding, we have build hedging portfolio assuming the set of hedging instruments is fixed during the life of the product. Market constraints such as liquidity imply that common market practice differs from the theory. IR derivatives books are hedged globally. A set of hedging instruments is chosen daily in order to optimize the execution cost. The assets S in the hedging portfolio are thus changing daily, and the assets in the portfolio are not necessarily liquidated. In particular, for liquidity reasons we always trade at par IRS. The fixed leg rate of standard IRS is settled every day by the market. In this section we build theory of hedging which takes into account these practical considerations.

Changing the hedging instruments

We start by analyzing the problem in the one-dimensional factorial IR setup. Let us consider a derivative contract maturing at T. Assume that we change the hedging instruments at dates $0 = t_1 < \cdots < t_n = T$, meaning that in the interval $[t_i, t_{i+1})$ we can only trade the asset $S^{(i)}$.

One asset portfolio Let us initially assume that in the interval $[t_i, t_{i+1})$ the hedging portfolio is made of a quantity H of asset S^i and H^0 of cash. Meaning that the other assets are liquidated prior to date t_i. In the interval $[t_1, t_2)$, we build our portfolio as follows

$$\begin{aligned} H_t &= \frac{\partial_X V(t, X_t)}{\partial_X S^{(1)}(t, X_t)} \\ H_t^0 &= \tilde{V}_t - H_t \tilde{S}_t^{(1)}, \end{aligned}$$

this portfolio is self-financed and replicates the price of the option in the interval $[t_1, t_2)$. At t_2^- we sell the portfolio which by continuity of the price process is worth V_{t_2} and we build the portfolio for the next interval $[t_2, t_3)$

$$\begin{aligned} H_t &= \frac{\partial_X V(t, X_t)}{\partial_X S^{(2)}(t, X_t)} \\ H_t^0 &= \tilde{V}_t - H_t \tilde{S}_t^{(2)}. \end{aligned}$$

We follow this procedure up to the maturity of the contract. The hedging portfolio is then written as follows

$$H_t = \sum_{i=1}^{n-1} 1_{[t_i, t_{i+1})} \frac{\partial_X V(t, X_t)}{\partial_X S^{(i)}(t, X_t)} \tag{2.32}$$

$$H_t^0 = \tilde{V}_t - \sum_{i=1}^{n-1} 1_{[t_i, t_{i+1})} H_t \tilde{S}_t^{(i)}. \tag{2.33}$$

This strategy is naturally self financed and replicates the option. The weights of the portfolio are discontinuous, with discontinuity at the times at which the hedging instruments change $t_1, ..., t_{n-1}$.

Multiple assets portfolio We assume that in the interval $[t_i, t_{i+1})$ we add a quantity H of asset S^i to the hedging portfolio and H^0 of cash. We assume that the other assets are liquidated prior to date t_i. In the interval $[t_1, t_2)$, we build our portfolio as follows

$$\begin{aligned} H_t &= \frac{\partial_X V(t, X_t)}{\partial_X S^{(1)}(t, X_t)} \\ H_t^0 &= \tilde{V}_t - H_t \tilde{S}_t^{(1)}, \end{aligned}$$

this portfolio is self-financed and replicates the price of the option in the interval $[t_1, t_2)$. At t_2^- by continuity of the price process the portfolio is worth V_{t_2} and we build the portfolio for the next interval $[t_2, t_3)$, as mentioned before we cannot sell the hedging portfolio, we thus want to hedge the net position in the portfolio which is given by $V_{t_2} - \frac{\partial_X V(t_2, X_{t_2})}{\partial_X S^{(1)}(t_2, X_{t_2})} S^{(1)}_{t_2}$

$$
\begin{aligned}
H_t &= \frac{\partial_X V(t, X_t) - H_{t_2^-} \partial_X S^{(1)}(t, X_t)}{\partial_X S^{(2)}(t, X_t)} \\
H_t^0 &= \tilde{V}_t - H_{t_2^-} \tilde{S}_t^{(1)} - H_t \tilde{S}_t^{(2)}.
\end{aligned}
$$

We follow this procedure up to the maturity of the contract. The hedging portfolio is then written as follows

$$
\begin{aligned}
H_t &= \frac{\partial_X V(t, X_t) - \sum_{i=2}^{\gamma(t)} H_{t_i^-} \partial_X S^{(i-1)}(t, X_t)}{\partial_X S^{(i-1)}((t, X_t)} \\
H_t^0 &= \tilde{V}_t - \sum_{i=2}^{\gamma(t)} H_{t_i^-} \tilde{S}_t^{(i-1)} - H_t \tilde{S}_t^{(\gamma(t))}
\end{aligned}
$$

where $\gamma(t) = \max\{i : t_i \leq t\}$. This strategy is naturally self financed and replicates the option.

2.7.5 Continuous framework

Rebalancing of the hedging portfolios occurs regularly in practice, and the hedging instruments are changed with the same frequency. It is thus relevant to write a theory of hedging in continuous time, assuming that the hedging instrument changes continuously (the at par swap contract). Miming the discrete framework we have described above, we will analyze two different situations. First we assume that we instantaneously resell the hedging portfolio, so that the portfolio at instant t is only made of the asset $S^{(t)}$. Secondly we assume that the portfolio is build by successively adding a quantity of the asset $S^{(t)}$ with no liquidation of the asset part of the portfolio.

To find the continuous analogue of the discrete strategies presented above, we follow the general framework of [PT11]. Since we will continuously change the hedging instrument, we need to consider a continuum of hedging instruments. At date t, we consider a continuum of asset prices $(S_t^{(y)})_k$ (typically the set of IRS with fixed leg of rate y).

General framework Let $\left(S_t^{(\cdot)}\right)_t$ be a continuous adapted price process taking values in the space of real valued functions from $\mathbb{R}$ to $\mathbb{R}$ (this allows to have negative rates). For any $y \in \mathbb{R}$, $\left(\tilde{S}_t^{(y)}\right)_t$ is the discounted price of an asset parameterized by y. It is a martingale under the risk-neutral measure Q and:

$$
d\tilde{S}_t^{(y)} = K_t^{(y)} dZ_t,
$$

which in the factorial framework gives

$$d\tilde{S}_t^{(y)} = \partial_X S^{(y)}(t, X_t) b(X_t) dZ_t.$$

Definition 10 — *A self-financing portfolio is defined by a finite random adapted measure on $\mathbb{R}$ ϕ, and an adapted process η such that the dynamics of the value of the portfolio $V(\phi_t) = \int_{\mathbb{R}} \phi_t(dy) S_t^{(y)} + \eta_t S_t^0$ is given by*

$$dV(\phi_t) = \int_{\mathbb{R}} \phi_t(dy) dS_t^{(y)} + \eta_t dS_t^0. \tag{2.34}$$

Reselling the portfolio If we assume that the portfolio is continuously liquidated, and that at date t the portfolio is only made of one asset. Then it is immediate to build a self-financing replicating portfolio by taking

$$\phi_t(dy) = \frac{K_t}{K_t^{(Y_t)}} \delta_{Y_t}(dy),$$

where $\delta_u(dy)$ denotes the Dirac mass at the point $u \in \mathbb{R}$ and Y_t is an adapted process. The cash holding process is given by

$$\eta_t = \tilde{V}_t - \frac{K_t}{K_t^{(Y_t)}} \tilde{S}_t^{(Y_t)}.$$

In the factorial framework we would have

$$\phi_t(dy) = \frac{\partial_X V(t, X_t)}{\partial_X S^{(Y_t)}(t, X_t)} \delta_{Y_t}(dy).$$

Let us illustrate through an example, consider we want to hedge a derivative with maturity T. Take $S_t^{(y)}$ to be the time t value of the forward interest rate swap starting at T and of tenor D (the tenor will most likely be chosen w.r.t. the tenor of the swap rate involved in the payoff) with fixed rate y. Then a natural choice for the process $(Y_t)_t$ is the forward swap rate of maturity T and tenor D, $S_t(T, D)$. In this case we will have

$$\begin{aligned} \phi_t(dy) &= \frac{K_t}{K^{(S_t(T,D))}} \delta_{S_t(T,D)}(dy) \\ \eta_t &= \tilde{V}_t. \end{aligned}$$

Note that the above strategy corresponds to the limit when the time discretization step goes to zero of the discrete strategy defined by (2.32). Let us note that as in [PT11] the hedging portfolios provided above are given by discrete finite measures on $\mathbb{R}$. Following this procedure the set of hedging instruments in a Book of derivatives will also be finite and discrete.

Adding to the portfolio We assume now that we are not able to liquidate the hedging portfolio, and that we instantaneously add a new asset $S^{(y_t)}$ to the portfolio, without liquidating the assets already in the portfolio. In this case we cannot find a discrete measure as the solution of the hedging problem anymore. The hedging measure $\phi_t(dy)$ is necessarily continuous. A self-financing hedging portfolio must satisfy the following equations

$$\begin{aligned} dV(\phi_t) &= \int_{\mathbb{R}} \phi_t(dy) dS_t^{(y)} + \eta_t S_t^0 \\ V_t &= \int_{\mathbb{R}} \phi_t(dy) S_t^{(y)} + \eta_t S_t^0. \end{aligned}$$

We thus must have

$$\begin{aligned} \tilde{V}_t &= \int_{\mathbb{R}} \phi_t(dy) \tilde{S}_t^{(y)} + \eta_t \\ K_t &= \int_{\mathbb{R}} \phi_t(dy) K_t^{(y)}. \end{aligned}$$

First let us note that we can transcript the strategy (2.34) in the continuous asset space framework. The discrete hedging portfolio for $t \in [t_n, t_{n+1}]$ is defined by

$$\phi_t(dy) = \delta_{Y_{t_n}} \left(K_{t_n} - K_{t_{n-1}} \frac{K_{t_n}^{(Y_{t_{n-1}})}}{K_{t_{n-1}}^{(Y_{t_{n-1}})}} - \cdots - K_{t_0} \frac{K_{t_1}^{(Y_{t_0})}}{K_{t_0}^{(Y_{t_0})}} \times \cdots \times \frac{K_{t_n}^{(Y_{t_{n-1}})}}{K_{t_{n-1}}^{(Y_{t_{n-1}})}} \right) / K_{t_n}^{(Y_{t_n})} + \cdots + \delta_{Y_{t_0}} K_{t_0} / K_{t_0}^{(Y_{t_0})}$$

Part II

An affine stochastic variance-covariance model

Chapter 3

Modeling dependence through affine process: theory and Monte Carlo framework

This chapter is devoted to the definition of the theoretical and numerical framework for the underlying state variables dynamics. Our state variables dynamics is a stochastic variance-covariance version of the state variables dynamics (2.15). We give a clear interpretation of the underlying state variables of the model, separating them in yield curve factors and volatility factors. The yield curve factors are assumed to follow a multi-dimensional Ornstein-Uhlenbeck (OU) dynamics with stochastic variance-covariance matrix. To keep the model tractable we chose an affine (Wishart-like) process in the space of positive semidefinite variance-covariance. Affine processes in the space of positive semidefinite matrixes have been studied in details by a number of papers ([Bru91], [CFMT11], [AA13]). These processes have naturally been coupled with a vectorial process to model different markets such as the equity market [DFGT08a] and the interest rate market ([Ben10], [GS07]). Our own contribution to the study of the coupled process is that we are able to develop an efficient Monte Carlo simulation framework for the coupled dynamics even in presence of a correlation structure between the process driving the variance-covariance and the process driving the vector. Relying on the extensive work of [AA13] on the simulation of Wishart processes and their affine extension we are able to provide similar discretization schemes for Monte Carlo simulation.

We present some important properties concerning the Laplace transform of the factors and the ergodicity of the state variable process. We are able to fully characterize the distribution of the process using the infinitesimal generator and the Laplace/Fourier transform of the marginal laws. The affine structure of the process (X, Y) allows us to give formulas for the Laplace transform of the marginal laws by the mean of a Matrix Riccati Differential Equations (MRDE). By studying such equations we are able to provide analytical conditions on the model parameters under which the Laplace transform is well defined and the state variable process verifies ergodicity property.

The chapter is organized as follows, section 3.1 defines the underlying state variables dynamics and provides some theoretical results that characterize its distribution, section 3.2 describes the construction of efficient discretization schemes for Monte Carlo simulation.

3.1 A stochastic variance-covariance Ornstein-Uhlenbeck process

This section is devoted to the definition of a general affine SDE that will be used to model the underlying state variables dynamics. We chose to model a stochastic covariation dependence between state variables through an affine diffusion on the space of positive semidefinite symmetrical matrixes. Wishart-like processes thus appear as a natural candidate. Our state variable dynamics is close to log prices dynamics of model proposed in [DFGT08a].

The underlying state variables are assumed to be solution of the following SDE system:

$$\begin{aligned} Y_t^{x,y} &= y + \int_0^{\top} K\left(\theta - Y_s^{x,y}\right) + \int_0^{\top} c\sqrt{X_s^x}\left[\bar{\rho} dZ_s + dW_s \rho\right] \qquad (3.1) \\ X_t^x &= x + \int_0^{\top} \left(\Omega + (d-1)a^\top a + bX_s^x + X_s^x b^\top\right) ds + \int_0^{\top} \sqrt{X_s^x} dW_s a + a^\top dW_s^\top \sqrt{X_s^x}. \end{aligned}$$

Here, and throughout the paper, $(W_t, t \geq 0)$ denotes a d-by-d square matrix made of independent standard Brownian motions, and Z a p-vector of independent Brownian motions independent of W,

$$\begin{aligned} & x, \Omega \in \mathcal{S}_d^+(\mathbb{R}), a \in \mathcal{M}_d(\mathbb{R}), b \in \mathcal{M}_d(\mathbb{R}) \\ & y, \theta \in \mathbb{R}^p, K \in \mathcal{M}_p(\mathbb{R}), c \in \mathcal{M}_{p\times d}(\mathbb{R}), \rho \in \mathbb{R}^d, |\rho|^2 \leq 1 \quad \text{and} \quad \bar{\rho} = \sqrt{1 - |\rho|^2} \qquad (3.2) \end{aligned}$$

We will consider the particular case of a Wishart process for X which corresponds to the case where

$$\exists \quad \alpha \geq 0 : \Omega = \alpha a^T a. \qquad (3.3)$$

The dependence structure between Y and X through the driving Brownian motions is the same as the one proposed by Da Fonseca, Grasselli and Tebaldi [DFGT08a]. As explained in [DFGT08a], this is the most general way to get a non trivial instantaneous correlation between Y and Z while keeping the affine structure. We have added a drift term that allows to account for properties such as the mean reversion which is important in the interest rates market. From (3.1), we easily get

$$e^{Kt} Y_t = y + \int_0^t e^{Ks} K\theta ds + \int_0^t e^{Ks} c\sqrt{X_s}\left[\bar{\rho} dZ_s + dW_s \rho\right].$$

Therefore, the process Y is uniquely determined once the processes Z, W and X are given. We know by Cuchiero et al. [CFMT11] that the SDE on X has a unique weak solution when $x \in \mathcal{S}_d^+(\mathbb{R})$ and $\Omega \in \mathcal{S}_d^+(\mathbb{R})$, and a unique strong solution if we assume besides that x is invertible and $\Omega - 2a^T a \in \mathcal{S}_d^+(\mathbb{R})$. This leads to the following result.

Proposition 11 — *If $x \in \mathcal{S}_d^+(\mathbb{R})$, $\Omega \in \mathcal{S}_d^+(\mathbb{R})$ there exists a unique weak solution of the SDE (3.1). If we assume moreover that $\Omega - 2a^T a \in \mathcal{S}_d^+(\mathbb{R})$ and $x \in \mathcal{S}_d^{+,*}(\mathbb{R})$ is positive definite, there is a unique strong solution to the SDE (3.1). We denote by*

$$AFF((p, y, K, \theta, c), (d, x, \Omega, b, a), \rho)$$

the law of $(X^x, Y^{x,y})$, and by

$$AFF\left[(p, y, K, \theta, c), (d, x, \Omega, b, a), \rho); t\right]$$

the marginal law of $(X_t^x, Y_t^{x,y})$.

Throughout the document , when we use the notation $AFF((p, y, K, \theta, c), (d, x, \Omega, b, a), \rho)$ or $AFF\left[(p, y, K, \theta, c), (d, x, \Omega, b, a), \rho); t\right]$ we implicitly assume that $\Omega \in \mathcal{S}_d^+(\mathbb{R})$ (resp. $\alpha \geq 0$).

3.1.1 The infinitesimal generator on $\mathbb{R}^p \times \mathcal{S}_d^+(\mathbb{R})$

The infinitesimal generator for the affine process (3.1) is defined by:

Let $y \in \mathbb{R}^p, x \in \mathcal{S}_d^+(\mathbb{R})$,

$$L^{\mathcal{M}} f(x, y) = \lim_{t \to 0^+} \frac{\mathbb{E}\left[f(X_t^x, Y_t^{x,y})\right] - f(x, y)}{t}, \quad f \in \mathcal{C}^2(\mathcal{M}_d(\mathbb{R}) \times \mathbb{R}^p) \quad \text{with bounded derivatives.}$$

Proposition 12 — *Infinitesimal generator on $\mathcal{M}_d(\mathbb{R}) \times \mathbb{R}^p$.*
Let $(X^x, Y^{x,y}) \sim AFF((p, y, K, \theta, c), (d, x, \Omega, b, a), \rho)$ be an affine process. Its infinitesimal generator on $\mathbb{R}^p \times \mathcal{M}_d(\mathbb{R})$ is given by:

$$\begin{aligned} L^{\mathcal{M}} &= \mathrm{Tr}\left(\left[\Omega + (d-1)a^T a + bx + xb^\top\right] D^{\mathcal{M}}\right) + K(\theta - y)D^{\mathcal{V}} + \frac{1}{2}\mathrm{Tr}\left(cxc^T D^{\mathcal{V}} \left(D^{\mathcal{V}}\right)^T\right) \\ &+ \frac{1}{2}\left\{2\mathrm{Tr}(xD^{\mathcal{M}} a^T a D^{\mathcal{M}}) + \mathrm{Tr}(xD^{\mathcal{M}} a^T a (D^{\mathcal{M}})^T) + \mathrm{Tr}(x(D^{\mathcal{M}})^T a^T a D^{\mathcal{M}})\right\} \\ &+ \frac{1}{2}\left\{\mathrm{Tr}\left(xc^T D^{\mathcal{V}} \rho^T a D^{\mathcal{M}}\right) + \mathrm{Tr}\left(xc^T D^{\mathcal{V}} \rho^T a \left(D^{\mathcal{M}}\right)^T\right)\right\} \end{aligned} \tag{3.4}$$

where

$$\begin{aligned} D^{\mathcal{M}} &= (\partial_{x_{ij}})_{ij} \\ D^{\mathcal{V}} &= (\partial_{y_k})_k \end{aligned}$$

Proof : Since the drift and diffusion functions of the process (3.1) are linear they are also lipschitz. Then lemma 46 given in appendix B allows to compute directly the infinitesimal generator $L^{\mathcal{M}}$. □

Here, we have given the infinitesimal generator on $\mathcal{M}_d(\mathbb{R})$, while we know that the affine process $(X_t^x)_{t\geq 0}$ takes values in $\mathcal{S}_d^+(\mathbb{R}) \subset \mathcal{S}_d(\mathbb{R})$. Thus, we can also look at the infinitesimal generator of the diffusion (3.1) on $\mathbb{R}^p \times \mathcal{S}_d(\mathbb{R})$, which is defined by:

Let $y \in \mathbb{R}^p, x \in \mathcal{S}_d^+(\mathbb{R})$,

$$L^{\mathcal{S}} f(x,y) = \lim_{t\to 0^+} \frac{\mathbb{E}\left[f(X_t^x, Y_t^{x,y})\right] - f(x,y)}{t} \quad \text{for} \quad f \in \mathcal{C}^2(\mathcal{S}_d(\mathbb{R})\times\mathbb{R}^p) \quad \text{with bounded derivatives.}$$

For $x \in \mathcal{S}_d^+(\mathbb{R})$, we denote by $x_{\{i,j\}} = x_{ij} = x_{ji}$ the value of the coordinates (i,j) and (j,i) of x, so that $x = \sum_{i,j=1}^d x_{\{ij\}}(e_d^{i,j} + 1_{i\neq j} e_d^{j,i})$ (where $e^{i,j}$ is a matrix with the element column j and row i equal to 1 and all the other elements of the matrix equal to zero). For $f \in C^2(\mathcal{S}_d^+(\mathbb{R}) \times \mathbb{R}^p)$, we then denote by $\partial_{x_{\{i,j\}}} f$ its derivative with respect to the coordinates $x_{\{i,j\}}$. We introduce

$$\pi : \underset{x \to \frac{x+x^T}{2}}{\mathcal{M}_d(\mathbb{R}) \to \mathcal{S}_d(\mathbb{R})}$$

that is such that $\pi(x) = x$ for $x \in \mathcal{S}_d(\mathbb{R})$. Obviously, $f \circ (I_p, \pi) \in \mathcal{C}^2(\mathcal{M}_d(\mathbb{R}) \times \mathbb{R}^p, \mathbb{R})$ and we have

$$L^{\mathcal{S}} f(x,y) = L^{\mathcal{M}} f(\pi(x), y).$$

By the chain rule, we have for $x \in \mathcal{S}_d^+(\mathbb{R}), \partial_{x_{ij}} f(\pi(x), y) = (1_{i=j} + \frac{1}{2} 1_{i\neq j})\partial_{x_{\{i,j\}}} f$ and get the following result.

Corollary 13 — *Infinitesimal Generator on $\mathbb{R}^p \times \mathcal{S}_d(\mathbb{R})$. The infinitesimal generator on $\mathbb{R}^p \times \mathcal{S}_d(\mathbb{R})$ associated to $AFF((p,y,K,\theta,c),(d,x,\Omega,b,a),\rho)$ is given by:*

$$\begin{aligned} L^{\mathcal{S}} = {} & \mathrm{Tr}\left(\left[\Omega + (d-1)a^\top a + bx + xb^\top\right] D^{\mathcal{S}}\right) + K(\theta - y)D^{\mathcal{V}} + 2\mathrm{Tr}(xD^{\mathcal{S}}a^T a D^{\mathcal{S}}) \\ & + \frac{1}{2}\mathrm{Tr}\left(cxc^T D^{\mathcal{V}}\left(D^{\mathcal{V}}\right)^T\right) + \mathrm{Tr}\left(xc^T D^{\mathcal{V}}\rho^T a D^{\mathcal{S}}\right) \end{aligned} \tag{3.5}$$

where

$$\begin{aligned} D^{\mathcal{S}} &= (1_{i=j} + \frac{1}{2} 1_{i\neq j})\partial_{x_{\{i,j\}}} \\ D^{\mathcal{V}} &= (\partial_{y_k})_k. \end{aligned}$$

The proofs and results of Proposition 12, and Corollary 13 are closely related to [DFGT08a]. Of course, the generators $L^{\mathcal{M}}$ and $L^{\mathcal{S}}$ are equivalent: one can be deduced from the other. However, $L^{\mathcal{S}}$ already embeds the fact that the process X lies in $\mathcal{S}_d(\mathbb{R})$, which reduces the dimension from $p + d \times d$ to $p + d(d+1)/2$ and gives in practice shorter formulas. This is why we will mostly work in the sequel with infinitesimal generators on $\mathcal{S}_d(\mathbb{R})$. Unless it is necessary to make the distinction with $L^{\mathcal{M}}$, we will simply denote $L = L^{\mathcal{S}}$.

3.1.2 The Laplace transform of affine processes

The affine structure of the process $(X^x, Y^{x,y})$ allows us to give formulas for the Laplace transform of the marginal laws by the mean of a Matrix Riccati Differential Equations (MRDE). Similar calculations have been made in equity modelling by Da Fonseca et al. [DFGT08a] or Benabid et al. [BBEK08]. The following proposition states the precise result, which is useful for the pricing of Zero-Coupon bonds. By a slight abuse of notations we will drop the dependence of the process $(X^x, Y^{x,y})$ on his initial condition (x, y) and simply denote by (X, Y) the process.

Proposition 14 — *Let $\Lambda, \lambda \in \mathbb{R}^p$, $-\Gamma, \bar{\Gamma} \in \mathcal{S}_d(\mathbb{R})$ and $K \in \mathcal{G}_d(\mathbb{R})$ such that*

$$-\Gamma \in \mathcal{S}_d^+(\mathbb{R}), \tag{3.6}$$

$$-\bar{\Gamma} - \frac{1}{2} c^\top \lambda(t) \lambda^\top(t) c \in \mathcal{S}_d^+(\mathbb{R}), \quad t \geq 0 \tag{3.7}$$

where

$$\lambda^\top(t) = e^{-K^\top t}\Lambda + (K^\top)^{-1}\left(I_p - e^{-K^\top t}\right)\bar{\Lambda}. \tag{3.8}$$

Then, the following system of differential equations

$$\begin{cases} \dot{g} = 2 g a^T a g + g(b + \frac{1}{2} a \rho \lambda^\top c) + (b + \frac{1}{2} a \rho \lambda^\top c)^\top g + \frac{1}{2} c^\top \lambda \lambda^\top c + \bar{\Gamma}, \ g(0) = \Gamma, \\ \dot{\eta} = \lambda^\top K \theta + \mathrm{Tr}\left(g \bar{\alpha}\right), \quad \eta(0) = 0, \end{cases} \tag{3.9}$$

has a unique solution, which is defined on $\mathbb{R}_+$. It satisfies $-g(t) \in \mathcal{S}_d^+(\mathbb{R})$ for any $t \geq 0$. Besides, we have for any $0 \leq t \leq T$:

$$\mathbb{E}\left[\exp\left(\mathrm{Tr}(\Gamma X_T) + \Lambda^\top Y_T + \int_t^T \mathrm{Tr}\left(\bar{\Gamma} X_s\right) + \bar{\Lambda}^\top Y_s ds\right) | \mathcal{F}_t\right] = \exp(\eta(T-t) + \mathrm{Tr}(g(T-t)X_t) + \lambda(T-t)^\top Y_t) \tag{3.10}$$

Proof : The proof is quite standard for affine diffusion. First, we notice that if (3.10) holds, we necessarily have that $M_t = \exp\left(\int_0^t \mathrm{Tr}\left(\bar{\Gamma} X_s\right) + \bar{\Lambda}^\top Y_s ds\right) \exp(\eta(T-t) + \mathrm{Tr}(g(T-t)X_t) +$

$\lambda(T-t)^\top Y_t)$ is a martingale. By (3.5), this yields to

$$\begin{aligned}&\bar\Gamma X_t + \bar\Lambda^\top Y_t - \dot\eta(T-t) - \mathrm{Tr}(\dot g(T-t)X_t) - \dot\lambda(T-t)^\top Y_t + \mathrm{Tr}(g(T-t)[\bar\alpha + bX_s + X_s b^\top])\\ &+ \lambda(T-t)^\top K(\theta - Y_t) + 2\mathrm{Tr}(Xg(T-t)a^\top a g(T-t)) + \frac{1}{2}\mathrm{Tr}(Xc^\top \lambda(T-t)\lambda^\top(T-t)c)\\ &+ \frac{1}{2}\mathrm{Tr}(X[c^\top\lambda(T-t)\rho^\top a^\top g(T-t) + g(T-t)a\rho\lambda^\top(T-t)c]) = 0.\end{aligned}$$

By identifying the constant term and the linear terms with respect to Y_t and X_t, we get (3.9) and $\dot\lambda = K\lambda + \bar\Lambda, \quad \lambda(0) = \Lambda$, which leads to (3.8). By applying Lemma 17 to $-g$, the solution of (3.9) exists and is well defined for $t \geq 0$. Besides, $-g$ stays in $\mathcal{S}_d^+(\mathbb{R})$ by using (3.6) and (3.7).

Then, it remains to check that we have indeed (3.10), and it is sufficient to check it for $t = 0$. To do so, we apply Itô's formula to M and get

$$dM_s = M_s\left[\mathrm{Tr}(g(T-s)[\sqrt{X_s}dW_s a + a^\top dW_s^\top\sqrt{X_s}]) + \lambda(T-s)^\top c\sqrt{X_s}\,[\bar\rho dZ_s + dW_s\rho]\right].$$

Thus, M is a positive local martingale and thus a supermartingale, which gives $M_0 \geq \mathbb{E}[M_T]$. To prove that $M_0 = \mathbb{E}[M_T]$, we use the argument presented by Rydberg [Ryd97]. We define $N_t = M_t/M_0$ in order to work with probability measures. We define for $K > 0$, $\tau_K = \inf\{t \geq 0, \mathrm{Tr}(X_t) \geq K\}$, $\pi_K(x) = \mathbb{1}_{\mathrm{Tr}(x)\leq K}x + \mathbb{1}_{\mathrm{Tr}(x)\geq K}\frac{K}{\mathrm{Tr}(x)}x$ for $x \in \mathcal{S}_d^+(\mathbb{R})$ and consider $N_t^{(K)}$ the solution of

$$\begin{aligned}dN_s^{(K)} =& N_s^{(K)}\big[\mathrm{Tr}(g(T-s)[\sqrt{\pi_K(X_s)}dW_s a + a^\top dW_s^\top\sqrt{\pi_K(X_s)}])\\ &+ \lambda(T-s)^\top c\sqrt{\pi_K(X_s)}\,[\bar\rho dZ_s + dW_s\rho]\big],\\ N_0^{(K)} =& 1.\end{aligned}$$

Clearly, $\mathbb{E}[N_T^{(K)}] = 1$, and under $\frac{d\mathbb{P}^{(K)}}{d\mathbb{P}} = N_T^{(K)}$,

$$dW_t^{(K)} = dW_t - 2\sqrt{\pi_K(X_t)}g(T-t)a - \sqrt{\pi_K(X_t)}c^\top\lambda(T-t)\rho^\top dt$$

is a matrix Brownian motion under $\mathbb{P}^{(K)}$.

We now write $\mathbb{E}[N_T] = \mathbb{E}[N_T\mathbb{1}_{\tau_K\geq T}] + \mathbb{E}[N_T\mathbb{1}_{\tau_K<T}]$. By the dominated convergence theorem, we have $\mathbb{E}[N_T\mathbb{1}_{\tau_K\geq T}] \underset{K\to+\infty}{\to} 0$. Besides, $\mathbb{E}[N_T\mathbb{1}_{\tau_K<T}] = \mathbb{E}[N_T^{(K)}\mathbb{1}_{\tau_K<T}] = \mathbb{P}^{(K)}(\tau_K < T)$, and we have to prove that this probability goes to 1. To do so, we focus on the following SDE

$$\begin{aligned}d\tilde X_t =& (\bar\alpha + (b + 2ag(T-t) + a\rho\lambda^\top(T-t)c)\tilde X_t + \tilde X_t(b^\top + 2g(T-t)a^\top + c^\top\lambda(T-t)\rho^\top a^\top))dt\\ &+ \left(\sqrt{\tilde X_t}dW_t a + a^\top dW_t^\top\sqrt{\tilde X_t}\right)\end{aligned}$$

starting from $\tilde X_0 = X_0$. We check that X solves before τ_K and under $\mathbb{P}^{(K)}$ the same SDE as $\tilde X$ under $\mathbb{P}$. This yields to $\mathbb{P}^{(K)}(\tau_K < T) = \mathbb{P}(\inf\{t \geq 0, \mathrm{Tr}(\tilde X_t) \geq K\} < T)$. Since the SDE

satisfied by $\tilde{X}$ is the one of an affine diffusion on $\mathcal{S}_d^+(\mathbb{R})$, it is well defined for any $t \geq 0$. In particular $\max_{t\in[0,T]} \mathrm{Tr}(\tilde{X}_t) < \infty$ a.s., which gives $\mathbb{P}(\inf\{t \geq 0, \mathrm{Tr}(\tilde{X}_t) \geq K\} < T) \underset{K\to+\infty}{\to} 1$.
□

To the best of our knowledge there is no explicit characterization of the set of convergence of the Laplace transform. The above proposition only allows us to explicitly characterize a subset of $\mathcal{D}_{t,T}$ for the class of state variables we will consider in our modeling framework.

With the Laplace transform (3.10), we have a mathematical tool to check if the process (X, Y) is stationary. This is important for our modeling perspective: unless for some transitory period, one may expect that the factors are stable around some equilibrium. The next proposition give a simple sufficient condition that ensures stationarity.

We recall the following useful result

$$\forall x, y \in \mathcal{S}_d^+(\mathbb{R}),\ \mathrm{Tr}(xy) \geq 0, \tag{3.11}$$

which comes easily from $\mathrm{Tr}(xy) = \mathrm{Tr}(\sqrt{x}y\sqrt{x})$ and $\sqrt{x}y\sqrt{x} \in \mathcal{S}_d^+(\mathbb{R})$.

Proposition 15 — *If $-(b + b^\top) \in \mathcal{S}_d^+(\mathbb{R})$ is positive definite, the process (X, Y) is stationary.*

Proof : It is sufficient to prove the proposition for $a = \epsilon I_d^n$, as we will see later that -without loss of generality- we can reduce the general SDE (3.1) to this particular case. For $x, y \in \mathcal{S}_d(\mathbb{R})$, we use the notation $x \leq y$ if $y - x \in \mathcal{S}_d^+(\mathbb{R})$. By assumption, there is $\mu > 0$ such that $2\mu I_d \leq -(b + b^\top)$. We now apply Proposition 14 with $\bar{\Lambda} = 0$ and $\bar{\Gamma} = 0$. Since $\|\lambda(t)\| \leq \|\Lambda\|$, there is a constant $h > 0$ small enough such that for any $\Lambda \in \mathbb{R}^d$ satisfying $\|\Lambda\| < h$ we have

$$\forall t \geq 0, \mu I_d \leq -[b + \frac{\epsilon}{2} I_d^n \rho \lambda^\top c + (b + \frac{\epsilon}{2} I_d^n \rho \lambda^\top c)^\top] \text{ and } \frac{1}{2} c^\top \lambda \lambda^\top c \leq \frac{\mu^2}{8\epsilon^2} I_d$$

By choosing $\Upsilon = \frac{\mu}{4\epsilon^2} I_d$, we see that the condition (3.7) is satisfied since $\frac{\mu^2}{4\epsilon^2} I_d - \frac{\mu^2}{8\epsilon^2} I_d^n - \frac{1}{2} c^\top \lambda \lambda^\top c \in \mathcal{S}_d^+(\mathbb{R})$ for all $t \geq 0$. Thus, the conclusions of Proposition 14 hold for any $\Lambda \in \mathbb{R}^d$ and $\Gamma \in \mathcal{S}_d(\mathbb{R})$ such that $\|\Lambda\| < h$ and $\Gamma \leq \frac{\mu}{4\epsilon^2} I_d$, and we have $g(t) \leq \frac{\mu}{4\epsilon^2} I_d$ for any $t \geq 0$. We now want to prove that $\lambda(t) \underset{t\to+\infty}{\to} 0$, $g(t) \underset{t\to+\infty}{\to} 0$ and $\eta(t)$ converges when $t \to +\infty$. This will prove the convergence to the stationary law by Lévy's theorem.

From (3.9), we have

$$\frac{1}{2} \frac{d}{dt} \mathrm{Tr}(g^2) = 2\epsilon^2 \mathrm{Tr}(g I_d^n g^2) + \mathrm{Tr}(g^2 [b + \frac{\epsilon}{2} I_d^n \rho \lambda^\top c + (b + \frac{\epsilon}{2} I_d^n \rho \lambda^\top c)^\top]) + \mathrm{Tr}(g \frac{1}{2} c^\top \lambda \lambda^\top c).$$

By (3.11), we get

$$\frac{1}{2} \frac{d}{dt} \mathrm{Tr}(g^2) \leq \frac{\mu}{2} \mathrm{Tr}(g I_d^n g) - \mu \mathrm{Tr}(g^2) + \frac{\mu}{4\epsilon^2} \mathrm{Tr}(\frac{1}{2} c^\top \lambda \lambda^\top c).$$

Since $\mathrm{Tr}(gI_d^n g) \leq \mathrm{Tr}(g^2)$, we get by Gronwall's lemma

$$\frac{1}{2}\mathrm{Tr}(g(t)^2) \leq \frac{1}{2}\mathrm{Tr}(\Gamma^2)e^{-\mu t} + \frac{\mu}{4\epsilon^2}\int_0^t \mathrm{Tr}\left(\left[\frac{1}{2}c^\top \lambda(s)\lambda^\top(s)c\right]^2\right) e^{-\mu(t-s)}ds.$$

We now use that the entries of λ decay exponentially. Since $\int_0^t e^{-\mu' s}e^{-\mu(t-s)}ds \underset{t\to+\infty}{=} O(e^{-\frac{\min(\mu,\mu')}{2}t})$ for $\mu, \mu' > 0$, we get that there exists $C, \nu > 0$ such that $\frac{1}{2}\mathrm{Tr}(g(t)^2) \leq Ce^{-\nu t}$. This gives that $g(t) \underset{t\to+\infty}{\to} 0$ and that $\eta(t) = \int_0^t \lambda^\top(s)\kappa\theta + \mathrm{Tr}\left(g(s)(\Omega + \epsilon^2(d-1)I_d^n)\right) ds$ converges. □

Remark 16 — *The Laplace transform of Wishart processes has been studied in details in [GG14b]. The authors are able to derive an explicit semi-analytic expression for the Laplace transform of the couple $(\bar{X}_t, \int_0^t \bar{X}_s ds, t \geq 0)$, where $\bar{X}$ is a Wishart process. The analytical form allows to derive an explicit expression for the definition domain. Unfortunately the formula cannot be applied to our state variable process (Y, X) for a number of reasons : the first is that we are considering an affine extension of Wishart processes, the second is the mean reverting drift of the process Y.*

3.1.3 A detailed study of the matrix Riccati differential equation

The MRDE (3.9) plays a key role in our modeling framework. Not only it characterizes the distribution of the underlying state variables but, as we will see later, it determines the functional mapping between the factors and fundamental market quantities such as ZCBs. Therefore it is important to be able to characterize the behavior of the solution as much as possible. The solution of (3.9) may explode in finite time, and this is a behavior we want to avoid since it would produce a degenerated yield curve dynamics.

A linearization solution for MRDE

We follow Levin [Lev59] and solve (3.9) by a well known linearization method. We introduce the so-called hamiltonian matrix $H(t)$ defined by

$$H(t) = \begin{pmatrix} (b + \frac{1}{2}a\rho\lambda^\top c)^\top & \bar{\Gamma} - \frac{1}{2} \\ a^\top a & -(b + \frac{1}{2}a\rho\lambda^\top c) \end{pmatrix}. \tag{3.12}$$

It is equivalent to solve the MRDE (3.9) and to solve the linear differential system

$$\begin{cases} (\dot{F} \quad \dot{G}) = (F \quad G)H \\ (F(0) \quad G(0)) = (I_n \quad \Gamma). \end{cases} \tag{3.13}$$

Where F and G are in $\mathcal{M}_d(\mathbb{R})$. It is possible to show that there exists a unique solution[1] of (3.9) if and only if there exists a unique solution of (3.13) and $F(t)$ is non-singular for $t \geq 0$. See [Rei70]. The blow-up time, or equivalently the interval in which the solution of (3.9) is well defined is entirely characterized by singularity of the matrix $F(t)$. Let us denote by t_e the blow-up time of g i.e.

$$t_e = \inf\left\{t : \lim_{s\to t} \max_{1\leq i,j\leq d} |g_{ij}(s)| = \infty\right\}. \tag{3.14}$$

Then t_e is characterized as follows:

$$t_e = \inf\left\{t : \det\left(F(t)\right) \leq 0\right\}. \tag{3.15}$$

Unfortunately this expression of the blow-up time does not leads to an analytical characterization of the convergence set $\mathcal{D}_t$ in the general case. An analytical characterization of the blow-up time can be obtained when the hamiltonian matrix is constant. In this case, the blow-up time depends on the hyperbolic sinus of the eigenvalues of the hamiltonian matrix. We refer to [Sha86] for a detailed study of the blow up phenomena of MRDE.

A sufficient condition for non explosion

We now derive an explicit and easy to verify sufficient condition for non explosion of the solution of (3.9). Let us first note that the explosion behavior of g comes from the quadratic term of the Riccati differential equation $2ga^\top ag$. This term is a positive semidefinite matrix, so the explosion phenomena can only occur in the cone of positive semidefinite matrixes. One way to avoid explosion is to keep the solution outside this cone.

The following lemma allows to build a solution which stays negative.

Lemma 17 — *Let $b : \mathbb{R} \to \mathcal{M}_d(\mathbb{R})$ and $a, c : \mathbb{R} \to \mathcal{S}_d^+(\mathbb{R})$ be continuous functions. We consider the following matrix Riccati differential equation:*

$$\dot{X} + Xa(t)X = b(t)X + Xb(t)^\top + c(t),\ X(0) \in \mathcal{S}_d^+(\mathbb{R}). \tag{3.16}$$

Then, the solution X does not explode and is well defined for $t \in \mathbb{R}_+$. Besides, we have $X(t) \in \mathcal{S}_d^+(\mathbb{R})$ for any $t \geq 0$.

Proof : A proof of this lemma is given in [DE94], here we propose an alternative proof. We follow Levin [Lev59], and consider the following time-dependent linear differential equation on $\mathcal{M}_{2d}(\mathbb{R})$

$$\frac{d}{dt}\begin{bmatrix} M_1(t) & M_2(t) \\ M_3(t) & M_4(t)\end{bmatrix} = \begin{bmatrix} b(t) & c(t) \\ a(t) & -b(t)^\top\end{bmatrix}\begin{bmatrix} M_1(t) & M_2(t) \\ M_3(t) & M_4(t)\end{bmatrix},\ t \geq 0,$$

[1]There exists a solution means that the solution exists and doesn't blow-up in finite time.

with $M_1(0) = M_4(0) = I_d$ and $M_2(0) = M_3(0) = 0$. This linear differential equation has a unique solution that is well defined for $t \geq 0$, and we define

$$y(t) = M_1(t)X(0) + M_2(t),\ x(t) = M_3(t)X(0) + M_4(t).$$

We observe that

$$\dot{x}(t) = [a(t)M_1(t) - b(t)^\top M_3(t)]X(0) + a(t)M_2(t) - b(t)^\top M_4(t) = a(t)y(t) - b(t)^\top x(t),$$
$$\dot{y}(t) = [b(t)M_1(t) + c(t)M_3(t)]X(0) + b(t)M_2(t) + c(t)M_4(t) = b(t)y(t) + c(t)x(t),$$

Let $\tau = \inf\{t \geq 0, \det(x(t)) = 0\}$, with $\inf \emptyset = +\infty$. For $t \in [0, \tau)$, we define

$$X(t) = y(t)x(t)^{-1}.$$

We have $\dot{X}(t) = \dot{y}(t)x(t)^{-1} - X(t)\dot{x}(t)x(t)^{-1} = b(t)X(t) + c(t) - X(t)[a(t)X(t) - b(t)^\top]$, and therefore X is the solution of (3.16) by the Cauchy-Lipschitz theorem. Also, $X(t)$ is symmetric since $X(t)^\top$ solves the same ODE (3.16).

We now use an argument proposed by Knobloch in [Kno95], and consider $v(t) = x^\top(t)X(t)x(t) = x^\top(t)y(t)$. We have $\dot{x}(t) = [a(t)X(t) - b(t)^\top]x(t)$ and thus

$$\begin{aligned}\dot{v}(t) &= x(t)^\top \left[(X(t)a(t) - b(t))X(t) + \dot{X}(t) + X(t)(2X(t) - b(t)^\top) \right] x(t) \\ &= x(t)^\top c(t)x(t).\end{aligned}$$

From $v(0) = X(0) \in \mathcal{S}_d^+(\mathbb{R})$ and $c(t) \in \mathcal{S}_d^+(\mathbb{R})$, we get that $v(t) \in \mathcal{S}_d^+(\mathbb{R})$ for any $t \in [0, \tau)$. Since $x(t)$ is invertible, this gives $X(t) \in \mathcal{S}_d^+(\mathbb{R})$ for $t \in [0, \tau)$. It remains to prove that the solution X cannot explode. We have $x(0) = I_d$ and $\frac{d}{dt}\det(x(t)) = \det(x(t))\mathrm{Tr}[x(t)^{-1}\dot{x}(t)] = \det(x(t))\mathrm{Tr}[a(t)X(t) - b(t)]$. This yields to

$$t \in [0, \tau),\ \det(x(t)) = \exp\left(\int_0^t \mathrm{Tr}[a(s)X(s) - b(s)]ds \right) \geq \exp\left(- \int_0^t \mathrm{Tr}[b(s)]ds \right),$$

since $a(s) \in \mathcal{S}_d^+(\mathbb{R})$ and thus $\mathrm{Tr}[a(s)X(s)] = \mathrm{Tr}[\sqrt{a(s)}X(s)\sqrt{a(s)}] \geq 0$ for $s \in [0, \tau)$. This inequality necessarily implies $\tau = +\infty$. □

It is sufficient to apply this lemma with $X = -g$ and it follows that the condition:

$$-\Gamma \in \mathcal{S}_d^+(\mathbb{R}) \quad \text{and} \quad \bar{\Gamma} - \frac{1}{2}c\lambda\lambda^T c^T \in \mathcal{S}_d^+(\mathbb{R}), \quad t \geq 0, \tag{3.17}$$

is sufficient for $g(t)$ to exists $t \geq 0$ and $-\gamma(t) \in \mathcal{S}_d^+(\mathbb{R}), t \geq 0$. Note that $\lambda(t)$ admits an analytical expression and thus condition (3.17) can be easily verified.

3.1.4 Numerical resolution of the MRDE with time dependent coefficients.

In this subsection we discuss the numerical resolution of the MRDE (3.9), and present the numerical scheme which has been used in our implementation of the model. MRDE with constant coefficients can be solved explicitly by solving the matrix linear equation (3.13) using matrix exponentials[2]. The linearization method cannot be applied when the coefficients of the MRDE are time dependent. We refer to [DE94] for some results on the properties of the numerical solutions of the MRDE of type (3.9). One of the aspects discussed in the paper is the question of wether the scheme preserves the positiveness of the solution. The paper shows that the only direct schemes that preserve the positiveness of the solution are of order one, and propose an higher order (indirect) scheme which preserves the positiveness of the solution.

The scheme we propose is a simple direct scheme which consists in approximating the time dependent coefficients by piecewise constant functions and solving the MRDE analytically in the intervals where the coefficients are constants. Let $t_0 = 0 < t_1 < \cdots < t_n = T$ be a discretization of the interval $[0, T]$, we define the piecewise constant approximation of $\tilde{\lambda}$ of the vector function $\lambda(t)$ by

$$\tilde{\lambda}(t) = \lambda\left(\frac{t_{i-1} + t_i}{2}\right), \quad t \in [t_{i-1}, t_i), \quad i = 1, ..., n. \tag{3.18}$$

In the interval $[t_{i-1}, t_i)$, we solve analytically the (3.9) replacing λ by $\tilde{\lambda}$ using the linearization method. As mentioned before the method is direct and under the assumptions of proposition 14, the approximated solution is positive defined. On our numerical experiments, a step size between 0.01 and 0.05 is sufficient to achieve a satisfactory convergence of the scheme.

3.1.5 Some identities in law for affine processes

In this section we give some identities in law for affine processes that will be useful to build discretization schemes for the general process. Let us first note that the infinitesimal generator (3.5) only depends on the the parameter a through the matrix $a^T a$, which immediately implies the following identity in law

$$AFF((p, y, K, \theta, c), (d, x, \Omega, b, a), \rho) \underset{Law}{=} AFF((p, y, K, \theta, c), (d, x, \Omega, b, \sqrt{a^T a}), \rho) \tag{3.19}$$

Linear transformations preserve the affine structure of the diffusion process. Let q in $\mathcal{G}_d(\mathbb{R})$ and r in $\mathcal{G}_p(\mathbb{R})$ and define $K_r = rKr^{-1}$, $a_q = aq^\top$, $\theta_r = r\theta$, $c_{rq} rc(q^\top)^{-1}$, $\Omega_q = q^\top \Omega q$,

[2] Let us also mention that in the case of constant coefficients it is also possible to explicitly characterize the definition domain and the explosion time of the solution of the MRDE using sinh and cosh functions, and the eigenvalues of the Hamiltonian matrix.

$b_q = q^\top b (q^\top)^{-1}$ and $a_q = a q^\top$. Let $(X^x, Y^{x,y}) \underset{Law}{=} AFF((p, y, K, \theta, c), (d, x, \Omega, b, a), \rho)$, we have

$$(q^\top X^x q, rY^{x,y}) \underset{Law}{=} AFF((p, ry, K_r, \theta_r, c_{rq}), (d, q^\top x q, \Omega_q, b_q, a_q), \rho).$$

Since the two processes solve the same martingale problem. An interesting consequence of this linear transformation is given in the following corollary. It states that any affine process can be obtained as a linear transformation of an affine process for which we have $a = I_d^n$. This has two major implications: in terms of Monte Carlo simulation it implies that the sampling of the general process reduces to this special case, in terms of term structure modeling it implies that without loss of generality we can restrict ourselves to this particular case of state variables dynamics.

Corollary 18 — *Let $(X^x, Y^{x,y}) \sim AFF((p, y, K, \theta, c), (d, x, \Omega, b, a), \rho)$ and $n = \mathrm{Rk}(a)$ be the rank of $a^T a$. Then, there exist a non singular matrix $u \in \mathcal{G}_d(\mathbb{R})$ such that $a^T a = u^T I_d^n u$. Defining*

$$(\bar{X}_t^x, \bar{Y}_t^{x,y})_{t\geq 0} \underset{Law}{=} AFF((p, y, K, \theta, c), (d, (u^{-1})^\top x u^{-1}, (u^{-1})^\top \Omega u^{-1}, (u^{-1})^\top b u^\top, I_d^n), \rho),$$

we have

$$(X_t^x, Y_t^{x,y})_{t\geq 0} \underset{Law}{=} (u^\top \bar{X}_t^x u, \bar{Y}_t^{x,y})_{t\geq 0}. \tag{3.20}$$

The proof of this corollary is a direct consequence of the identities in law of affine diffusions in the space $\mathcal{S}_d^+(\mathbb{R})$ presented in [AA13].

3.2 Second order discretization schemes for Monte Carlo simulation

This section is devoted to the construction of high order discretization schemes to simulate the underlying state variables of the model. In a high dimensional framework Monte Carlo schemes are necessary for efficient pricing of path dependent products. Note also that while we are able to provide time efficient deterministic methods for pricing vanilla options, these lack of precise error estimates, and their accuracy strongly depends on the set of models parameters. A numerical benchmark is therefore necessary for validation of these methods.

We use the splitting technique that is already used by Ahdida and Alfonsi [AA13] for Wishart processes. We explain here briefly the main line of this method and refer to Alfonsi [Alf10] for precise statements in a framework that embeds affine diffusions. The simulation of Wishart processes and their affine extensions has been studied in details in [AA13]. The key idea is to exploit on one hand the fact any Wishart diffusion is a linear transformation of a canonical

Wishart process of the form $WIS_d(x, \alpha, 0, I_d^n)$, on the other hand some remarkable splitting properties of this canonical process. When the correlation parameter $\rho = 0$, these schemes directly apply to the general process $AFF((p, y, K, \theta, c), (d, x, \Omega, b, a), 0)$. It is not immediate that such splitting property still holds for the generator of the general diffusion process $AFF((p, y, K, \theta, c), (d, x, \Omega, b, a), \rho)$ because of the correlation structure between the vector process Y and the Wishart process X.

3.2.1 Some preliminary results on discretization schemes for SDEs

Let us consider a domain $D \subset \mathbb{R}^\zeta, \zeta \in \mathbb{N}^*$. Let us consider a general $\mathbb{R}^\zeta$ valued SDE:

$$V_t = v + \int_0^t dsb(V_s^v) + \int_0^t \sigma(V_s^v)dW_s. \tag{3.21}$$

In the context of our work the domain of interest will usually be $D = \mathbb{R}^p \times \mathcal{S}_d^+(\mathbb{R})$ viewed as a subspace of $\mathbb{R}^{p+d(d+1)/2}$. For a multi-index $\gamma = (\gamma_1, ..., \gamma_\zeta) \in \mathbb{N}^\zeta$, we define $\partial_\gamma = \partial_1^{\gamma_1}, ..., \partial_\zeta^{\gamma_\zeta}$ and $|\gamma|_1 = \sum_{i=1}^{\zeta} \gamma_i$. We define the set of functions

$$C_{pol}^\infty(D) = \{f \in \mathcal{C}^\infty(D, \mathbb{R}), \forall \gamma \in \mathbb{N}^\zeta, \exists C_\gamma > 0, e_\gamma \in N^*, \forall v \in D, |\partial_\gamma f(v)|_1 \leq C_\gamma(1 + |v|^{e_\gamma}\}$$

where $|\cdot|$ is a norm on $\mathbb{R}^\zeta$. We will say that $(C_\gamma, e_\gamma)_{\gamma\in\mathbb{N}^\zeta}$ is a good sequence for $f \in \mathcal{C}_{pol}^\infty(D)$ if $\partial_\gamma f(v) \leq C_\gamma(1 + |x|^{e_\gamma})$.

Definition 19 — *Let $b : D \to \mathbb{R}^\zeta, \sigma : D \to \mathcal{M}_\zeta(\mathbb{R})$. The operator L defined for $f \in \mathcal{C}^2(D, \mathbb{R})$ by:*

$$Lf(x) = \sum_{i=1}^{\zeta} b_i(v)\partial_i f(v) + \frac{1}{2}\sum_{i,j=1}^{\zeta} (\sigma\sigma^T)_{ij}(v)\partial_{ij}^2 f(v) \tag{3.22}$$

is said to satisfy the required assumptions on D if the following conditions hold:

- $\forall i, j \in \{1, ..., \zeta\}, b_i(v), (\sigma\sigma^T)_{ij}(v) \in \mathcal{C}_{pol}^\infty(D)$,
- *for any $v \in D$, the SDE $V_t^v = v + \int_0^t dsb(V_s^v) + \int_0^t \sigma(V_s^v)dW_s$ has a unique weak solution defined for $t \geq 0$, and thus $P(\forall t \geq 0, V_t^v \in D) = 1$.*

In the case of affine diffusions, $b_i(v)$ and $(\sigma\sigma^T)_{ij}(v)$ are affine functions of x and the operator satisfies the required assumption on the appropriated domain. Let us note that if L statisfies the required assumptions on D and $f \in \mathcal{C}_{pol}^\infty(D)$, all the iterated functions $L^k f(v)$ are well defined on D and belong to $\mathcal{C}_{pol}^\infty(D)$ for any $k \in \mathbb{N}$.

We now turn to discretization schemes for the SDEs. Let us fix a time horizon $T > 0$. We will consider the whole interval $[0, T]$ and the regular time discretization $t_i^N = iT/N, i = 0, 1, ...n$.

Definition 20 — *A family of transition probabilities $(\hat{p}_v(t)(dz), t > 0, v \in D)$ on D is such that $\hat{p}_v(t)$ is a probability law on D for $t > 0$ and $v \in D$.*

A discretization scheme with transition probabilities $(\hat{p}_v(t)(dz), t > 0, x \in D)$ is a sequence $(\hat{V}^N_{t_i^N}, 0 \leq i \leq N)$ of D-valued random variables such that:

- *for $0 \leq i \leq N$, $\hat{V}^N_{t_i^N}$ is a $\mathcal{F}_{t_i^N}$-measurable random variable on D*
- *the law of $\hat{V}^N_{t_{i+1}^N}$ given the information $\mathcal{F}_{t_i^N}$ is given by $\mathbb{E}\left[f(V^N_{t_{i+1}^N})|\mathcal{F}_{t_i^N}\right] = \int_D f(z)\hat{p}_{\hat{V}^N_{t_i^N}}(T/N)(dz)$ and thus only depends on $\hat{V}^N_{t_i^N}$ and T/N.*

For ease of notations, for $t > 0$ and $v \in D$ we will denote by $\hat{V}^v_t$ a r.v. distributed according to the probability $\hat{p}_v(t)(dz)$. Since the law of a discretization scheme is entirely determined by its initial value and by the transition probabilities. We will always take the initial value to be equal to the initial value of the SDE, by a slight abuse of language we will identify the scheme $\hat{V}^N_{t_i^N}$ with its transition probabilities ($\hat{p}_v(t)(dz)$ or $\hat{V}^v_t$).

Definition 21 — *Assume that the operator L associated to the SDE (3.21) verifies the required assumptions. Let us denote $\mathcal{C}^\infty_K(D,\mathbb{R})$ the set of the $\mathcal{C}^\infty$ real valued functions with a compact support in D. Let $v \in D$. A discretization scheme $(\hat{V}^N_{t_i^N}, 0 \leq i \leq N)$ is a weak ν^{th}-order scheme for the SDE (3.21) if:*

$$\forall f \in \mathcal{C}^\infty_K(D,R), \exists K > 0 : \left|\mathbb{E}\left(f(V^v_T)\right) - \mathbb{E}\left(f(V^N_{t_N^N})\right)\right| \leq K/N^\nu.$$

The quantity $\mathbb{E}\left(f(V^v_T)\right) - \mathbb{E}\left(f(V^N_{t_N^N})\right)$ is called the weak error associated to f.

It is hard to prove that a scheme is a weak ν^{th}-order scheme directly, the useful notion in practice is the notion of potential weak ν^{th}-order scheme.

Definition 22 — *Assume that the operator L associated to the SDE (3.21) verifies the required assumptions. A scheme $(\hat{p}_v(t)(dz), t > 0, v \in D)$ is a potential weak ν^{th}-order scheme for the operator L if for any function $f \in \mathcal{C}^\infty_{pol}(D)$ with a good sequence $(C_\gamma, e_\gamma)_{\gamma\in\mathbb{N}^\gamma}$, there exist positive constants C, E, and η depending on (C_γ, e_γ) only such that*

$$\forall t \in (0,\eta), \left|\mathbb{E}\left[f(\hat{V}^v_t)\right] - \left[f(v) + \sum_{k=1}^{\nu} \frac{1}{k!} t^k L^k f(v)\right]\right| \leq C t^{\nu+1}(1 + |v|^E). \tag{3.23}$$

The following theorem of [Alf10], which is a consequence of the results of [TT90], links the notion of potential weak ν^{th}-order scheme and the weak error of the scheme. It provides the technical conditions under which a potential weak ν^{th} order scheme is sufficient to control the weak error.

Theorem 23 — *Let us consider an operator L that satisfies the required assumptions on D and a discretization scheme $(\hat{V}^N_{t^N_i}, 0 \leq i \leq N)$ with transition probabilities $\hat{p}_v(t)(dz)$ on D that starts from $V^N_{t^N_0} = v_0 \in D$. We assume that*

1. *$\hat{p}_v(t)(dz)$ is a potential weak ν^{th} order scheme for the operator L*
2. *the scheme has bounded moments, i.e.*
$$\forall q \in \mathbb{N}^*, \exists N(q) \in \mathbb{N} \quad : \quad \sup_{N \geq N(q), 0 \leq i \leq N} \mathbb{E}\left[\left|V^N_{t^N_i}\right|\right] < \infty \tag{3.24}$$
3. *$f : D \to \mathbb{R}$ is a function such that $u(t,v) = \mathbb{E}\left[f(V^v_{T-t})\right]$ is defined on $[0,T] \times D$, $\mathcal{C}^\infty$, solves $\forall t \in [0,T]$, $\forall v \in D$, $\partial_t u(t,v) = -Lu(t,v)$, and satisfies :*
$$\forall l \in \mathbb{N}, \gamma \in \mathbb{N}^\zeta, \exists C_{l,\gamma}, e_{l,\gamma} > 0 \quad : \quad \forall v \in D, t \in [0,T], |\partial^l_t \partial_\gamma u(t,v)| \leq C_{l,\gamma}(1 + |v|^{e_{l,\gamma}}). \tag{3.25}$$

Then, there is $K > 0$ and $N_0 \in \mathbb{N}$, such that $\left|\mathbb{E}\left[f(V^N_{t^N_N})\right] - \mathbb{E}\left[f(V^v_T)\right]\right| \leq K/N^\nu$ for $N \geq N_0$.

In the following we will propose some potential second order schemes for the state variables (3.1). In order to prove that these schemes allow to control the weak error one would need to verify the technical conditions (3.24) and (3.25) of the above theorem. In [AA13] these conditions have been verified for Wishart processes. We will not verify these conditions for the schemes we propose. The verification of the conditions (3.24) and (3.25) for the process (X,Y) would require some tedious calculations, and we don't value it as a crucial point of our work. Instead we privilege an empirical approach and investigate the question numerically. This provides extensive numerical proof of the control of the weak error.

The last important result for our purpose is the fact the property of being a potential ν^{th} order scheme is preserved by composition of the schemes. Let us consider two schemes $\hat{p}^1_v(t)(dz)$ and $\hat{p}^2_v(t)(dz)$, the composition of the second scheme with the first $\hat{p}^2(t_2) \circ \hat{p}^1_v(t_1)(dz)$ is defined by

$$\hat{p}^2(t_2) \circ \hat{p}^1_v(t_1)(dz) = \int_D \hat{p}^2_y(t_2)(dz)\hat{p}^1_v(t_1)(dy). \tag{3.26}$$

More generally we can define the composition of m discretization schemes $\hat{p}^1_v, ..., \hat{p}^m_v$ on D by

$$\hat{p}^m(t_m) \circ \cdots \circ \hat{p}^1_v(t_1)(dz) = \hat{p}^m(t_m) \circ \left(\hat{p}^{m-1}(t_{m-1}) \circ \cdots \circ \hat{p}^1_v(t_1)(dz)\right).$$

Proposition 24 — *Let L_1, L_2 be two operators satisfying the required assumptions on D. Let $\hat{p}^1_v$ and $\hat{p}^2_v$ be respectively two potential weak ν^{th} order schemes on D for L_1 and L_2.*

- *If $L_1 L_2 = L2 L_1$ then $\hat{p}^2(t) \circ \hat{p}^1_v(t)(dz)$ is a potential weak ν^{th} order scheme for $L1 + L_2$.*

- *If $\nu \geq 2$, $\hat{p}^2(t/2) \circ \hat{p}^1(t) \circ \hat{p}^2_v(t/2)(dz)$ and $\frac{1}{2}\left(\hat{p}^1(t) \circ \hat{p}^2_v(t)(dz) + \hat{p}^2(t) \circ \hat{p}^1_v(t)(dz)\right)$ are potential weak second order schemes for $L_1 + L_2$.*

The above proposition gives a straightforward way to build second order schemes for an operator L as soon as we can split it in simple operators for which we have a potential weak order scheme of order higher then two. Let us assume that $\hat{\xi}^{i,x}_t$ is a second order scheme for L_i. Let B be an independent Bernoulli variable with parameter 1/2. Then, the following schemes

$$\hat{\xi}_{t/2}^{1,\hat{\xi}_t^{2,\hat{\xi}^{1,x}_{t/2}}} \text{ and } B\hat{\xi}_t^{2,\hat{\xi}_t^{1,x}} + (1-B)\hat{\xi}_t^{1,\hat{\xi}_t^{2,x}} \tag{3.27}$$

satisfy

$$\mathbb{E}[f(\hat{\xi}^x_t)] = f(x) + tLf(x) + \frac{t^2}{2}L^2 f(x) + O(t^3) \tag{3.28}$$

and are thus second order schemes for L. Therefore, a strategy to construct a second order scheme is to split the infinitesimal generator into elementary pieces for which second order schemes or even exact schemes are known. The underlying idea of the discretization schemes we propose is to find some splitting property of the operator of the diffusion (3.1) and then build the scheme using the composition rule (3.27).

3.2.2 High order discretization schemes for L

Let us first note that the identity in law (3.20) implies that the sampling of the general process

$$AFF((p,y,K,\theta,c),(d,x,\Omega,b,a),\rho)$$

reduces to the sampling of the process

$$AFF\left((p,y,K,\theta,c),(d,x,\Omega,b,I^n_d),0\right).$$

In all the following let (X,Y) follow $AFF\left((p,y,K,\theta,c),(d,x,\Omega,b,I^n_d),\rho\right)$.

To use this splitting technique, we first have to calculate the infinitesimal generator of (X,Y). From (3.5), we easily get

$$\begin{aligned} L = &\sum_{m=1}^{p}(K(\theta-y))_m\partial_{y_m} + \sum_{1\leq i,j\leq d}(\Omega+(d-1)I^n_d+bx+xb^\top)_{i,j}\partial_{x_{i,j}} \\ &+\frac{1}{2}\sum_{m,m'=1}^{p}(cxc^\top)_{m,m'}\partial_{y_m}\partial_{y_{m'}} + \frac{1}{2}\sum_{m=1}^{p}\sum_{1\leq i,j\leq d}[(cx)_{m,i}(I^n_d\rho)_j + (cx)_{m,j}(I^n_d\rho)_i]\partial_{x_{i,j}}\partial_{y_m} \\ &+\frac{1}{2}\sum_{1\leq i,j,k,l\leq d}[x_{i,k}(I^n_d)_{j,l} + x_{i,l}(I^n_d)_{j,k} + x_{j,k}(I^n_d)_{i,l} + x_{j,l}(I^n_d)_{i,k}]\partial_{x_{i,j}}\partial_{x_{k,l}}. \end{aligned}$$

When $\rho = 0$, this operator is simply the sum of the infinitesimal generators for X and the generator for Y when X is frozen. We know from [AA13] a second order scheme for X. When X is frozen, Y follows an Ornstein-Uhlenbeck process and the law of Y_t is a Gaussian vector that can be sampled exactly. By using the composition rule (3.27), we get a second order scheme for (X, Y).

Thus, the difficulty here comes from the correlation between X and Y that has to be handled with care. We first make some simplifications. The first term $\sum_{m=1}^{p}(K(\theta-y))_m\partial_{y_m}$ is the generator of the linear Ordinary Differential Equation $y'(t) = K(\theta - y(t))$ that is solved exactly by $y(t) = e^{-Kt}y(0) + (I_p - e^{-Kt})\theta$. Therefore, it is sufficient to have a second order scheme for $L - \sum_{m=1}^{p}(K(\theta-y))_m\partial_{y_m}$, which is the generator of $AFF\,((p, y, 0, 0, c), (d, x, \Omega, b, I_d^n), \rho)$. In which case, we have $Y_t = y + c(\tilde{Y}_t - \tilde{Y}_0)$ with

$$\tilde{Y}_t = \tilde{Y}_0 + \int_0^t \sqrt{X_s}\,[\bar{\rho}dZ_s + dW_s\rho]$$

We can then focus on getting a second order scheme for $(X, \tilde{Y})$, which amounts to work with $p = d$ and $c = I_d$. It is therefore sufficient to find a second order scheme for the SDE

$$\begin{aligned} Y_t &= y + \int_0^t \sqrt{X_s}\,[\bar{\rho}dZ_s + dW_s\rho]\,, \\ X_t &= x + \int_0^t \left(\Omega + (d-1)I_d^n + bX_s + X_sb^\top\right)\right) ds + \int_0^t \sqrt{X_s}dW_sI_d^n + I_d^ndW_s^\top\sqrt{X_s}, \end{aligned}$$

with the infinitesimal generator

$$\begin{aligned} L = &\sum_{1\le i,j\le d}(\Omega + (d-1)I_d^n + bx + xb^\top)_{i,j}\partial_{x_{i,j}} + \frac{1}{2}\sum_{m=1}^{d}\sum_{1\le i,j\le d}[x_{m,i}(I_d^n\rho)_j + x_{m,j}(I_d^n\rho)_i]\partial_{x_{i,j}}\partial_{y_m} \\ &+ \frac{1}{2}\sum_{m,m'=1}^{d}x_{m,m'}\partial_{y_m}\partial_{y_{m'}} + \frac{1}{2}\sum_{1\le i,j,k,l\le d}[x_{i,k}(I_d^n)_{j,l} + x_{i,l}(I_d^n)_{j,k} + x_{j,k}(I_d^n)_{i,l} + x_{j,l}(I_d^n)_{i,k}]\partial_{x_{i,j}}\partial_{x_{k,l}}. \end{aligned} \tag{3.29}$$

3.2.3 A second order scheme

For $1 \le q \le d$, we define $e_d^q \in \mathcal{S}_d^+(\mathbb{R})$ by $(e_d^q)_{k,l} = \mathbb{1}_{k=l=q}$ and $g_d^q \in \mathbb{R}^d$ by $(g_d^q)_k = \mathbb{1}_{q=k}$ so that $I_d^n = \sum_{q=1}^{n} e_d^q$ and $I_d^n\rho = \sum_{q=1}^{n}\rho_q g_d^q$. We define

$$\begin{aligned} L_q^c = &(d-1)\partial_{x_{q,q}} + \frac{1}{2}\sum_{m=1}^{d}\sum_{1\le i,j\le d}\rho_q[x_{m,i}(g_d^q)_j + x_{m,j}(g_d^q)_i]\partial_{x_{i,j}}\partial_{y_m} + \frac{\rho_q^2}{2}\sum_{m,m'=1}^{d}x_{m,m'}\partial_{y_m}\partial_{y_{m'}} \\ &+ \frac{1}{2}\sum_{1\le i,j,k,l\le d}[x_{i,k}(e_d^q)_{j,l} + x_{i,l}(e_d^q)_{j,k} + x_{j,k}(e_d^q)_{i,l} + x_{j,l}(e_d^q)_{i,k}]\partial_{x_{i,j}}\partial_{x_{k,l}}. \end{aligned} \tag{3.30}$$

We consider the splitting $L = L' + L'' + \sum_{q=1}^{n} L_q^c$ of the operator (3.29), with

$$L' = \sum_{1\le i,j\le d} (\Omega + bx + xb^\top)_{i,j}\partial_{x_{i,j}},$$
$$L'' = \left(1 - \sum_{q=1}^{n} \rho_q^2\right) \frac{1}{2} \sum_{m,m'=1}^{d} x_{m,m'}\partial_{y_m}\partial_{y_{m'}}.$$

The operator L' is the one of the linear ODE $x'(t) = \Omega + (d-1)I_d^n + bx + xb^\top$ that can be solved exactly and stays in the set of semidefinite positive matrices, see Lemma 27 in [AA13]. The operator L'' is the one of $Y_t'' = y'' + \sqrt{1 - \sum_{q=1}^{n} \rho_q^2}\sqrt{x}Z_t$, which can be sampled exactly since it is a Gaussian vector with mean y'' and covariance matrix $(1-|\rho|^2)tx$. The operator L_q^c is the infinitesimal generator of the following SDE

$$\begin{cases} Y_t &= y + \rho_q \int_0^t \sqrt{X_s} dW_s g_d^q, \\ X_t &= x + \int_0^t (d-1)e_d^q ds + \int_0^t \sqrt{X_s} dW_s e_d^q + e_d^q dW_s^\top \sqrt{X_s}. \end{cases} \tag{3.31}$$

Thus, X follows an elementary Wishart process and stays in $\mathcal{S}_d^+(\mathbb{R})$. Using the notation of [AA13], X_t follows the law $WIS_d(x, d-1, 0, e_d^q, t)$. Theorems 9 and 16 in [AA13] gives respectively an exact and a second (or higher) discretization scheme for this process. We now explain how to calculate Y_t once that X_t has been sampled. From (3.31), we have for $1 \le i \le d$,

$$d(Y_t)_i = \rho_q \sum_{j=1}^{d} (\sqrt{X_t})_{i,j}(dW_t)_{j,q},$$
$$d(X_t)_{q,i} = \sum_{j=1}^{d} (\sqrt{X_t})_{i,j}(dW_t)_{j,q} + \mathbb{1}_{i=q}\left[(d-1)dt + \sum_{j=1}^{d} (\sqrt{X_t})_{q,j}(dW_t)_{j,q}\right].$$

This yields to

$$(Y_t)_i = y_i + \rho_q((X_t)_{q,i} - x_{q,i}), \text{ if } i \neq q,$$
$$(Y_t)_q = y_i + \frac{\rho_q}{2}[(X_t)_{q,q} - x_{q,q} - (d-1)t].$$

Using these formula together with the exact (resp. second order) scheme for X_t, we get an exact (resp. second order) scheme for (3.31). By using the composition rules (3.27), we get a second order scheme for (3.29).

3.2.4 A faster second order scheme when $\Omega - I_d^n \in \mathcal{S}_d^+(\mathbb{R})$

As explained in [AA13], the sampling of each elementary Wishart process in L_q requires a Cholesky decomposition that has a time complexity of $O(d^3)$. Since the second order scheme proposed above calls $n \le d$ times this routine, the whole scheme requires at most $O(d^4)$ operations. However, by adapting an idea that has been already used in [AA13] for Wishart processes, it is possible to get a faster scheme if we assume in addition that $\Omega - I_d^n$. This alternative scheme only requires $O(d^3)$ operations and is also easier to implement.

We consider the splitting $L = \tilde{L}' + \tilde{L}'' + \hat{L}$ of the operator (3.29), with

$$\tilde{L}' = \sum_{1\le i,j\le d} (\Omega - I_d^n + bx + xb^\top)_{i,j}\partial_{x_{i,j}}$$

$$\hat{L} = \sum_{1\le i\le n} d\partial_{x_{i,i}} + \frac{1}{2}\sum_{m=1}^{d}\sum_{1\le i,j\le d}[x_{m,i}(I_d^n\rho)_j + x_{m,j}(I_d^n\rho)_i]\partial_{x_{i,j}}\partial_{y_m} + \frac{\sum_{q=1}^n \rho_q^2}{2}\sum_{m,m'=1}^{d} x_{m,m'}\partial_{y_m}\partial_{y_{m'}}$$
$$+\frac{1}{2}\sum_{1\le i,j,k,l\le d}[x_{i,k}(I_d^n)_{j,l} + x_{i,l}(I_d^n)_{j,k} + x_{j,k}(I_d^n)_{i,l} + x_{j,l}(I_d^n)_{i,k}]\partial_{x_{i,j}}\partial_{x_{k,l}}.$$

Again, $\tilde{L}'$ is the operator of the linear ODE $x'(t) = \Omega - I_d^n + (d-1)I_d^n + bx + xb^\top$ that can be solved exactly and stays in the set of semidefinite positive matrices by Lemma 27 in [AA13] since $\Omega - I_d^n \in \mathcal{S}_d^+(\mathbb{R})$. We have already seen above that the generator L'' can be sampled exactly, and we focus now on the sampling of $\hat{L}$. It relies on the following result.

Lemma 25 — *For $x \in \mathcal{S}_d^+(\mathbb{R})$ we consider $c \in \mathcal{M}_d(\mathbb{R})$ such that $c^\top c = x$. We define $U_t = c + W_t I_d^n$, $X_t = U_t^\top U_t$ and $Y_t = y + \int_0^t M_s^\top dW_s I_d^n \rho$. Then, the process (X, Y) has the infinitesimal generator $\hat{L}$.*

Proof : For $1 \le i,j,m \le d$, we have $d(X_t)_{i,j} = \sum_{k=1}^d ((U_t)_{k,i}(dW_t)_{k,j}\mathbb{1}_{j\le n} + (U_t)_{k,j}(dW_t)_{k,i}\mathbb{1}_{i\le n}) + \mathbb{1}_{i=j\le n}ddt$ and $d(Y_t)_m = \sum_{k,l=1}^d (U_t)_{k,m}(dW_t)_{k,l}(I_d^n\rho)_l$. This leads to

$$\begin{aligned}
\langle d(Y_t)_m, d(Y_t)_{m'}\rangle &= \sum_{k,l=1}^d (U_t)_{k,m}(U_t)_{k,m'}(I_d^n\rho)_l^2 dt = \left(\sum_{l=1}^n \rho_l^2\right)(X_t)_{m,m'}dt,\\
\langle d(Y_t)_m, d(X_t)_{i,j}\rangle &= [(I_d^n\rho)_j(X_t)_{m,i} + (I_d^n\rho)_i(X_t)_{m,j}]dt,\\
\langle d(X_t)_{i,j}, d(X_t)_{k,l}\rangle &= [(X_t)_{i,k}(I_d^n)_{j,l} + (X_t)_{i,l}(I_d^n)_{j,k} + (X_t)_{j,k}(I_d^n)_{i,l} + (X_t)_{j,l}(I_d^n)_{i,k}]dt,
\end{aligned}$$

which precisely gives the generator $\hat{L}$. □

Thanks to Lemma 25, it is sufficient to construct a second order scheme for (U, Y). Since $\langle d(Y_t)_m, d(U_t)_{i,j}\rangle = \epsilon(U_t)_{i,m}(I_d^n\rho)_j dt$, the infinitesimal generator $\bar{L}$ of (U, Y) is given by

$$\bar{L} = \frac{1}{2}\sum_{i=1}^d\sum_{j=1}^n \partial^2_{u_{i,j}} + \frac{1}{2}\sum_{i,m=1}^d\sum_{j=1}^n \rho_j u_{i,m}\partial_{u_{i,j}}\partial_{y_m} + \frac{\sum_{q=1}^n \rho_q^2}{2}\sum_{m,m'=1}^d (u^\top u)_{m,m'}\partial_{y_m}\partial_{y'_m}.$$

We use now the splitting $\bar{L} = \sum_{q=1}^{n} \bar{L}_q$ with

$$\bar{L}_q = \frac{1}{2}\sum_{i=1}^{d} \partial^2_{u_{i,q}} + \frac{1}{2}\sum_{i,m=1}^{d} \rho_q u_{i,m}\partial_{u_{i,q}}\partial_{y_m} + \frac{\rho_q^2}{2}\sum_{m,m'=1}^{d}(u^\top u)_{m,m'}\partial_{y_m}\partial_{y'_m}.$$

By straightforward calculus, we find that $\bar{L}_q$ is the generator of the following SDE

$$dY_t = \rho_q U_t^\top dW_t g_d^q,\ dU_t = dW_t e_d^q.$$

We note that only the q^{th} row of U is modified. For $1 \le i \le d$ we have $d(U_t)_{i,q} = (dW_t)_{i,q}$ and $d(Y_t)_m = \rho_q \sum_{j=1}^{d}(U_t)_{j,m}(dW_t)_{j,q}$. This yields to

$$(Y_t)_m = (Y_0)_m + \rho_q \sum_{j=1}^{d}(U_0)_{j,m}(W_t)_{j,q} \text{ for } m \neq q,$$

$$(Y_t)_q = (Y_0)_q + \rho_q \sum_{j=1}^{d}(U_0)_{j,q}(W_t)_{j,q} + \frac{\rho_q}{2}\sum_{j=1}^{d}\{(W_t)^2_{j,q} - t\}.$$

By using these formulas, we can then sample exactly (U_t, Y_t) and then get a second order scheme for $\hat{L}$. We note that the simulation cost of $\bar{L}_q$ requires $O(d)$ operations and then the one of $\bar{L}$ requires $O(d^2)$ operations. Since a matrix multiplication requires $O(d^3)$ operations, this second order schemes for $\bar{L}$ and then for L requires $O(d^3)$ operations instead of $O(d^4)$ for the scheme described in Subsection 3.2.3.

Remark 26 — *As already mentioned, the dependence between the processes X and Y is the same as the one proposed by Da Fonseca, Grasselli and Tebaldi [DFGT08a] for a model on asset returns. Therefore, we can use the same splittings as the one proposed in Subsections 3.2.3 and 3.2.4 to get a second order scheme for their model.*

3.2.5 Numerical results

We now turn to the empirical analysis of the convergence of the discretization schemes we have proposed. Unlike the case of the Wishart process we were not able to provide an exact scheme for our state variables dynamics (3.1). Also we cannot achieve schemes of order higher then 2 since we are composing operators that don't commute. In this paragraph we adopt the following terminology,

- Scheme 1 is the second order scheme given in Subsection 3.2.3, where we use the exact sample of the Wishart part and the exact simulation the Gaussian variables.
- Scheme 2 is the second order scheme given in Subsection 3.2.3, where we use the second order scheme for the Wishart part and replacing the simulation of Gaussian variables by random variables that matches the five firs moments, see Theorem 16 and equation (36) in [AA13].

- Scheme 3 is the second order scheme given in Subsection 3.2.4.

In order to assess that the potential second order schemes we have proposed for L are indeed second order schemes we start by analyzing the weak error for quantities we can compute analytically such as the Fourier transform $\mathbb{E}\left[\exp\left(-i\left(\mathrm{Tr}(\Gamma X_T) + \Lambda^T Y_T\right)\right)\right]$. We compare the values obtained by Monte Carlo simulation and the values obtained by solving the system of o.d.e (3.9). In the next chapter the computation of option prices will give us the occasion to further investigate the convergence of the schemes for more complicated functions of the underlying state variables.

As shown by Figures 3.1, 3.2 and 3.3 the weak error of the Fourier transform for the 3 schemes we have considered is converges in $O(1/N^2)$. In terms of performance, scheme 1 seems to converge much faster then the others, this is due to the fact we use the exact simulation of the Wishart part of the generator L in this scheme and this certainly increases its accuracy. The scheme 3, which is extremely cost-efficient compared to the other schemes, has a performance which is comparable to scheme 2. Let us note that as illustrated by Figure 3.3 the convergence of the schemes is much slower when we introduce consider the full process with a mean reverting drift.

Unfortunately we were not able to investigate the convergence of the schemes for path-wise expectations. Very few theoretical results exist on this topic and since we don't have an exact scheme for the simulation of L, we are not able to conduce a numerical investigation on the subject.

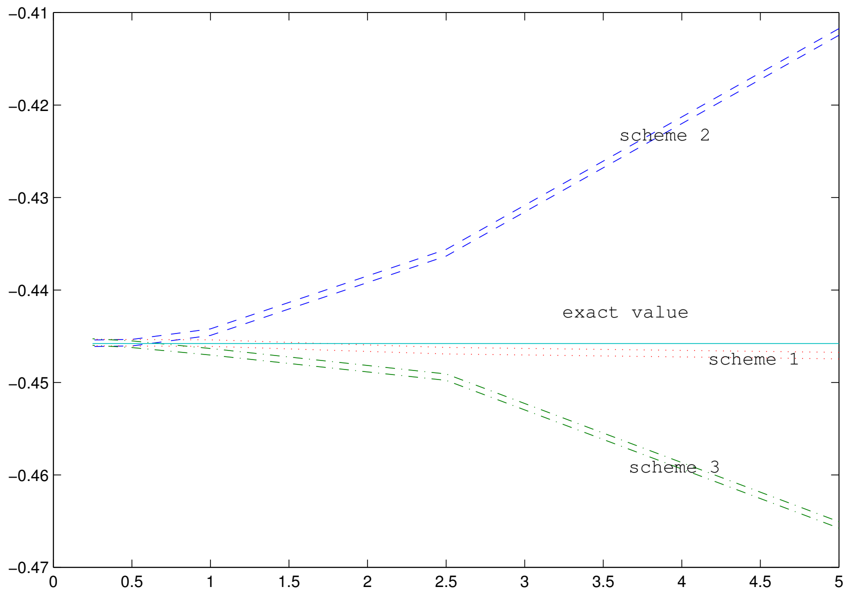

Figure 3.1: Convergence of the discretization schemes. $p = d = 3$, 10^7 Monte Carlo samples, $T = 5$. The real value of $\mathbb{E}\left[\exp\left(-i\left(\mathrm{Tr}(\Gamma X_T) + \Lambda^T Y_T\right)\right)\right]$, as a function the time step T/N. $\Gamma = 0.05I_d, \Lambda = 0.021_{\mathbb{R}^d}$ and the diffusion parameters $x = 0.4I_d, y = 0.2\mathbf{1}_p, \alpha = 4.5, a = I_d, \rho = 0, b = 0, K = 0, c = I_d, A = 0$. The value obtained by solving the o.d.e is -0.445787. For each scheme the two curves represent the upper and lower bound of the 95% confidence interval.

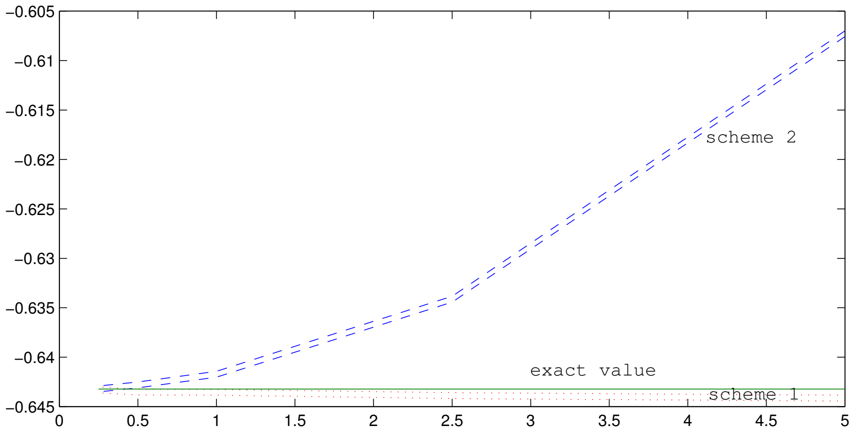

Figure 3.2: Convergence of the discretization schemes. $p = d = 3$, 10^7 Monte Carlo samples, $T = 5$. The imaginary value of $\mathbb{E}\left[\exp\left(-i\left(\mathrm{Tr}(\Gamma X_T) + \Lambda^T Y_T\right)\right)\right]$, as a function the time step T/N. $\Gamma = 0.05 I_d, \Lambda = 0.02\mathbf{1}_p$ and the diffusion parameters $x = 0.4 I_d, y = 0.2 1_{\mathbb{R}^d}, \alpha = 2.5, a = I_d, \rho = -0.3\mathbf{1}_p, b = 0, K = 0, c = I_d, A = 0$. The value obtained by solving the o.d.e is -0.643222. For each scheme the two curves represent the upper and lower bound of the 95% confidence interval.

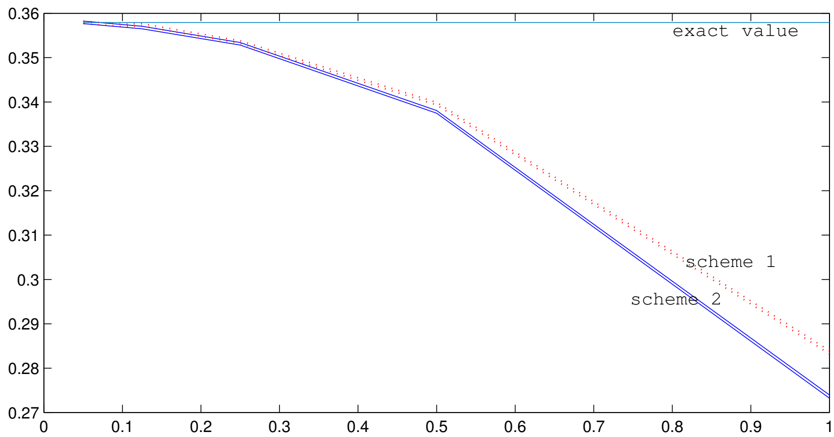

Figure 3.3: Convergence of the discretization schemes. $p = d = 3$, 10^7 Monte Carlo samples, $T = 1$. The real value of $\mathbb{E}\left[\exp\left(-i\left(\mathrm{Tr}(\Gamma X_T) + \Lambda^T Y_T\right)\right)\right]$, as a function the time step T/N. $\Gamma = 0.2 I_d + 0.04 q, \Lambda = 0.2$ and the diffusion parameters $x = 0.4 I_d + 0.2 * q, y = 0.2\mathbf{1}_p, \alpha = 2.5, a = I_d, \rho = -0.3\mathbf{1}_p, b = -0.5 I_d, K = 0.1 I_p, c = I_d, A = 0$. The value obtained by solving the o.d.e is 0.357901. For each scheme the two curves represent the upper and lower bound of the 95% confidence interval.

Chapter 4

A stochastic variance-covariance affine term structure model

In this chapter we present the model that we study in this thesis. The model is presented in a mono-curve setting. This model is a stochastic variance-covariance perturbation of the LGM (see section 2.4.1). We chose a quite general specification that keeps the model affine and gives a stochastic instantaneous covariance for the factors, which will generate a smile for the Caplets and Swaptions, as well as a stochastic correlation between rates. The resulting model belongs to the general class of ATSM defined by [DK96b]. The affine structure of the model makes it reasonably tractable. We derive semi-analytical bond reconstruction formulas for zero-coupon bonds and for Fourier/Laplace transforms of asset prices under the equivalent martingale measures. Recently Wishart processes have been used to model stochastic volatility of interest rates. The first proposals were formulated in [GMS10], [GS03], [GS07], [Gou06] both in discrete and continuous time, for interest rates and equity underlyings. Applications to multifactor volatility and stochastic correlation can be found in [DFGT08b], [DFGT08a], [DFGI11], [DFGI14], [DFG11], [BPT10], [BL13], [BCT08], both in option pricing and portfolio management. These contributions consider the case of continuous path Wishart processes, [GT08] and [Cuc11] investigate processes lying in the more general symmetric cones state space domain, including the interior of the cone $\mathcal{S}_d^+(\mathbb{R})$. A Wishart ATSM has been defined by Benabid, Bensusan and El Karoui and studied by Bensusan in his Phd thesis [Ben10]. Their model is inspired by the canonical specification of ATSM given by [DS00], and uses Wishart processes to define a rich volatility dynamics. The model we study here differs in that: first we use the Wishart-like process to model stochastic variance-covariance, and secondly we consider a correlation between the factors and the drivers of the volatility. In a series of papers by Gnoatto [Gno13] and Gnoatto and Grasselli, [GG14a], and Gnoatto et al. [BGH13] have considered affine term structure models, in which the yield curve is driven by a Wishart process. A mean reverting Wishart model for the log-prices of commodities has been considered in [CWZ15].

Section 3.2 provides an efficient Monte Carlo simulation framework for the model. We also investigate less time consuming numerical methods for pricing vanilla interest rates options.

Transform methods, and asymptotic expansion of the density through Gram-Charlier and Edgeworth series prove to be very efficient due to the affine structure of the model. The semi-analytical expressions of the Laplace/Fourier transforms allow an efficient computation of the moments of asset prices. Unfortunately these methods lack of precise error estimates. We apply perturbation techniques to derive the volatility smile at order two in the volatility-of-volatility. This allows for an explicit representation of the volatility smile, allowing for full understanding of the model parameters and factors role. The arguments we use to obtain the expansion have been developed in the book of Fouque et al. [FPS00]. They rely on an expansion of the infinitesimal generator. Recently, this technique was applied by Bergomi and Guyon [BG12] to provide approximation under a multi factor model for the forward variance. Here, we have to take into account some specific features of the fixed income and work under the appropriate probability measure to apply these arguments.

The chapter is organized as follows: section 4.1 defines the model, discusses the specification, provides the bond reconstruction formulas and characterizes the distribution of the state variables under the forward measures. Section 4.2 presents two fast numerical methods for pricing caplets and swaptions. Namely the Fourier inversion and the Fast Fourier Transform (FFT) methods, and the series expansion methods of Gram-Charlier and Edgeworth type. Section 4.3 presents the expansion for caplets and swaptions. Finally section 4.4 discusses the hedging in the model.

4.1 Model definition

Since [DK96b] it is well known that to define an ATSM one only needs to consider a set of state variables following an affine diffusion process under the risk neutral measure and assume that the short rate is an affine function of these variables. We define the stochastic variance covariance ATSM (SCVATSM) as a short rate model driven by state variables with distribution $AFF\left((p, y, K, \theta, c), (d, x, \Omega, b, \epsilon I_d^n), \rho\right)$ under the risk-neutral measure. The specification of the model is designed to allow for an interpretation of the model factors. While keeping the the interpretation of the yield curve factors as described in the LGM model, we also identify the volatility factors that drive the implied volatility of swaptions and caplets. The model exhibits a volatility smile allowing for market option prices calibration and stochastic correlation between rates. These features make the model suitable for exotic options pricing and option books management. We tried to reduce the number of parameters of the model as much as possible. Note that the numerical framework that will be developed can be applied to any asset pricing model driven by a general dynamics of the type $AFF((p, y, K, \theta, c), (d, x, \Omega, b, a), \rho)$. However, we will show in subsection 4.1.1 that considering a more general specification would not make the model richer.

Definition 27 — *We assume that* $(X_t, Y_t)_{t\geq 0}$ *follows* $AFF\left((p, y, K, \theta, c), (d, x, \Omega, b, \epsilon I_d^n), \rho\right)$

under a risk-neutral measure. Then, we define the short interest rate by

$$r_t = \varphi + \sum_{i=1}^{p} Y_t^i + \mathrm{Tr}\left(\gamma X_t\right), \tag{4.1}$$

with $\varphi \in \mathbb{R}$ *and* $\gamma \in \mathcal{S}_d(\mathbb{R})$.
In the following we denote by $M\left\{\varphi, \gamma, (p, y, K, \theta, c), (d, \Omega, b, n, \epsilon), \rho\right\}$ *the stochastic variance-covariance affine term structure model (SCVATSM) defined by* (4.1).

In the LGM model, it is generally assumed distinct speed of mean-reversion for each factors. This leads to assume that

$$K = \mathrm{diag}(k_1, \dots, k_p) \text{ with } 0 < k_1 < \cdots < k_p,$$

and we work under this assumption in the sequel. It can be easily checked (see for example Andersen and Piterbarg [PA10]) that any linear Gaussian model such that K has distinct positive eigenvalues can be rewritten, up to a linear transformation of the factors, within the present parametrization. From Proposition 14, we easily get the following result on the Zero-Coupon bonds.

Corollary 28 ***Bond reconstruction formula****. Let* $0 \leq t \leq T$ *and* $P_{t,T} = \mathbb{E}[\exp(-\int_t^T r_s ds)|\mathcal{F}_t]$ *denote the price at time* t *of a zero-coupon bond with maturity* T*. Let us assume that*

$$\gamma - \frac{1}{2}\left(\sum_{i=1}^{p} \frac{1}{k_i^2}\right) c^\top c \in \mathcal{S}_d^+(\mathbb{R}). \tag{4.2}$$

Then, $P_{t,T}$ *is given by*

$$P_{t,T} = \exp(A(T-t) + \mathrm{Tr}(D(T-t)X_t) + B(T-t)^\top Y_t), \tag{4.3}$$

with $A(t) = \eta(t) - \varphi t$*,* $D(t) = g(t)$ *and* $B(t) = \lambda(t)$*, where* (η, g, λ) *is the solution of* (3.9) *with* $\Lambda = 0$*,* $\Gamma = 0$*,* $\bar{\Gamma} = -\gamma$ *and* $\bar{\Lambda} = -\mathbf{1}_p$ *(i.e.* $\bar{\Lambda}_i = -1$ *for* $1 \leq i \leq p$*). In particular, we have* $-D(T-t) \in \mathcal{S}_d^+(\mathbb{R})$.

Let us make now some comments on the model.

- In order to keep the same factors as in the LGM, one would like to take $\gamma = 0$. However, this choice is possible only if the perturbation around the LGM is small enough provided that $-(b + b^\top)$ is positive definite, see Remark 29. Besides, even if $P_{t,T}$ may be well defined for $T - t$ small enough, it would be then given by the same formula, and therefore the yield curve dynamics depends anyway on the factor X.
- In order to have a clear interpretation of the volatility factor X on the factor Y, a possible choice is to consider $d = q \times p$ with $q \in \mathbb{N}^*$ and $c_{i,j} = \mathbb{1}_{(i-1)\times p < j \leq i \times p}$. Thus, from (B.2), the principal matrix $(X_{k,l})_{(i-1)\times p < k,l \leq i\times p}$ rules the instantaneous quadratic variation of the factor Y_i while the submatrix $(X_{k,l})_{(i-1)\times p < k \leq i \times p, (j-1)\times p < l \leq j\times p}$ rules the instantaneous covariation between the factors Y^i and Y^j.

- The model does not prevent from having a negative short rate or from having $\mathbb{E}[|P_{t,T}|^k] = \infty$ for any $k > 0$, unless we consider the degenerated case ($p = 0$) where the yield curve is driven by the volatility factors X and the factors Y are null. This particular model has been studied by Gnoatto in [Gno13].

Remark 29 — *The condition (4.2) is sufficient to get that $P_{t,T}$ is well-defined. However, this condition does not depend on ϵ while we know that for $\epsilon = 0$, $P_{t,T}$ is well-defined since X is deterministic and Y is a Gaussian process. We can get a complementary sufficient condition when $-(b + b^\top)$ is positive definite, which is a reasonable assumption since it leads to a stationary process by Proposition 15. In this case, there exists $\mu > 0$ such that $-(b + b^\top) - \mu I_d \in \mathcal{S}_d^+(\mathbb{R})$. By using Proposition 14 with $\Upsilon = \frac{\mu}{4\epsilon^2} I_d$, we get that (3.7) is satisfied if we have*

$$\forall t \geq 0, \frac{\mu^2}{8\epsilon^2} I_d - \frac{\mu}{8\epsilon}(I_d^n \rho \lambda^\top c + I_d^n \rho \lambda^\top c) - \frac{1}{2} c^\top \lambda \lambda^\top c + \gamma \in \mathcal{S}_d^+(\mathbb{R}).$$

Since for $t \geq 0$, $\lambda(t)$ takes values in a compact subset of $\mathbb{R}^p$, there is $\epsilon_0 > 0$ such that this condition is satisfied for any $\epsilon \in (0, \epsilon_0)$.

Stationarity is considered a desirable property for a term structure model. The function mapping the factors (X, Y) into the yield curve (which is defined by the functions D, B and A) is stationary. Choosing between a stationary model and a non stationary model often results of a tradeoff between the flexibility needed for market prices calibration and the extent to which we want our model to exhibit a robust dynamical behavior. Too often excessive importance is accorded to the reproduction of the static stylized facts observed in the market (volatility smile, skew, output correlations...) neglecting the dynamical properties of the model. Non stationarity often results in poor dynamical properties of the model and can lead to serious mishedging issues (see paragraph 2.6 for an example). It is possible to allow for some non-stationarity of the model for the initial yield curve calibration while keeping a stationarity in the way the yield curve and volatility cube move. In practice this means we can consider a time dependent function $\varphi(t)$.

We give a clear interpretation of the underlying state variables of the model, separating them between yield curve factors Y and volatility factors X. Similarly to what we have done in the multi-factors gaussian model (see section 2.4.1), we want to identify the factors as drivers of the market prices, for example the first diagonal factor can be associated to the main driver of the short expiry volatility of long maturity swaptions. In a stationary model we can hedge against the factors movements. Assuming these factors have been identified with some market prices drivers the model would allow us to hedge the risks associated to the market movements. Furthermore, the model also contains additional degrees of freedom, such as a stochastic correlation between rates. This makes the model a good candidate to price and risk manage exotic products which depend on the correlation between rates such as CMS spread options.

In all the following we denote by $M\left\{\varphi,\gamma,(p,y,K,\theta,c),(d,\Omega,b,n,\epsilon),\rho\right\}$ the model defined above.

4.1.1 Model reduction

Relying on the identities in law of affine processes given in paragraph 3.1.5 we show that considering a more general dynamics for the state variables would not increase the number of degrees of freedom of the model.

Let $a \in \mathcal{M}_d(\mathbb{R})$, and consider the model $r_t = \varphi + \sum_{i=1}^p Y_t^i + \mathrm{Tr}\left(\tilde{\gamma}\tilde{X}_t\right)$ with

$$
\begin{aligned}
Y_t &= y + \int_0^t K(\theta - Y_s)ds + \int_0^t \tilde{c}\sqrt{\tilde{X}_s}\left[\sqrt{1-|\tilde{\rho}|^2}dZ_s + dW_s\tilde{\rho}\right] \\
\tilde{X}_t &= \tilde{x} + \int_0^t \left(\tilde{\Omega} + (d-1)\epsilon^2 a^\top a + \tilde{b}\tilde{X}_s + \tilde{X}_s\tilde{b}^\top\right)\right) ds + \epsilon\int_0^t \sqrt{\tilde{X}_s}dW_s a + a^\top dW_s^\top\sqrt{\tilde{X}_s},
\end{aligned}
$$

and $\tilde{\gamma} \in \mathcal{S}_d(\mathbb{R})$, $\tilde{x}, \tilde{\Omega} \in \mathcal{S}_d^+(\mathbb{R})$, $\tilde{c}, \tilde{b} \in \mathcal{M}_d(\mathbb{R})$, $\tilde{\rho} \in \mathbb{R}^d$ such that $|\tilde{\rho}| \leq 1$. This model may seem a priori more general, but this is not the case. In fact, let n be the rank of a and $u \in \mathcal{M}_d(\mathbb{R})$ be an invertible matrix such that $a^\top a = (u^{-1})^\top I_d^n (u^{-1})$. Then, $X_t = u^\top \tilde{X}_t u$ solves

$$
dX_t = [\Omega + (d-1)\epsilon^2 I_d^n + bX_t + X_t b^\top]dt + u^\top\sqrt{\tilde{X}_t}dW_t a u + u^\top a^\top dW_t^\top\sqrt{\tilde{X}_t}u,
$$

with $b = u^\top\tilde{b}(u^{-1})^\top$, $\Omega = u^\top\tilde{\Omega}u \in \mathcal{S}_d^+(\mathbb{R})$ and starting from $x = u^\top\tilde{x}u \in \mathcal{S}_d^+(\mathbb{R})$. After some calculations, we obtain $\langle d(Y_t)_m, d(Y_t)_{m'}\rangle = (\tilde{c}\tilde{X}_t\tilde{c}^\top)_{m,m'}dt = (cX_tc^\top)_{m,m'}dt$ with $c = \tilde{c}(u^{-1})^\top$; $\langle d(X_t)_{i,j}, d(X_t)_{k,l}\rangle = \epsilon^2\left[(X_t)_{i,k}\mathbb{1}_{j=l\leq n} + (X_t)_{i,l}\mathbb{1}_{j=k\leq n} + (X_t)_{j,k}\mathbb{1}_{i=l\leq n} + (X_t)_{j,l}\mathbb{1}_{i=k\leq n}\right]dt$ and

$$
\begin{aligned}
\langle d(Y_t)_m, d(X_t)_{i,j}\rangle &= \epsilon\left[(u^\top\tilde{X}_t\tilde{c}^\top)_{m,i}(u^\top a^\top\tilde{\rho})_j + (u^\top\tilde{X}_t\tilde{c}^\top)_{m,j}(u^\top a^\top\tilde{\rho})_i\right]dt \\
&= \epsilon\left[(X_tc^\top)_{m,i}(u^\top a^\top\tilde{\rho})_j + (X_tc^\top)_{m,j}(u^\top a^\top\tilde{\rho})_i\right]dt.
\end{aligned}
$$

Since the law of (X,Y) is characterized by its infinitesimal generator, we can assume without loss of generality that $\tilde{\rho} \in \ker(u^\top a^\top)^\perp = \mathrm{Im}(au)$. Therefore, there is $\rho' \in \mathbb{R}^d$ such that $\tilde{\rho} = au\rho'$, and we set $\rho_i = \rho'_i$ for $i \leq n$ and $\rho_i = 0$ for $n < i \leq d$. We have $|\rho|^2 = (\rho')^\top I_d^n\rho' = |\tilde{\rho}|^2 \leq 1$, and therefore (X,Y) follows the law $AFF((p,y,K,\theta,c),(d,x,\Omega,b,a),\rho)$, and we have $r_t = \varphi + \sum_{i=1}^p Y_t^i + \mathrm{Tr}(\gamma X_t)$ with $\gamma = u^{-1}\tilde{\gamma}(u^{-1})^\top$.
Therefore, the (reduced) specification of the model we have given maintains all the degrees of freedom of Wishart-like diffusions. The reduced form of the model simplifies most of the calculations, and in particular allows for a remarkable application of the splitting property of affine process highlighted in [AA13] to the model both for Monte Carlo simulation and implied volatility expansions.

4.1.2 Change of measure and Laplace transform

In the fixed income market, the pricing of vanilla products is often (if not always) made under a suitably chosen equivalent martingale measure different from the risk-neutral measure. It is thus important to characterize the distribution of the underlying market prices under these measure. The forward-neutral measures are probably the most important example of such pricing measures. The ATSM framework implies that the diffusion of the underlying state variables is of the form (3.1) (but with time-dependent coefficients) under all equivalent martingale measures. Furthermore, given the expression of zero-coupon bonds

$$P_{t,T} = \exp(A(T-t) + B(T-t)^T Y_t + \mathrm{Tr}(D(T-t)X_t))$$

we can provide semi-analytical expressions of the moments of zero-coupon bonds, by means of the Laplace transform of the couple (X, Y) under the corresponding measure.

Diffusion under the forward neutral probabilities

We assume that the condition (4.2) holds. Let Q^U denote the U-forward neutral probability, which is defined on $\mathcal{F}_U$ by

$$\frac{dQ^U}{d\mathbb{P}} = \frac{e^{-\int_0^U r_s d_s}}{P_{0,U}}.$$

This is the measure associated with the numeraire $P_{t,U}$. It comes from the martingale property of discounted asset prices that for $t \in (0, U)$,

$$\begin{aligned}
\frac{d\left(e^{-\int_0^t r_s d_s} P_{t,U}\right)}{e^{-\int_0^t r_s d_s} P_{t,U}} &= 2\epsilon \mathrm{Tr}(D(U-t)\sqrt{X_t} dW_t I_d^n) + B(U-t)^\top c\sqrt{X_t} dW_t \rho + \bar{\rho} B(U-t)^\top c \sqrt{X_t} dZ_t \\
&= \mathrm{Tr}([2\epsilon I_d^n D(U-t)\sqrt{X_t} + \rho B(U-t)^\top c\sqrt{X_t}] dW_t) + \bar{\rho} B(U-t)^\top c\sqrt{X_t} dZ_t.
\end{aligned}$$

From Girsanov's theorem, the processes

$$\begin{aligned}
dW_t^U &= dW_t - \sqrt{X_t}(2\epsilon D(U-t) I_d^n + c^\top B(U-t)\rho^\top) dt \\
dZ_t^U &= dZ_t - \bar{\rho}\sqrt{X_t} c^\top B(U-t) dt
\end{aligned}$$

are respectively matrix and vector valued Brownian motions under Q^U and are independent. This yields to the following dynamics for Y and X under Q^U:

$$\begin{aligned}
dX_t &= (\Omega + (d-1)\epsilon^2 I_d^n + b^U(t) X_t + X_t (b^U(t))^\top) dt + \epsilon\left(\sqrt{X_t} dW_t^U I_d^n + I_d^n (dW_t^U)^\top \sqrt{X_t}\right) \quad (4.4) \\
dY_t &= K(\theta - Y_t) dt + c X_t c^\top B(U-t) dt + 2\epsilon c X_t D(U-t) I_d^n \rho dt + c\sqrt{X_t}(dW_t^U \rho + \bar{\rho} dZ_t^U), \quad (4.5)
\end{aligned}$$

with $b^U(t) = b + 2\epsilon^2 I_d^n D(U-t) + \epsilon I_d^n \rho B(U-t)^\top c$.

Fourier/Laplace transform

One of the most appealing properties of affine processes is that their Fourier and Laplace transforms admit semi-analytical expression (see [DK96a]). The diffusion of the state variables is affine under the forward neutral measures, implying that the Fourier and Laplace transforms of the state variables under these measures admit a semi-analytic expression. Furthermore, the ZC bonds prices are exponential affine functions of the state variables, and thus the Fourier transforms of log ZC prices can be viewed as Fourier transforms of the state variables.

We are now interested in calculating the law of (X_T, Y_T) under the U-forward measure for $T \leq U$. More precisely, we calculate $\mathbb{E}^{Q^U}\Big[\exp(\mathrm{Tr}(\Gamma X_T) + \Lambda^\top Y_T)|\mathcal{F}_t\Big]$ for $t \in [0,T]$ by using again Proposition (14). We assume that condition (4.2) holds and have

$$\begin{aligned}
&\mathbb{E}^{Q^U}\Big[\exp(\mathrm{Tr}(\Gamma X_T) + \Lambda^\top Y_T)|\mathcal{F}_t\Big] \\
&= \frac{1}{P_{t,U}}\mathbb{E}\Big[\exp(\mathrm{Tr}(\Gamma X_T) + \Lambda^\top Y_T - \varphi(U-t) - \int_t^U \mathbf{1}_p^\top Y_s ds - \int_t^U \mathrm{Tr}(\gamma X_s) ds)|\mathcal{F}_t\Big] \\
&= \frac{e^{\mathrm{Tr}(\Gamma + D(U-T)X_T) + (\Lambda + B(U-T))^\top Y_T + A(U-T) - \varphi(T-t) - \int_t^T \mathbf{1}_p^\top Y_s ds - \int_t^T \mathrm{Tr}(\gamma X_s) ds}}{\exp(A(U-t) + \mathrm{Tr}(D(U-t)X_t) + B(U-t)^\top Y_t)}.
\end{aligned}$$

We consider $\Gamma \in \mathcal{S}_d(\mathbb{R})$ and $\Lambda \in \mathbb{R}^p$ such that

$$-\Gamma \in \mathcal{S}_d^+(\mathbb{R}) \text{ and } |\Lambda_i| \leq e^{-k_i(U-T)}/k_i, \text{ for } 1 \leq i \leq p,$$

in order to have $|\Lambda_i + B_i(U-T) \leq 1/k_i$ and $-(\Gamma + D(U-T)) \in \mathcal{S}_d^+(\mathbb{R})$. By Proposition (14), we get that the expectation is finite and that

$$\mathbb{E}^{Q^U}\Big[\exp(\mathrm{Tr}(\Gamma X_T) + \Lambda^\top Y_T)|\mathcal{F}_t\Big] = \exp(A^U(t,T) + \mathrm{Tr}(D^U(t,T)X_t) + B^U(t,T)^\top Y_t), \quad (4.6)$$

with $F^U(t,T) = \tilde{F}(T-t) + F(U-T) - F(U-t)$ for $F \in \{A, D, B\}$, where $(\tilde{B}, \tilde{D}, \tilde{A})$ is the solution of (3.9) with $\tilde{B}(0) = \Lambda + B(U-T)$, $\tilde{D(0)} = \Gamma + D(U-T)$, $\tilde{A}(0) = 0$, $\bar{\Lambda} = \mathbf{1}_p$ and $\bar{\Gamma} = -\gamma$.

Corollary 30 *Let* (4.2) *hold. For* $\Gamma \in \mathcal{S}_d(\mathbb{R})$ *and* $\Lambda \in \mathbb{R}^p$ *such that* $-\Gamma \in \mathcal{S}_d^+(\mathbb{R})$ *and* $|\Lambda_i| \leq e^{-k_i(U-T)}/k_i$ *for* $1 \leq i \leq p$, $\mathbb{E}^{Q^U}\Big[\exp(\mathrm{Tr}(\Gamma X_T) + \Lambda^\top Y_T)|\mathcal{F}_t\Big] < \infty$ *a.s. for any* $t \in [0,T]$ *and is given by* (4.6).

Let us mention that in practice, the formula above for $A^U(t,T)$, $D^U(t,T)$ and $B^U(t,T)$ requires to solve two different ODEs. It may be more convenient to use the following one that can be easily deduced from dynamics of (X,Y) under the U-forward measure.

$$\begin{cases} \frac{\partial B^U}{\partial t}(t,T) &= K^\top B^U(t,T),\ B^U(T,T)=\Lambda, \\ -\frac{\partial D^U}{\partial t}(t,T) &= 2\epsilon^2 D^U I_d^n D^U + D^U(b^U(t)+\frac{\epsilon}{2}I_d^n\rho(B^U)^\top c)+(b^U(t)+\frac{\epsilon}{2}I_d^n\rho(B^U)^\top c)^\top D^U \\ & +\frac{1}{2}c^\top B^U(B^U)^\top c + c^\top B^U B(U-t)^\top c + \epsilon D(U-t)I_d^n\rho(B^U)^\top c + \epsilon c^\top B^U\rho^\top I_d^n D(U-t), \\ & D^U(T,T)=\Gamma, \\ -\frac{\partial A^U}{\partial t}(t,T) &= B^U(t,T)^\top K\theta + \mathrm{Tr}\left(D^U(t,T)(\Omega+\epsilon^2(d-1)I_d^n)\right),\ A^U(T,T)=0. \end{cases} \tag{4.7}$$

4.2 Numerical framework

In this section we investigate fast numerical methods for vanilla options pricing, intended to be used for model calibration. As already mentioned, one of the main appeals of SC-VATSM is that Laplace and Fourier transform of the underlying market assets log prices can be computed by means of analytical or semi-analytical formulas. Transform based methods and moments methods are thus very efficient in term of computational time.

4.2.1 Series expansion of distribution

In this paragraph we present pricing methods based on the expansion of the distribution of the underlying asset prices. These methods have previously been applied for pricing swaptions in [CDG02] and [TYW05], both papers do not provide theoretical arguments for the convergence of the expansion series, nor for the bounds of the error. Numerical tests are provided for classic ATSM, in particular for the multi-factors gaussian and CIR model. The interest rates vanilla options market is made of caps, floors and swaptions. As discussed in paragraph 1.3.2, it is standard to write swaptions prices as call or put options on coupon bonds. Suppose we want to compute

$$C(K) = \mathbb{E}^Q\left[e^{-\int_t^T ds r_s}(Y_T-K)^+|\mathcal{F}_t\right],$$

where Q is the risk-neutral probability and $Y_t = \sum_{i=1}^n C_i P_{T,T_i}$. We have the following decomposition of the price

$$C(K) \;=\; \sum_{i=1}^n C_i\mathbb{E}\left[e^{-\int_t^T ds r_s}P_{T,T_i}1_{Y_T>K}|\mathcal{F}_t\right] - K\mathbb{E}^Q\left[e^{-\int_t^T ds r_s}1_{Y_T>K}|\mathcal{F}_t\right].$$

By standard change of numeraire arguments [GREK95], we can write the following change of measure $P_{t,T_i}\frac{dQ^{T_i}}{dQ}\mid_{\mathcal{F}_t} = e^{-\int_t^T ds r_s}P_{T,T_i}$. Which leads to the following expression

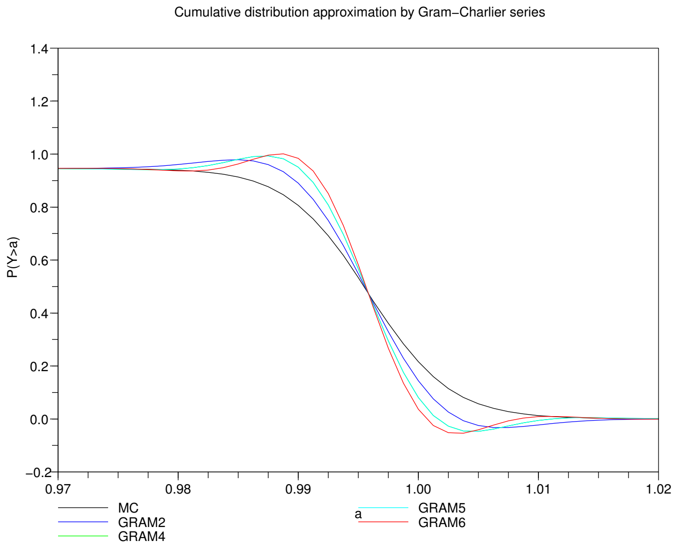

Figure 4.1: Quality of the approximation of $Q^T(P_{T,T+\delta} > a)$ by the Gram-Charlier series at different orders, with $T = 4$ and $\delta = 3m$. The model parameters are $x = 10^{-4}I_d$, $y = (-0.02, 0.01)$, $\alpha = 2.5$, $b = \text{diag}(-0.11, -1.1)$, $K = \text{diag}(1, 0.1)$, $c = I_d$. The higher order moments of the ZCB are computed using the numerical resolution of the MRDE described in 3.1.4, with a time discretization step of 0.01.

$$C(K) = \sum_{i=1}^{n} C_i P_{t,T_i} Q^{T_i} \left(Y_T > K|\mathcal{F}_t\right) - KQ^T \left(Y_T > K|\mathcal{F}_t\right).$$

Moments expansion pricing methods are based on the idea of building an approximation of the probabilities $Q^{T_i}\left(Y_T > K|\mathcal{F}_t\right)$ through Gram-Charlier or Edgeworth series. The implementation requires efficient methods for computing the moments of Y_T, which are easily given by linear combinations of the moments of the zero-coupon bonds P_{T,T_i}.

In the following we present two different ways to approximate the density function of Y by series, the cumulative distribution function can then easily be approximated as the integral of such series. Without loss of generality we assume that Y is centered i.e. $\mathbb{E}[Y] = 0$ and denote by σ^2 its variance.

Gram-Charlier expansions

Gram-Charlier series are widely used in various domains of applied mathematics and physics for cumulative distribution functions estimation. For a function in $L^2(\mathbb{R})$, this series can be viewed as the canonical representation in the Hilbert basis

$$e^{-x^2/2} H_j(x), \quad j \geq 1,$$

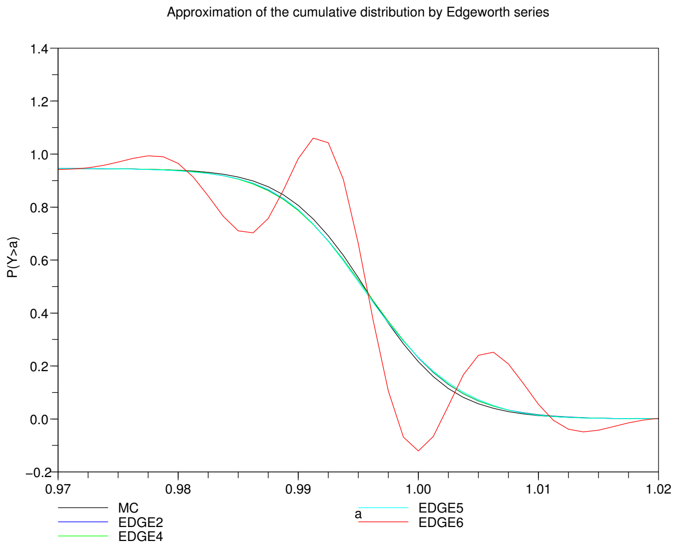

Figure 4.2: Quality of the approximation of $Q^T(P_{T,T+\delta} > a)$ by the Edgeworth series at different orders, with $T = 4$ and $\delta = 3m$. The model parameters are $x = 10^{-4} I_d$, $y = (-0.02, 0.01)$, $\alpha = 2.5$, $b = \text{diag}(-0.11, -1.1)$, $K = \text{diag}(1, 0.1)$, $c = I_d$. The higher order moments of the ZCB are computed using the numerical resolution of the MRDE described in 3.1.4, with a time discretization step of 0.01.

where H_j denotes the Hermite polynomial of degree j, which is defined by

$$\frac{d^j}{(dx)^j}\left(\frac{e^{-x^2/2}}{\sqrt{2\pi}}\right) = (-1)^j H_j(x)\frac{e^{-x^2/2}}{\sqrt{2\pi}}.$$

The polynomials H_j admit the following explicit expression

$$H_j(x) = n! \sum_{k=0}^{[n/2]} \frac{(-1)^k x^{n-2k}}{k!(n-2k)!2^k}.$$

The following theorem gives sufficient conditions under which the convergence of the Gram Charlier series hold.

Theorem 31 *—(Cramer [Cra57]) Let f be a density function such that*

$$\int_{\mathbb{R}} dx e^{\frac{x^2}{4}} f(x) < \infty, \tag{4.8}$$

and f is of bounded variations, then the Gram-Charlier series converges in every continuity point of f

$$f(x) = \sum_{k\geq 0} \frac{c_k}{k!}\phi^{(k)}(x),$$

where $\phi^{(k)}$ is the k-th derivative of the normal density and c_k is defined by

$$c_k = \int_{\mathbb{R}} dx f(x) H_k(x).$$

For practical purposes it is important to measure the quality of the approximation given by the truncated series. Unfortunately there exist very few results on the error of approximation of the truncated series. The following theorems give (under some very restrictive conditions) the rate of convergence of Gram-Charlier series.

Theorem 32 — *Let f be a density function such that, f has k-th continuous derivative of bounded variation, and*

$$x e^{\frac{x^2}{4}} f^{(k)}(x), x^2 e^{\frac{x^2}{4}} f^{(k-1)}(x), ..., x^{k+1} e^{\frac{x^2}{4}} f(x)$$

are of bounded variations on $\mathbb{R}$, then

$$\left| f(x) - \sum_{n=0}^{N} \frac{c_n}{n!} \phi^{(n)}(x) \right| < M N^{-k/2} (1+x^2)^{1/6} e^{-x^2/4},$$

where M is a constant independent of N.

Note that the sufficient convergence conditions (4.8) is not verified for zero-coupon bonds in our model. Let $E \in \mathbb{R}$, $B \in \mathbb{R}^p$, and $D \in \mathcal{S}_d(\mathbb{R})$, we have

$$\begin{aligned} \mathbb{E}^{Q^U}\left[\exp\left(\frac{e^{E+\mathrm{Tr}(DX_T)+B^T Y_T}}{4}\right)\right] &= \mathbb{E}^{Q^U}\left[\mathbb{E}^{Q^U}\left[\exp\left(\frac{e^{E+\mathrm{Tr}(DX_T)+B^T Y}}{4}\right)|X_s, s \le t\right]\right] \\ &= \mathbb{E}^{Q^U}\left[\mathbb{E}\left[\exp\left(\frac{1}{4} e^{a(x_s, s\le t) + b(x_s, s \le t)^T G}\right)\right] |_{x_s = X_s, s \le t}\right] \end{aligned}$$

where G is a gaussian variable. The last equality comes form the fact that conditionally to the path of the process X, the process Y is gaussian under the U-forward measure. Precisely, conditionally to the path of the process X, Y is an Ornstein-Uhlenbeck process. Since $\mathbb{E}\left[\exp(\eta \exp(\mu G)\right] = \infty$ for any $\eta > 0$ and $\mu \in \mathbb{R}^*$, it is clear that we must have

$$\mathbb{E}^{Q^U}\left[\exp\left(\frac{e^{E+\mathrm{Tr}(DX_T)+B^T Y_T}}{4}\right)\right] = \infty,$$

which implies that the sufficient conditions for the convergence of the Gram-Charlier series are not verified. In fact the series diverge in many situations of practical interest (see Figure 4.1).

Edgeworth expansion

In the limit Edgeworth and Gram-Charlier series are the same, but Edgeworth series overcome some of the problems of Gram-Charlier by assembling the terms of the same order together. Let us denote by f the density of a generic random variable Y and by $G(k)$ its characteristic function

$$G(k) = \mathbb{E}\left[e^{ikY}\right] = \int_{\mathbb{R}} dy e^{iky} f(y).$$

Provided that the function G admits a Taylor expansion when k goes to zero, the cumulants κ_i are defined by

$$\log G(k) = \sum_{j\geq 1} \kappa_j \frac{(ik)^j}{j!}.$$

Cumulants can be computed from the moments μ_n of Y, we have

$$\kappa_n = n! \sum_{\{k_m\}} (-1)^{r-1}(r-1)! \prod_{m=1}^{n} \frac{1}{k_m!} \left(\frac{\mu_m}{m!}\right)^{k_m},$$

here the summation extends over all non-negative integers $\{k_m\}$ satisfying

$$k_1 + 2k_2 + \cdots + nk_n = n, \tag{4.9}$$

and $r = k_1 + k_2 + \cdots + k_n$. Let us denote by g the characteristic function of Y/σ,where $\sigma = \sqrt{Var(Y)}$, then $g(k) = G(k/\sigma)$, and thus we have

$$g(k) = e^{-\frac{k^2}{2}} \exp\left(\sum_{n\geq 3} \frac{S_n \sigma^{n-2}}{n!} (ik)^n\right),$$

where $S_n = \kappa_n/\sigma^{2n-2}$. Developing the exponential function as a formal series in powers of σ we get

$$\exp\left(\sum_{n\geq 3} \frac{S_n \sigma^{n-2}}{n!} (ik)^n\right) = 1 + \sum_{s\geq 1} P_s(ik)\sigma^s \tag{4.10}$$

where

$$P_s(ik) = \sum_{\{k_m\}} \prod_{m=1}^{s} \frac{1}{k_m!} \left(\frac{S_{m+2}(ik)^{m+2}}{(m+2)!}\right)^{k_m},$$

where again the summation extends over all non-negative integers $\{k_m\}$ satisfying (4.9). If the characteristic function G is integrable, injecting the expression of g in the Fourier inversion formula, we get the following expression for the density p of Y/σ

$$
\begin{aligned}
p(y) &= \frac{1}{2\pi}\int_{\mathbb{R}} dk e^{-iky} g(k) \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} dk e^{-iky} e^{-\frac{k^2}{2}} \left(1 + \sum_{s\geq 1}\sum_{\{k_m\}}\prod_{m=1}^{s}\frac{1}{k_m!}\left(\frac{S_{m+2}(ik)^{m+2}}{(m+2)!}\right)^{k_m}\right) \\
&= \phi(y) + \sum_{s\geq 1}\sigma^s \left\{\sum_{\{k_m\}}\prod_{m=1}^{s}\frac{1}{k_m!}\left(\frac{S_{m+2}(-1)^{m+2}}{(m+2)!}\frac{d^{m+2}}{dx^{m+2}}\right)^{k_m}\phi(x)\right\} \\
&= \phi(y)\left\{1 + \sum_{s\geq 1}\sigma^s \sum_{\{k_m\}} H_{s+2r}(y)\prod_{m=1}^{s}\frac{1}{k_m!}\left(\frac{S_{m+2}(-1)^{m+2}}{(m+2)!}\right)^{k_m}\phi(x)\right\}
\end{aligned}
$$

Likewise Gram-Charlier series, Edgeworth series also diverge in several cases of practical interest (see Figure 4.2).

Remark 33 — *The truncated series are polynomial functions which are not necessarily proper densities. In particular the truncated series are not necessarily positive or unimodal. Boundary regions for the moments of the distribution exist to impose that the Gram-Charlier or Edgeworth series define proper densities, we refer to [BZ88] for further details. To the best of our knowledge no tractable explicit boundary conditions exist for a generic order series.*

4.2.2 Transform pricing

We follow [CM99] and [Lee04] and recall some results for transform pricing methods. Consider a general framework with an asset S which is a martingale under a generic martingale measure Q. Suppose we want to price a Call on S of maturity T

$$C(SK) = \mathbb{E}^Q\left[(S_T - SK)^+ | \mathcal{F}_t\right].$$

In the following we will drop the Q in the expectation, if not specified, the expectations are always meant to be taken under the martingale measure Q. We introduce the notations

$$
\begin{aligned}
s_t &= \log(S_t) \\
k &= \log(SK) \\
\phi_T(u) &= \mathbb{E}^Q\left[\exp(ius_T)\right]
\end{aligned}
$$

With a slight abuse of notations, we will note either $C(SK)$ or $C(k)$ the call price as a function of respectively the strike or the log strike. For a given random variable X, we define A_X as the interior of $\{v \in \mathbb{R} : \mathbb{E}\left[e^{vX}\right] < \infty\}$.

Existence theorem and Fourier inversion

The following theorem gives an existence result for the modified Fourier transform of s_T.

Theorem 34 *—(Roger Lee [Lee04]) Assume* $1 \in A_{s_T}$*. Then there exists* $\alpha > 0$ *such that* $1+\alpha \in A_{s_T}$ *and for any such* α*,* $\hat{c}_\alpha$ *exists and*

$$\hat{c}_\alpha(u) = \frac{\phi_T(u-(1+\alpha)i)}{\alpha^2+\alpha-u^2+i(2\alpha+1)u},$$

where $\hat{c}_\alpha$ *denotes the Fourier transform of* $c_\alpha(k) = e^{\alpha k}C(k)$*.*

Let us note that $1 \in A_{s_T}$ means that $\mathbb{E}[S_T] < \infty$ and that there exists $\epsilon > 0$ such that $\mathbb{E}[S_T^{1+r}] < \infty$ for $r < \epsilon$. The following theorem gives the expression of the Call price obtained by Fourier transform

Theorem 35 *—(Carr-Madan [CM99]) Assume* $1 \in A_{s_T}$*. Then there exists* $\alpha > 0$ *such that*

$$\begin{aligned} C(SK) &= \frac{e^{-\alpha k}}{2\pi}\int_{\mathbb{R}} due^{-iuk}\hat{c}_\alpha(u) \\ &= \frac{e^{-\alpha k}}{\pi}\int_{\mathbb{R}_+} due^{-iuk}\mathrm{Re}\left(\hat{c}_\alpha(u)\right) \end{aligned}$$

Transform pricing methods are based on a discrete approximation of the above integral.

Approximation error bounds

In [Lee04], the author provides bounds for both sampling and truncation errors, which gives bounds for the approximation error of the call price. Let us introduce the following notations:

$$\begin{aligned} \Sigma^N(k) &= e^{-\alpha k}\frac{\Delta}{\pi}\mathrm{Re}\left[\sum_{n=0}^{N-1} e^{-i(n+1/2)k\Delta}\hat{c}_\alpha((n+1/2)\Delta)\right] \\ \Sigma^\infty(k) &= \lim_{N\to\infty}\Sigma^N(k). \end{aligned}$$

We have provided here the price approximation given by the Riemann sum of the integral, more generally, we can use the following approximation of the price

$$C(SK) = \frac{\exp(-\alpha k)}{\pi}\sum_{n=0}^{N-1}\exp(-iu(n)k)\hat{c}_\alpha(u(n))\omega_n(\Delta),$$

where $u(n), \omega(\Delta)$ are respectively the sampling points and the weights depending on the choice of numerical integration method, and Δ is the step of the integration method. For example, the Simpson's rule points and weights are given by

$$\begin{aligned} u(n) &= n\Delta \\ \omega_n(\Delta) &= \frac{\Delta}{3}(3 + (-1)^{n-1} - \delta_n). \end{aligned}$$

The following theorems provide bound estimates for the sampling error and the truncation error. Error bounds estimates depend respectively on the finite moments of the underlying, and on the asymptotic (decreasing) properties of the modified Fourier transform.

Theorem 36 *—(Roger Lee [Lee04]) Assume* $\mathbb{E}[\exp(s_T)] < \infty$.

If ϕ_T *is such that* $\hat{c}_\alpha$ *decays as a power* $|\hat{c}_\alpha(u)| \leq \frac{\psi(u)}{u^{1+\gamma}}$ *for* $u > u_0$ *large enough, where* $\gamma > 0$ *and* ψ *is a decreasing function, then the truncation error*

$$|\Sigma^N(k) - \Sigma^\infty(k)| \leq \frac{\psi(N\Delta)}{\pi e^{\alpha k}\gamma(N\Delta)^\gamma}, \quad for \quad N\Delta > u_0.$$

If ϕ_T *is such that* $\hat{c}_\alpha$ *decays exponentially* $|\hat{c}_\alpha(u)| \leq \psi(u)e^{-\gamma u}$ *for* $u > u_0$ *large enough, where* $\gamma > 0$ *and* ψ *is a decreasing function, then the truncation error*

$$|\Sigma^N(k) - \Sigma^\infty(k)| \leq \frac{\Delta\psi((N + 1/2)\Delta)}{2\pi e^{\alpha k + \gamma N\Delta}\sinh(\gamma\Delta/2)}, \quad for \quad N\Delta > u_0.$$

In order to exploit the above theorem we should be able to characterize the asymptotic properties of the modified fourier transform of the log-price. Again, this means we should be able to characterize the asymptotic properties of the solutions of MRDE, which is not always easy. Yet, we can always find a rough upper bound for $|\hat{c}_\alpha(u)|$, we have

$$|\hat{c}_\alpha(u)| \leq \frac{\mathbb{E}\left[S_T^{1+\alpha}\right]}{|\alpha^2 + \alpha - u^2 + i(2\alpha + 1)u|} = \frac{\mathbb{E}\left[S_T^{1+\alpha}\right]}{u^4 + u^2(2\alpha(\alpha + 1) + 1) + \alpha^2(\alpha + 1)^2}.$$

Then there exists a constant C and a u_0 such that for $u > u_0$

$$|\hat{c}_\alpha(u)| \leq \frac{C\mathbb{E}\left[S_T^{1+\alpha}\right]}{u^2}.$$

While rough this bound is useful if we don't have explicit expressions of $\hat{c}_\alpha$, and cannot work out the limit behavior of its semi-analytical expression. As expected the truncation error is bounded by a quantity depending on $U = N\Delta$ which is the upper bound of Fourier inversion integral to which we chose to truncate. Applying the above theorem with our rough upper bound, we know that the truncation error is $O(1/U)$ when U is large enough. Numerical tests proved that the norm of the modified Fourier transform $|\hat{c}_\alpha(u)|$ decreases faster then $O(1/u^{1+\gamma})$ in several cases of practical interest. Figure 4.5 shows the convergence $\Sigma^N(k)$ for different values of the model parameters to the price of 1Yx1Y caplet price. Clearly the norm of the modified Fourier transform $|\hat{c}_\alpha(u)|$ decreases as an en exponential function of u, which ensures a convergence faster then $O(1/U)$.

The following theorem provides an upper bound for the sampling error.

Theorem 37 *—(Roger Lee [Lee04]) Assume* $1 \in A_{s_T}$ *and* $(1+\alpha) \in A_{s_T}$.

$$|C(SK) - \Sigma^\infty(k)| \leq \inf_{p>\alpha:p+1\in A_{s_T}} \left[\frac{e^{-2\pi\alpha/\Delta}\phi_T(-i)}{1-e^{-4\pi\frac{\alpha}{\Delta}}} + \frac{e^{2\pi(\alpha-p)/\Delta}\phi_T(-i(p+1))}{(1-e^{-4\pi\frac{\alpha}{\Delta}})(p+1)e^{pk}} \left(\frac{p}{p+1}\right)^p \right].$$

The error depends on the moments of the underlying asset price. Again we do not have an explicit expression for the moments of the underlying asset prices in our model, and it is difficult to exploit the above bound.

FFT algorithm setup

FFT is an efficient algorithm to compute the sum

$$y_m = \sum_{n=0}^{N-1} e^{-i\frac{2\pi}{N}nm} x_n \quad \text{for} \quad m = 0, ..., N-1.$$

In order to be able to apply the above algorithm to compute the sum $\Sigma^N(k)$, we rearrange the expression. We chose a strike discretization $k_m = \underline{k} + \lambda m$, then we have the following approximation

$$\Sigma^N(k_m) = \frac{\exp(-\alpha k_m)\Delta}{\pi} \sum_{n=0}^{N-1} \exp(-inm\lambda\Delta) e^{-in\underline{k}\Delta} \hat{c}_\alpha((n+1/2)\Delta).$$

As soon as we impose that

$$\Delta\lambda = \frac{2\pi}{N}, \tag{4.11}$$

the above expression fits into the FFT framework with the vectors

$$\begin{aligned} x_n &= e^{-in\underline{k}\Delta}\hat{c}_\alpha((n+1/2)\Delta) \\ y_m &= \Sigma^N(k_m). \end{aligned}$$

Application to Caplets pricing in the SCVATSM

Caplets can be written as call or puts respectively on forward ZC bond and on ZC bond. We have

$$\text{Caplet}(t, T, \delta, K) = P_{t,T} \frac{1+\delta K}{\delta} \mathbb{E}^{Q^T} \left[\left(\frac{1}{1+\delta K} - P_{T,T+\delta} \right)^+ \right],$$

By Call-Put parity, the pricing problem reduces to the pricing of the following quantities

$$V_t^T(SK) \;=\; \mathbb{E}^T\left[(P_{T,T+\delta} - SK)^+\right].$$

We can thus apply the FFT framework to the above quantity. We introduce the notations

$$\begin{aligned} S_t &= P_{t,T+\delta} \\ s_t &= \log(S_t) \\ k &= \log(SK). \end{aligned}$$

and ψ_T is the modified Fourier transform of s_T

$$\begin{aligned} \hat{v}_\alpha^T(v) &= \frac{\phi_T(v-(\alpha+1)i)}{\alpha^2+\alpha-v^2+i(2\alpha+1)v} \\ \phi_T(u) &= \mathbb{E}^{Q^T}\left[\exp(ius_T)|\mathcal{F}_t\right] \\ &= \mathbb{E}^{Q^T}\left[\exp(iu(\mathbb{E}(\delta)+\mathrm{Tr}(D(\delta)X_T)+B(\delta)^T Y_T))|\mathcal{F}_t\right], \end{aligned}$$

which can be easily computed using (3.10).

Let us note that the above method Fourier transform based methods are not directly applicable for the pricing of swaptions. The reason for this is that neither the Fourier transform of the forward swap rate, nor the Fourier transform of coupon bonds are admit semi-analytical expressions. In order to apply Fourier methods for pricing swaptions, one needs to use the drift freezing approximations introduced in 1.3.2. We refer to Schrager and Pelsser [SP06] and Singleton and Umantsev [SU02] for a detailed description of these methods which directly apply to our model.

4.2.3 Numerical results

We test the accuracy of the methods for pricing caplets and swaptions. Our benchmark is given by the Monte Carlo price. We will test the accuracy of the methods for different values of the model parameters.

Let us now present the models that will be tested. In the following we will denote by `model 1` the model defined by the following set of parameters:

$$K = \mathrm{diag}(0.1, 1.), c = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, b = \mathrm{diag}(-0.41, -0.011), \Omega = -bx_\infty - x_\infty b^T$$
$$x = \begin{pmatrix} 1.2\% & -4.32\cdot 10^{-5} \\ -4.32\cdot 10^{-5} & 1.8\% \end{pmatrix}, x_\infty = \begin{pmatrix} 0.6\% & -3\cdot 10^{-6} \\ -3\cdot 10^{-6} & 1\% \end{pmatrix}, \epsilon = 2\cdot 10^{-3}.$$

The parameters values have been chosen in such a way that the yield curve and volatility levels generated by the model are in line with todays US and EUR interest rates market levels. In particular, $\theta^T = (0.02, -0.01), y^T = (0.02, -0.03)$, the parameter φ is set in such a way that the spot rate is equal to 0.5%, and the parameter γ is set as follows:

$$\gamma = \frac{1}{2} c^\top (K^\top)^{-1} \mathbf{1}_p \mathbf{1}_p^\top (K)^{-1} c + 0.0001 I_d. \tag{4.12}$$

Let us recall that under the condition $\gamma - \frac{1}{2} c^\top B(\tau) B(\tau)^\top c \in \mathcal{S}_d^+(\mathbb{R})$ for all $\tau \geq 0$ we have that D is well defined on $\mathbb{R}_+$ and $-D \in \mathcal{S}_d^+(\mathbb{R})$. Since $B(0) = 0$, the condition is verified at $\tau = 0$ as soon as $\gamma \in \mathcal{S}_d^+(\mathbb{R})$. By defining γ as above, we have imposed that the positiveness condition is verified when τ goes to infinity.

Series expansion

Due to the divergence of the Gram-Charlier series, increasing the number of terms in the expansion does not necessarily leads to more accurate approximations of the cumulative distribution function of the coupon bond Y. See Figures 4.2 and 4.1.

Figures 4.3 and 4.4 show that the Edgeworth expansion approximation of the caplet price is accurate across different expiries and maturities. The approximated price is inside the confidence interval obtained with 10^5 simulation paths and 30 discretization steps in a reasonably wide range around the ATM price.

Transform methods

We have investigated the pricing of caplets using FFT. As mentioned before we are not able to describe the asymptotic behavior of the modified fourier transform of the log-price $\hat{c}_\alpha$ defined in theorem . Numerical evidence indicate that $|\hat{c}_\alpha|$ decreases exponentially with u. See Figure 4.5.

4.3 Volatility expansions for caplets and swaptions

The goal of this section is to provide the asymptotic behavior of the Caplet and Swaption prices when the volatility parameter ϵ is close to zero. The practical interest of these formulas is to give a proxy for these prices. Thus, they give a tool to calibrate the model parameters to the smile. Getting an expansion on the smile is complementary with respect to the transform based and density expansion methods we have presented in the previous section. On the one hand, it is less accurate to calculate a single price since we only calculate here the expansion up to ϵ^2. On the other hand, it is more tractable for a first calibration of the model and gives a good approximation for key quantities on the smile.

The arguments that we use in this section to obtain the expansion have been developed in the book of Fouque et al. [FPS00]. They rely on an expansion of the infinitesimal generator with respect to ϵ. Recently, this technique was applied by Bergomi and Guyon [BG12] to provide approximation under a multi factor model for the forward variance. Here, we have to take into account some specific features of the fixed income and work under the appropriate probability measure to apply these arguments. Not surprisingly the zero order term in the expansion is exactly the volatility of the LGM with a time-dependent variance-covariance matrix. More interestingly the higher order terms allow to confirm the intuitions on the role of the parameters and factors that determine the shape and dynamics of the volatility. An alternative expansion for an equity Wishart model similar to the one we model has been derived in [BBEK08], and studied in [DFG11].

Last, we have to mention that the calculations presented in this section are rather formal. In particular, we implicitly assume that the caplet and swaption prices are smooth enough and admits expansions with respect to ϵ. A rigorous proof of these expansions is beyond the scope of this paper.

4.3.1 Price and volatility expansion for Caplets

It is standard market practice to view Caplets $\text{Caplet}(t, T, \delta)$ as Call options on the forward libor rate $L_t(T, \delta)$, and analyze the prices in terms of implied Normal volatility of the forward libor rate under the $T + \delta$-forward neutral measure $Q^{T+\delta}$. It is fairly equivalent to look at caplets as Call options on the forward ZC bond $\frac{P_{t,T}}{P_{t,T+\delta}}$ and analyze the prices in terms of log-normal implied volatility of the forward ZC bond under the $Q^{T+\delta}$. In fact these two volatilities are close one to the other. The only quantity of interest in order to understand the Caplets volatility cube is what we call the forward Caplet price

$$\text{FCaplet}(t, T, \delta) = \mathbb{E}^{T+\delta}\left[(L_T(T, \delta) - K)^+ \,|\mathcal{F}_t\right],$$

which can be rewritten as a call option on the forward zero coupon bond $\frac{P_{t,T}}{P_{t,T+\delta}}$

$$\text{FCaplet}(t, T, \delta) = \frac{1}{\delta}\mathbb{E}^{T+\delta}\left[\left(\frac{P_{T,T}}{P_{T,T+\delta}} - (1 + \delta K)\right)^+ \middle| \mathcal{F}_t\right].$$

Since (X, Y) is a Markov process, $\text{FCaplet}(t, T, \delta)$ is a function of (X_t, Y_t) and therefore we can define the forward price function

$$P(t, x, y) = \mathbb{E}^{T+\delta}\left[\left(\frac{P_{T,T}}{P_{T,T+\delta}} - (1 + \delta K)\right)^+ \middle| X_t = x, Y_t = y\right]. \tag{4.13}$$

The goal of Subsection 4.3.1 is to obtain the second order expansion (4.15) of P with respect to ϵ.

A convenient change of variable

We want to get an expansion of the caplet price with respect to ϵ. To do so, we first need to get an expansion to ϵ of the infinitesimal generator of the process (X, Y) under the probability $Q^{T+\delta}$. However, we can make before a change of variable that simplifies this approach. Thus, we define

$$H_t = \Delta A(t,T,\delta) + \mathrm{Tr}(\Delta D(t,T,\delta)X_t) + \Delta B(t,T,\delta)^\top Y_t,$$

with

$$\begin{aligned} \Delta A(t,T,\delta) &= A(t,T) - A(t,T+\delta) \\ (\Delta B, \Delta D)(t,T,\delta) &= (B,D)(T-t) - (B,D)(T+\delta-t) \end{aligned}$$

Thus, we have $\frac{P_{t,T}}{P_{t,T+\delta}} = e^{H_t}$. It is well known that $\frac{P_{t,T}}{P_{t,T+\delta}}$ is a martingale under $Q^{T+\delta}$, see e.g. Proposition 2.5.1 in Brigo and Mercurio [BM06]. Thus, we get by Itô calculus from (4.4) and (4.5) that (X, H) solve the following SDE

$$\begin{aligned} dX_t &= (\Omega + \epsilon^2(d-1)I_d^n + b^{T+\delta}(t)X_t + X_t(b^{T+\delta}(t))^\top)dt + \epsilon\sqrt{X_t}dW_t^{T+\delta}I_d^n + \epsilon I_d^n(dW_t^{T+\delta})^\top\sqrt{X_t}, \\ dH_t &= -\frac{1}{2}\left(\Delta B^\top cX_tc^\top\Delta B + 4\epsilon^2\mathrm{Tr}(\Delta D I_d^n \Delta D X_t) + 2\epsilon(\Delta B^\top cX_t\Delta D I_d^n\rho)\right)dt \\ &\quad + \Delta B^\top c\sqrt{X}_t(dW_t^{T+\delta}\rho + \bar{\rho}dZ_t^{T+\delta}) + 2\epsilon\mathrm{Tr}(\Delta D\sqrt{X}_t dW_t^{T+\delta}I_d^n). \end{aligned} \quad (4.14)$$

Therefore, $P(t,x,y) = \mathbb{E}^{T+\delta}\left[\left(e^{H_T} - (1+\delta K)\right)^+ |X_t = x, Y_t = y\right]$ only depends on (x, y) through (x, h) where $h = \Delta B(t,T,\delta)^\top y + \mathrm{Tr}\left(\Delta D(t,T,\delta)x\right) + \Delta A(t,T,\delta)$, and we still denote by a slight abuse of notations

$$P(t,x,h) = \mathbb{E}^{T+\delta}\left[\left(e^{H_T} - (1+\delta K)\right)^+ |X_t = x, H_t = h\right].$$

Let us emphasize that this change of variable is crucial in order to apply an expansion procedure similar to the one of Bergomi and Guyon [BG12]. It allows to reduce the dimensionality of the underlying state variable. The variable H is one-dimensional and it is the only variable that appears in the payoff of the caplet. Though this is obvious from the definition of the model, we insist on the fact the implied volatility of caplets is a function of the factors X only. This appears clearly in the SDE (4.14), H_t can be viewed as continuous version of the forward Libor rate and its volatility depends on the factors X only.

Expansion of the price

From the SDE (4.14) and 12, we get the following PDE representation of P:

$$\begin{aligned} \partial_t P + L(t)P &= 0 \\ P(T,x,h) &= (e^h - (1+\delta K))^+ \end{aligned}$$

where $L(t)$ is given by

$$
\begin{aligned}
L(t) =& \frac{1}{2}\left(\Delta B^\top c x c^\top \Delta B + 4\epsilon^2 \mathrm{Tr}(\Delta D I_d^n \Delta D x) + 2\epsilon(\Delta B^\top c x \Delta D I_d^n \rho)\right)(\partial_h^2 - \partial_h) \\
&+ \sum_{1\le i,j\le d} (\Omega + (d-1)\epsilon^2 I_d^n + b^{T+\delta} x + x(b^{T+\delta})^\top)_{i,j} \partial_{x_{i,j}} \\
&+ \frac{1}{2} \sum_{1\le i,j,k,l\le d} \epsilon^2 [x_{i,k}(I_d^n)_{j,l} + x_{i,l}(I_d^n)_{j,k} + x_{j,k}(I_d^n)_{i,l} + x_{j,l}(I_d^n)_{i,k}] \partial_{x_{i,j}} \partial_{x_{k,l}} \\
&+ \sum_{1\le k,l\le d} \left\{ \epsilon^2 \sum_{1\le i,j\le d} \Delta D_{i,j} [x_{i,k}(I_d^n)_{j,l} + x_{i,l}(I_d^n)_{j,k} + x_{j,k}(I_d^n)_{i,l} + x_{j,l}(I_d^n)_{i,k}] \right. \\
&\left. + \sum_{m=1}^{p} \epsilon \Delta B_m [(cx)_{m,k}(I_d^n \rho)_l + (cx)_{m,l}(I_d^n \rho)_k] \right\} \partial_{x_{k,l}} \partial_h.
\end{aligned}
$$

Some simplifications occur when we apply this operator to P since it is a function of a symmetric matrix, and therefore satisfies $\partial_{x_{i,j}} P = \partial_{x_{j,i}} P$. Thus, we have

$$
\forall x \in \mathcal{S}_d(\mathbb{R}), h \in \mathbb{R},\ L(t)P(t,x,h) = \tilde{L}(t)P(t,x,h),
$$

where

$$
\begin{aligned}
\tilde{L}(t) =& \frac{1}{2}\left(\Delta B^\top c x c^\top \Delta B + 4\epsilon^2 \mathrm{Tr}(\Delta D I_d^n \Delta D x) + 2\epsilon(\Delta B^\top c x \Delta D I_d^n \rho)\right)(\partial_h^2 - \partial_h) \\
&+ \sum_{1\le i,j\le d} (\Omega + (d-1)\epsilon^2 I_d^n + 2x(b^{T+\delta})^\top)_{i,j} \partial_{x_{i,j}} + 2 \sum_{1\le i,j,k,l\le d} \epsilon^2 x_{i,k}(I_d^n)_{j,l} \partial_{x_{i,j}} \partial_{x_{k,l}} \\
&+ \sum_{1\le k,l\le d} \left\{ 4\epsilon^2 \sum_{1\le i,j\le d} \Delta D_{i,j} x_{i,k}(I_d^n)_{j,l} + \sum_{m=1}^{p} 2\epsilon \Delta B_m [(cx)_{m,k}(I_d^n \rho)_l] \right\} \partial_{x_{k,l}} \partial_h.
\end{aligned}
$$

We assume that P admits a second order expansion

$$
P = P_0 + \epsilon P_1 + \epsilon^2 P_2 + o(\epsilon^2). \tag{4.15}
$$

Our goal is to explain how to calculate in a quite explicit way the value of P_0, P_1 and P_2, in order to show the tractability of the model. We assume in our derivations that these functions P_0, P_1 and P_2 are smooth enough. To determine the value of P_0, P_1 and P_2, we proceed as Bergomi and Guyon [BG12] and make an expansion of the generator $\tilde{L}(t) = \tilde{L}_0(t) + \epsilon \tilde{L}_1(t) + \epsilon^2 \tilde{L}_2(t) + \dots$ in order to obtain the PDEs satisfied by P_0, P_1 and P_2. We first note that B does not depend on ϵ and that

$$
b^{T+\delta}(t) = b + \epsilon I_d^n \rho B(T+\delta-t)^\top c + 2\epsilon^2 I_d^n D(T+\delta-t).
$$

We only need an expansion up to order 1 for D, and we get from (3.9) that $D(t) = D_0(t) + \epsilon D_1(t) + O(\epsilon^2)$ with $\dot{D}_0 = D_0 b + b D_0 + \frac{1}{2} c^\top B B^\top c - \gamma$, $D_0(0) = 0$ and $\dot{D}_1 = D_1 b + b D_1 +$

$\frac{1}{2}D_0 I_d^n \rho B^\top c + \frac{1}{2} c^\top B \rho^\top I_d^n D_0$, $D_1(0) = 0$. We then obtain

$$D_0(t) = e^{b^\top t}\left(\int_0^t e^{-b^\top s}\left(\frac{1}{2}c^\top B(s)B(s)^\top c - \gamma\right)e^{-bs}ds\right)e^{bt} \tag{4.16}$$

$$D_1(t) = \frac{1}{2}e^{b^\top t}\left(\int_0^t e^{-b^\top s}\left(c^\top B(s)\rho^\top I_d^n D_0(s) + D_0(s) I_d^n \rho B(s)^\top c\right)e^{-bs}ds\right)e^{bt}. \tag{4.17}$$

The order 0 term is easily determined:

$$\tilde{L}_0(t) = \left(\Delta B^\top c x c^\top \Delta B\right)\frac{1}{2}(\partial_h^2 - \partial_h) + \sum_{1\le i,j\le d}(\Omega + bx + xb)^\top)_{i,j}\partial_{x_{i,j}}.$$

This is the operator associated to the following diffusion

$$\begin{aligned} dX_t^0 &= (\Omega + bX_t^0 + X_t^0 b^\top)dt, \\ dH_t^0 &= -\frac{1}{2}\left(\Delta B^\top c X_t^0 c^\top \Delta B\right)dt + \Delta B^\top c\sqrt{X_t^0}dZ_t. \end{aligned}$$

The first and second order terms in the expansion of the generator are given by

$$\tilde{L}_1(t) = (\Delta B^\top c x \Delta D_0 I_d^n \rho)(\partial_h^2 - \partial_h) + 2\sum_{1\le i,j\le d}(xc^\top B\rho^\top I_d^n)_{i,j}\partial_{x_{i,j}} \tag{4.18}$$

$$+ 2\sum_{1\le k,l\le d}\sum_{m=1}^{p}\Delta B_m[(cx)_{m,k}(I_d^n\rho)_l]\partial_{x_{k,l}}\partial_h \tag{4.19}$$

$$\tilde{L}_2(t) = \left(2\mathrm{Tr}(\Delta D_0 I_d^n \Delta D_0 x) + (\Delta B^\top c x \Delta D_1 I_d^n \rho)\right)(\partial_h^2 - \partial_h) + \sum_{1\le i,j\le d}((d-1)I_d^n + 4xD_0 I_d^n)_{i,j}\partial_{x_{i,j}}$$

$$+ 2\sum_{1\le i,j,k,l\le d}x_{i,k}(I_d^n)_{j,l}\partial_{x_{i,j}}\partial_{x_{k,l}} + \sum_{1\le k,l\le d}\left\{4\sum_{1\le i,j\le d}(\Delta D_0)_{i,j}x_{i,k}(I_d^n)_{j,l}\right\}\partial_{x_{k,l}}\partial_h. \tag{4.20}$$

By plugging the expansions of the generator and the price in the pricing PDE and identifying each order in ϵ, we get the sequence of PDEs that are satisfied by P_0, P_1 and P_2:

$$\begin{aligned} \partial_t P_0 + \tilde{L}_0(t)P_0 &= 0, \quad P_0(T,x,h) = (e^h - (1+\delta K))^+, \\ \partial_t P_1 + \tilde{L}_0(t)P_1 + \tilde{L}_1(t)P_0 &= 0, \quad P_1(T,x,h) = 0, \\ \partial_t P_2 + \tilde{L}_0(t)P_2 + \tilde{L}_2(t)P_0 + \tilde{L}_1(t)P_1 &= 0, \quad P_2(T,x,h) = 0. \end{aligned}$$

Let $\mathrm{BS}(h,v) = \mathbb{E}\left[\left(\exp\left(h - \frac{1}{2}v + \sqrt{v}G\right) - (1+\delta K)\right)^+\right]$ with $G \sim N(0,1)$ denote the Black-Scholes price with realized volatility v. It satisfies the following PDE

$$\partial_v \mathrm{BS}(h,v) = \frac{1}{2}\left(\partial_h^2 - \partial_h\right)\mathrm{BS}(h,v), \tag{4.21}$$

that links the gamma, the delta and the vega. It will be used intensively in our calculations. We get that

$$P_0(t,x,h) = \mathrm{BS}(h, v(t,T,\delta,x)),$$

where

$$v(t,T,\delta,x) = \int_t^T \Delta B(u,T,\delta)^\top c e^{b(u-t)} \left(x + \int_t^u dz e^{-b(z-t)} \Omega e^{-b^\top(z-t)} \right) e^{b^\top(u-t)} c^\top \Delta B(u,T,\delta) du. \tag{4.22}$$

In fact, $X_u^0 = e^{b(u-t)} \left(x + \int_t^u dz e^{-b(z-t)} \Omega e^{-b^\top(z-t)} \right) e^{b^\top(u-t)}$ when $X_t^0 = x$. The higher order terms are given by

$$P_1(t,x,h) = \mathbb{E}\left[\int_t^T \tilde{L}_1(s) P_0(s, X_s^0, H_s^0) ds \middle| H_t^0 = h, X_t^0 = x \right] \tag{4.23}$$

$$P_2(t,x,h) = \mathbb{E}\left[\int_t^T (\tilde{L}_1(s) P_1(s, X_s^0, H_s^0) + \tilde{L}_2(s) P_0(s, X_s^0, H_s^0)) ds \middle| H_t^0 = h, X_t^0 = x \right] \tag{4.24}$$

We can calculate explicitly these quantities by using on the one hand the Gamma-Vega relationship (4.21) and on the other hand the fact that $\partial_h^k P_0(s, X_s^0, H_s^0)$ is a martingale. The details of the calculations are provided in Appendix E.1, and we obtain that

$$P_1(t,x,h) = \left(c_1(t,T,\delta,x)(\partial_h^3 - \partial_h^2) + c_2(t,T,\delta,x)(\partial_h^2 - \partial_h) \right) P_0(t,x,h) \tag{4.25}$$

and

$$\begin{aligned} P_2(t,x,h) = \Bigg[& \left(d_1(t,T,\delta,x)(\partial_h^2 - \partial_h)^2 + d_2(t,T,\delta,x)(\partial_h^2 - \partial_h)\partial_h + d_3(t,T,\delta,x)(\partial_h^2 - \partial_h) \right) \\ & + \Big(e_1(t,T,\delta,x)(\partial_h^2 - \partial_h)^2 \partial_h^2 + e_2(t,T,\delta,x)(\partial_h^2 - \partial_h)^2 \partial_h + e_3(t,T,\delta,x)(\partial_h^2 - \partial_h)^2 \\ & + e_4(t,T,\delta,x)(\partial_h^2 - \partial_h)\partial_h^2 + e_5(t,T,\delta,x)(\partial_h^2 - \partial_h)\partial_h + e_6(t,T,\delta,x)(\partial_h^2 - \partial_h) \Big) \Bigg] P_0(t,x,h). \end{aligned} \tag{4.26}$$

The formulas for the d_i's and the e_i's are given in Appendix E.1.

Expansion of the volatility

From the price expansion we can derive an expansion for the implied volatility. We denote by v_{Imp} the implied volatility of the Caplet and have by definition

$$\delta \text{FCaplet}(t,T,\delta) = \text{BS}(h, v_{\text{Imp}}).$$

We assume that the implied volatility satisfies an expansion with respect to ϵ up to order 2, i.e.

$$v_{\text{Imp}} = v_0 + \epsilon v_1 + \epsilon^2 v_2 + o(\epsilon^2).$$

By expanding the Black-Scholes price we get the following second order expansion for the caplet price:

$$\delta\text{FCaplet}(t,T,\delta) = \text{BS}(h,v_0) + \epsilon v_1 \partial_v \text{BS}(h,v_0) + \epsilon^2 \left(v_2 \partial_v \text{BS}(h,v_0) + \frac{v_1^2}{2} \partial_v^2 \text{BS}(h,v_0) \right) + o(\epsilon^2). \tag{4.27}$$

This has to be equal to $P_0(t,x,h) + \epsilon P_1(t,x,h) + \epsilon^2 P_2(t,x,h) + o(\epsilon^2)$, and we obtain by identification of each order in ϵ that

$$P_0(t,x,h) = \text{BS}(h,v_0), \; v_1 = \frac{P_1(t,x,h)}{\partial_v \text{BS}(h,v_0)} \text{ and } v_2 = \frac{P_2(t,x,h) - \frac{v_1^2}{2} \partial_v^2 \text{BS}(h,v_0)}{\partial_v \text{BS}(h,v_0)}.$$

The first equality gives

$$v_0 = v(t,T,\delta,x). \tag{4.28}$$

From (4.25), (4.21) and the formulas given in Appendix C, we obtain

$$\frac{v_1}{2} = c_2(t,T,\delta,x) + c_1(t,T,\delta,x) \left(\frac{1}{2} - \frac{h - \log(1+\delta K)}{v_0} \right). \tag{4.29}$$

With the same arguments, we can also obtain an explicit but cumbersome formula for v_2.

It is worth to underline at this stage that neither c_1 nor c_2 depend on the strike, see formulas (E.3) and (E.4). Therefore, the skew is at the first order in ϵ proportional to c_1, that is at its turn a linear function of ρ. We have in particular a flat smile at the first order when $\rho = 0$, as one may expect. However, we can also check that the second order term is different from zero when $\rho = 0$ and therefore the smile is not flat, at the second order. Besides ρ, it is interesting to notice that the coefficient c_1 depends on the factor x and on all the parameters involved in the (mean-reverting) drift, namely κ, θ, b and Ω. These coefficients determine the short and long term behaviour of the implied volatility surface.

4.3.2 Price and volatility expansion for Swaptions

From (1.23), the only quantity of interest in order to understand the swaptions volatility cube is what we call the annuity-forward swaption price

$$\text{AFSwaption}(t,T,m,\delta) = \mathbb{E}^A \left[(S_t(T,m,\delta) - K)^+ | \mathcal{F}_t \right].$$

It is standard to view swaptions as a basket option of forward Libor rates with stochastic weights, we have

$$S_t(T,m,\delta) = \sum_{i=1}^m \omega_t^i L_t(T + (i-1)\delta, T + i\delta) \tag{4.30}$$

$$\omega_t^i = \frac{P_{t,T+i\delta}}{\sum_{i=1}^m P_{t,T+i\delta}}. \tag{4.31}$$

The difficulty here comes from the fact that forward Libor rates, and the stochastic weights are complicated functions of the state variables (X, Y). The first implication is that the change of measure between $\mathbb{P}$ and Q^A is also complicated and the dynamics of the state variables under this new measure is quite unpleasant to work with. The second implication is that we cannot directly operate a convenient change of variable as we did for caplets. In order to derive an expansion for swaptions we thus proceed stepwise. First, we use a standard approximation that freezes the weights at their initial value (see for example Brigo and Mercurio [BM06] p. 239, d'Aspremont [d'A03] and Piterbarg [Pit09]). This is justified by the fact the variation of the weights is less important than the variation of the forward Libor rates[1]. Second, we use a similar approximation for the swap rate. Thus, the approximated swap rate is an affine function of the underlying state variables, which enables us to take advantage of the affine structure of the model. Let us mention that this technique is similar to the quadratic approximation of the swap rate proposed by Piterbarg in [Pit09]. Finally we perform our expansion on the affine approximation of the swap rates and obtain the second order expansion (4.38), which is the main result of Subsection 4.3.2.

Dynamics of the factors under the annuity measure

The annuity measure knowing the information up to date t, $Q^A|\mathcal{F}_t$ is defined by

$$\left.\frac{dQ^A}{d\mathbb{P}}\right|_{\mathcal{F}_t} = e^{-\int_t^T r_s ds}\frac{A_T(T,m,\delta)}{A_t(T,m,\delta)}.$$

It comes from the martingale property of discounted asset prices under the risk neutral measure that

$$\frac{d\left(e^{-\int_0^t r_s ds}A_t(T,m,\delta)\right)}{e^{-\int_0^t r_s ds}A_t(T,m,\delta)} =$$

$$\sum_{i=1}^m \omega_t^i\left(B(T+i\delta-t)^\top c\sqrt{X_t}(dW_t\rho+\bar{\rho}dZ_t) + 2\epsilon\mathrm{Tr}\left(D(T+i\delta-t)\sqrt{X_t}dW_tI_d^n\right)\right). \quad (4.32)$$

From Girsanov theorem, the change of measure is given by

$$\begin{aligned} dW_t^A &= dW_t - \sqrt{X_t}\left(2\epsilon\sum_{i=1}^m \omega_t^i D(T+i\delta-t)I_d^n + c^\top B(T+i\delta-t)\rho^\top\right)dt, \\ dZ_t^A &= dZ_t - \bar{\rho}\sqrt{X_t}c^\top\sum_{i=1}^m \omega_t^i B(T+i\delta-t)dt. \end{aligned}$$

[1] To the best of our knowledge there have been very few attempts to quantify either theoretically or numerically this statement. In [d'A03] d'Aspremont investigates the accuracy of the approximation for pricing swaptions in the log-normal BGM model, he shows that the approximation is less efficient for long maturities and long tenors.

This allows us to calculate from (3.1) the dynamics of the state variables under the annuity measure Q^A:

$$dY_t = \left(\kappa(\theta - Y_t) + cX_t c^\top \sum_{k=1}^m \omega_t^k B(T+k\delta-t) + 2\epsilon c X_t \sum_{k=1}^m \omega_t^k D(T+k\delta-t) I_d^n \rho\right) dt \tag{4.33}$$
$$+c\sqrt{X_t}(\bar\rho dZ_t^A + dW_t^A \rho),$$

$$dX_t = (\Omega + \epsilon^2(d-1)I_d^n + b^A(t)X_t + X_t(b^A(t))^\top)dt + \epsilon\left(\sqrt{X_t} dW_t^A I_d^n + I_d^n (dW_t^A)^\top \sqrt{X_t}\right), \tag{4.34}$$

where $b^A(t) = b + \epsilon I_d^n \rho \sum_{k=1}^m \omega_t^k B(T+k\delta-t)^\top c + 2\epsilon^2 I_d^n \sum_{k=1}^m \omega_t^k D(T+k\delta-t)$. The stochastic weights ω_t^k are complicated functions of the state variables and the solution of the above SDE is not easily characterized, in particular its infinitesimal generator is not a priori affine in the state variables. It is a rather classic approach to freeze the weights ω_t^k to their value at 0, this leads to an approximated SDE for the state variables under the measure Q^A which is affine.

An affine approximation of the forward swap rate

The forward swap rate is a martingale under the annuity measure Q^A. Therefore, we can only focus on the martingale terms when applying Itô's formula to $\frac{P_{t,T} - P_{t,T+m\delta}}{\sum_{i=1}^m P_{t,T+i\delta}}$, and we get from (4.3) that

$$\begin{aligned}
dS_t(T,m,\delta) &= \omega_t^0 \left(B(T-t)^\top c\sqrt{X_t}(dW_t^A\rho + \bar\rho dZ_t^A) + 2\epsilon \mathrm{Tr}\left(D(T-t)\sqrt{X_t} dW_t^A I_d^n\right)\right) \\
&- \omega_t^m \left(B(T+m\delta-t)^\top c\sqrt{X_t}(dW_t^A\rho + \bar\rho dZ_t^A) + 2\epsilon \mathrm{Tr}\left(D(T+m\delta-t)\sqrt{X_t} dW_t^A I_d^n\right)\right) \\
&- S_t(T,m,\delta)\sum_{k=1}^m \omega_t^k \left(B(T+k\delta-t)^\top c\sqrt{X_t}(dW_t^A\rho + \bar\rho dZ_t^A) + 2\epsilon \mathrm{Tr}\left(D(T+k\delta-t)\sqrt{X_t} dW_t^A I_d^n\right)\right) \\
&= \left[\omega_t^0 B(T-t)^\top - \omega_t^m B(T+m\delta-t)^\top - S_t(T,m,\delta)\sum_{k=1}^m \omega_t^k B(T+k\delta-t)^\top\right] c\sqrt{X_t}(dW_t^A\rho + \bar\rho dZ_t^A) \\
&+ 2\epsilon \mathrm{Tr}\left(\left[\omega_t^0 D(T-t) - \omega_t^m D(T+m\delta-t) - S_t(T,m,\delta)\sum_{k=1}^m \omega_t^k D(T+k\delta-t)\right]\sqrt{X_t} dW_t^A I_d^n\right)
\end{aligned} \tag{4.35}$$

By a slight abuse of notations, we will now drop the (T,m,δ) dependence of the swap rate and simply denote by S_t its time t value. We now use the standard approximation that consists in freezing the weights ω_t^k and the value of the swap rate S_t in the right-hand side to their value at zero. We then have

$$dS_t = B^S(t)^\top c\sqrt{X_t}(dW_t^A\rho + \bar\rho dZ_t^A) + 2\epsilon \mathrm{Tr}\left(D^S(t)\sqrt{X_t} dW_t^A I_d^n\right), \tag{4.36}$$

where

$$(B,D)^S(t) \;=\; \omega_0^0(B,D)(T-t) - \omega_0^m(B,D)(T+m\delta-t) - S_0(T,m,\delta)\sum_{k=1}^m \omega_0^k(B,D)(T+k\delta-t).$$

These coefficient are time-dependent and deterministic. We do the same approximation on X and get

$$dX_t = (\Omega+\epsilon^2(d-1)I_d^n + b_0^A(t)X_t + X_t(b_0^A(t))^\top)dt + \epsilon\left(\sqrt{X_t}dW_t^A I_d^n + I_d^n (dW_t^A)^\top \sqrt{X_t}\right),$$

where

$$b_0^A(t) = b + \epsilon I_d^n \rho \sum_{k=1}^m \omega_0^k B(T+k\delta-t)^\top c + 2\epsilon^2 I_d^n \sum_{k=1}^m \omega_0^k D(T+k\delta-t). \tag{4.37}$$

Thanks to this approximation, we remark that the process, that we still denote by (S_t, X_t) for simplicity, is now affine. This enables us to use again the same argument as for the Caplet prices to get an expansion of the price. The only difference lies in the fact the expansion is around the Gaussian model rather then around the log-normal model.

The swaption price expansion

Let $P^S(t,x,s) = \mathbb{E}^A\left[(S_t-K)^+ | S_t = s, X_t = x\right]$ denote the price of the Swaption at time $t \in [0,T]$. This is a function of the symmetric matrix x, and we assume that it is smooth enough for all the following derivations. Then, it solve the following pricing PDE

$$\partial_t P^S + \tilde{L}(t)P^S = 0,\ t\in(0,T),\ P^S(T,x,s) = (s-K)^+,$$

with

$$\begin{aligned}\tilde{L}^S(t) =& \frac{1}{2}\left((B^S(t))^\top cxc^\top B^S(t) + 4\epsilon^2 \mathrm{Tr}(D^S(t)I_d^n D^S(t)x) + 2\epsilon((B^S(t))^\top cxD^S(t)I_d^n\rho)\right)\partial_s^2 \\ &+ \sum_{1\le i,j\le d}(\Omega+(d-1)\epsilon^2 I_d^n + 2x(b_0^A(t))^\top)_{i,j}\partial_{x_{i,j}} + 2\sum_{1\le i,j,k,l\le d}\epsilon^2 x_{i,k}(I_d^n)_{j,l}\partial_{x_{i,j}}\partial_{x_{k,l}} \\ &+ \sum_{1\le k,l\le d}\left\{4\epsilon^2\sum_{1\le i,j\le d} D_{i,j}^S(t)x_{i,k}(I_d^n)_{j,l} + \sum_{m=1}^p 2\epsilon B_m^S(t)[(cx)_{m,k}(I_d^n\rho)_l]\right\}\partial_{x_{k,l}}\partial_s.\end{aligned}$$

Again, we assume that P^S admits a second order expansion

$$P^S = P_0^S + \epsilon P_1^S + \epsilon^2 P_2^S + o(\epsilon^2) \tag{4.38}$$

and that the functions P_0^S, P_1^S and P_2^S are smooth enough for what follows. To determine the value of P_0^S, P_1^S and P_2^S, we have to make as for the caplets an expansion of the generator

$\tilde{L}^S(t) = \tilde{L}^S_0(t) + \epsilon \tilde{L}^S_1(t) + \epsilon^2 \tilde{L}^S_2(t) + \dots$. We observe that B^S do not depend on ϵ, while we have

$$D^S = D^S_0 + \epsilon D^S_1 + o(\epsilon),$$

with

$$D^S_i(t) = \omega^0_0 D_i(T-t) - \omega^m_0 D_i(T+m\delta-t) - S_0(T,m,\delta)\sum_{k=1}^m \omega^k_0 D_i(T+k\delta-t), \quad i=0,1,$$

where the functions D_0 and D_1 are given by (4.16) and (4.17). For convenience, we also introduce the notation

$$(B,D_i)^A(t) = \sum_{k=1}^m \omega^k_0 (B,D_i)^A(T+k\delta-t).$$

From (4.37), we get

$$\begin{aligned}
\tilde{L}^S_0(t) =& \frac{1}{2}(B^S(t))^\top cxc^\top B^S(t)\partial^2_s + \sum_{1\le i,j\le d}(\Omega + (d-1)\epsilon^2 I^n_d + 2xb^\top)_{i,j}\partial_{x_{i,j}},\\
\tilde{L}^S_1(t) =& ((B^S(t))^\top cx D^S_0(t) I^n_d \rho)\partial^2_s + \sum_{1\le i,j\le d}(2xc^\top B^A(t)\rho^\top I^n_d)_{i,j}\partial_{x_{i,j}}\\
&+ \sum_{1\le k,l\le d}\sum_{m=1}^p 2B^S_m(t)[(cx)_{m,k}(I^n_d\rho)_l]\partial_{x_{k,l}}\partial_s, \qquad (4.39)\\
\tilde{L}^S_2(t) =& [2\mathrm{Tr}(D^S_0(t) I^n_d D^S_0(t)x) + (B^S(t))^\top cx D^S_1(t) I^n_d \rho]\partial^2_s + \sum_{1\le i,j\le d}(4xD^A_0(t)I^n_d)_{i,j}\partial_{x_{i,j}} + (d-1)\sum_{1\le i\le n}\partial_{x_{i,i}}\\
&+ 2\sum_{1\le i,j,k,l\le d} x_{i,k}(I^n_d)_{j,l}\partial_{x_{i,j}}\partial_{x_{k,l}} + \sum_{1\le k,l\le d}\left\{4\sum_{1\le i,j\le d}(D^S_0(t))_{i,j}x_{i,k}(I^n_d)_{j,l}\right\}\partial_{x_{k,l}}\partial_s. \qquad (4.40)
\end{aligned}$$

We note that $\tilde{L}^S_0(t)$ is the operator associated to the following diffusion

$$\begin{aligned}
dX^0_t &= (\Omega + bX^0_t + X^0_t b^\top)dt,\\
dS^0_t &= -\frac{1}{2}\left[(B^S(t))^\top cX^0_t c^\top B^S(t)\right]dt + (B^S(t))^\top c\sqrt{X^0_t}dZ_t.
\end{aligned}$$

Let $\mathrm{BH}(s,v) = \mathbb{E}\left[(s+\sqrt{v}G-K)^+\right]$ with $G \sim N(0,1)$ denote the European call price with strike K in the Bachelier model with realized volatility $v>0$ and spot price $s\in\mathbb{R}$. It satisfies the heat equation

$$\partial_v \mathrm{BH}(s,v) = \frac{1}{2}\partial^2_s \mathrm{BH}(s,v), \qquad (4.41)$$

that links the gamma and the vega. The order zero term is the price one would obtain by assuming that the forward swap rate is a Normal process with time dependent volatility given, and we have

$$P_0^S(t,x,s) = \mathrm{BH}(s, v^S(t,T,x)), \tag{4.42}$$

where

$$v^S(t,T,x) = \int_t^T B^S(r)^\top c \left(e^{b(r-t)} \left(x + \int_t^r e^{-b(u-t)} \Omega e^{-b^\top (u-t)} \right) e^{b^\top (r-t)} du \right) c^\top B^S(r) dr. \tag{4.43}$$

Besides, we have

$$\begin{aligned} P_1^S(t,x,s) =& \mathbb{E}\left[\int_t^T \tilde{L}_1^S(u) P_0^S(u, X_u^0, S_u^0) du \middle| S_t^0 = s, X_t^0 = x \right], \\ P_2^S(t,x,s) =& \mathbb{E}\left[\int_t^T (\tilde{L}_1^S(u) P_1^S(u, X_u^0, S_u^0) + \tilde{L}_2^S(u) P_0^S(u, X_u^0, S_u^0)) du \middle| S_t^0 = s, X_t^0 = x \right], \end{aligned}$$

exactly like (4.23) and (4.24) for the caplets.

We finally obtain

$$P_1^S(t,x,s) = \left(c_1^S(t,T,x)\partial_s^3 + c_2^S(t,T,x)\partial_s^2 \right) \mathrm{BH}(s, v^S(t,T,x)), \tag{4.44}$$

$$\begin{aligned} P_2^S(t,x,s) = & \Big[d_1^S(t,T,x)\partial_s^4 + d_2^S(t,T,x)\partial_s^3 + d_3^S(t,T,x)\partial_s^2 \\ & + e_1^S(t,T,x)\partial_s^6 + e_2^S(t,T,x)\partial_s^5 + e_3^S(t,T,x)\partial_s^4 \\ & + e_4^S(t,T,x)\partial_s^4 + e_5^S(t,T,x)\partial_s^3 + e_6^S(t,T,x)\partial_s^2 \Big] \mathrm{BH}(s, v^S(t,T,x)), \end{aligned} \tag{4.45}$$

where the coefficients c_i^S, d_i^S and e_i^S are calculated explicitly in Appendix E.2.

Once we have the price expansion, we can easily represent it in terms of implied volatility. Let v_{Imp}^S denote the Normal implied volatility, so that $P^S(t,x,s) = \mathrm{BH}(s, v_{\mathrm{Imp}}^S)$. We assume that the implied volatility satisfies the expansion $v_{\mathrm{Imp}}^S = v_0^S + \epsilon v_1^S + \epsilon^2 v_2^S + o(\epsilon^2)$ and get that

$$\mathrm{BH}(s, v_0^S) = P_0^S(t,x,s), \ v_1^S = \frac{P_1^S(t,x,s)}{\partial_v \mathrm{BH}(s, v_0^S)} \text{ and } v_2^S = \frac{P_2^S(t,x,s) - \frac{(v_1^S)^2}{2} \partial_v^2 \mathrm{BH}(s, v_0^S)}{\partial_v \mathrm{BH}(s, v_0^S)}.$$

This yields to $v_0^S = v^S(t,T,x)$, $\frac{v_1^S}{2} = c_2^S(t,T,x) - c_1^S(t,T,x)\frac{s-K}{v_0^S}$ by using the formulas in Appendix C. An explicit but cumbersome formula can also be obtained for v_2^S. As for the caplets, the skew is at the first order in ϵ proportional to c_1, which is at its turn a linear function of ρ. Thus, we have a flat smile at the first order when $\rho = 0$, but we can also check that v_2^S does not vanish when $\rho = 0$ and the smile is not flat at the second order.

4.3.3 Numerical results

We now assess on some examples the accuracy of the expansions we have developed. In practice we are interested in knowing up to what level of parameters and for what set of maturities and tenors the accuracy of the expansion is satisfactory. Let us recall that our expansion for caplets results from the combination of two expansions, the first on the support matrix function D up to the order 1 in ϵ is given by (4.16) and (4.17), the second on the infinitesimal generator of the Markov process (X, H) defined by (4.14). By construction the approximation of $D(\tau)$ will be more accurate for a small τ. As a consequence, for a given set of parameters, the full expansion will likely to be more accurate for short maturities, short tenors caplets. The expansion for swaptions results from a supplementary approximation step, which consists in freezing the weights ω^i in the diffusion of the Markov process (X, S) defined by (4.34) and (4.35). This approximation can be inaccurate for long maturities and long tenors swaptions. Therefore, we expect the full expansion to be more accurate for short maturities, short tenors swaptions. Figure 4.6 shows the contribution of the terms P_i in the expansion of the price of the caplet. The zero order expansion is the implied volatility in the LGM model, and gives a flat caplet N smile. The first order expansion captures the skew of the smile and gives the slope of the smile. Note that the first order expansion does not change the level of the at-the-money implied Normal volatility. The second order expansion gives the convexity of the smile, and changes the level of the at-the-money volatility. The higher the value of ϵ the more the at-the-money volatility will depart from the level given by the zero order term, and the higher the convexity of the smile. However this is only true for small values of ϵ, which will be discussed at the end of the section.

We assess the quality of the price expansion for caplets and swaptions. We compare the expanded price with the price computed using Monte Carlo simulation and the discretization scheme 1 described in Section 3.2.2 on a regular time grid. The expanded prices and the Monte Carlo prices are compared in terms of the Normal implied volatility of the forward Libor rate for caplets and of the forward swap rate for swaptions. The implied volatility is given in basis points (10^{-4}). In abscissa is indicated the difference between the strike and the at-the-money value, and the unit is one percent. A $6M \times 2Y$ caplet will denote a caplet with maturity $T = 2$ years and tenor $\delta = 0.5$ years, while a $5Y \times 2Y$ swaption will denote a swaption with maturity $T = 2$ years and tenor $m\delta = 5$ years.

We have tested different set of model parameters. The parameters values have been chosen in such a way that the yield curve and volatility levels generated by the model are in line with today's US and EUR interest rates market levels.

Parameters set 1 is a $p = 2, d = 2$ model with the following set of model parameters:

$$\kappa = \text{diag}(0.1, 1),\ c = I_d,\ b = -\text{diag}(0.41, 0.011),\ \Omega = -(bx_\infty + x_\infty b^\top) + 0.4 I_d,\ \gamma = 0.001 I_d$$
$$x = 10^{-4} \begin{pmatrix} 2.25 & -1.2 \\ -1.2 & 1. \end{pmatrix},\ x_\infty = 10^{-4} \begin{pmatrix} 1. & -0.125 \\ -0.125 & 0.25 \end{pmatrix}.$$

Parameters set 2 is a $p = 4, d = 4$ model with the following set of model parameters:

$$\kappa = \mathrm{diag}(0.1, 0.2, 0.4, 1),\ c = I_d,\ b = -\mathrm{diag}(0.41, 0.21, 0.11, 0.011),\ \gamma = 0.1I_d,$$
$$\mathrm{diag}(\sigma_1, \sigma_2, \sigma_3, \sigma_4) = (1.5, 1.2, 1, 0.8),\ \mathrm{diag}(\bar{\sigma}_1, \bar{\sigma}_2, \bar{\sigma}_3, \bar{\sigma}_4) = (1, 0.8, 0.6, 0.4)$$
$$\tilde{C}_{i,j} = \mathbf{1}_{i=j} - 0.4\mathbf{1}_{|i-j|=1} - 0.2\mathbf{1}_{|i-j|>1},\ \bar{C}_{i,j} = \mathbf{1}_{i=j} - 0.2\mathbf{1}_{|i-j|=1} - 0.1\mathbf{1}_{|i-j|>1}$$
$$x_{i,j} = \sigma_i\sigma_j\tilde{C}_{i,j},\ (x_\infty)_{i,j} = \bar{\sigma}_i\bar{\sigma}_j\bar{C}_{i,j},\ \Omega = -(bx_\infty + x_\infty b^\top) + 0.4I_d.$$

In both parameter sets, $-(b + b^\top) = -2b$ is positive definite. We know from Remark 29 that the condition of non-explosion will be verified in general for these set of parameters when ϵ is small enough, and we have checked that the yield curve given by these parameters sets is well defined up to 50 years.

In all the graphics the dotted line gives the Monte Carlo smile obtained with 100000 simulation paths, the solid line with small arrows is the expanded smile, the two continuous solid lines are the upper and lower bounds of the 95% confidence interval of the Monte Carlo price. Figures 4.7 and 4.8 show the accuracy of the expansion for the valuation of caplets. The approximation is accurate for expiries up to 2 years and less accurate with the same parameters for longer expiries. For maturities up to 2 years, the at-the-money volatility of the expanded smile is almost identical to the Monte Carlo smile and the whole expanded smile stays within the 95% confidence interval. Figures 4.9 show the accuracy of the expansion for the valuation of swaptions. We observe that the expansion is more accurate for negative values of the correlation parameters ρ (a similar behaviour is observed for Caplets). This can be intuitively understood from the Riccati equation (3.9): a negative ρ pushes D to zero while a positive one pushes D away from zero, and the expansion that we use on D (see (4.16) and (4.17)) is then less accurate. Overall the expansion is accurate at-the-money and is much less accurate out-of-the-money. For example, the graphic on the right hand side of Figure 4.8 shows that the expanded smile of the 6 months maturity 5 years expiry smile is quite inaccurate and the expanded smile fails to fit the skew of the Monte Carlo smile. However, the difference in the at-the-money volatility between the expanded price and Monte Carlo is around 1 bp. Figure 4.10 shows the accuracy of the expansion for caplets and swaptions in a $d = 4, p = 4$ model. The accuracy of the expansion shows similar patterns than in a $p = 2, d = 2$ model, but the expansion may be less accurate when the dimension of the model increases. In fact, we have taken in Figure 4.10 a smaller value of ϵ for the Swaption case in order to have a good approximation of the expansion.

In both parameter sets 1 and 2, the value of γ is relatively small. In fact, there can be some side effects of choosing a high value for γ. The specification of our model is designed to be a perturbation of the LGM model. The parameter ϵ controls the size of the perturbation. At first glance we would expect the convexity of the volatility smile to increase as the parameter increases. However, the effect of ϵ on the smile shape is not so straightforward, and depends on the value of γ. This is illustrated in details in the following paragraph.

4.3.4 On the impact of the parameter ϵ

The specification of our model is designed to be a perturbation of the LGM model. The parameter ϵ controls the size of the perturbation. Let us recall that the yield curve is driven by both the factors Y and X. When ϵ is small, the main driver of the curve are the factors Y, and the factors X acts as the volatility of the curve. As ϵ increases, we have two different effects. On the one hand the volatility of volatility of the factors Y increases, which leads to an increase in the volatility of volatility of the curve, and implies a larger convexity of the smile. On the other hand the contribution of the factors X to the yield curve dynamics becomes more important. This second effect implies that both the yield curve and its volatility are driven by the factors X, which eventually means that the model is close to a local volatility model. As a consequence, an increase in the parameter ϵ will not necessarily lead to an increase in the convexity of the smile. Let us mention also that the accuracy of the expansion will also decay when γ takes relatively large values with respect to ϵ. The heuristic reason is the same as the one for positive values of the entries of ρ: large values of γ will push D away from zero (see the Riccati equation (3.9) with $\bar{\Gamma} = -\gamma$), and the expansion that we use on D is then less accurate.

Let us illustrate this through an example. Consider a one-dimensional model i.e. $p = 1, d = 1$. The volatility of the forward ZCB $P_{t,T}/P_{,T+\delta}$ under the $T+\delta$-forward measure is given by

$$
\begin{aligned}
d\frac{P_{t,T}}{P_{t,T+\delta}} &= \frac{P_{t,T}}{P_{t,T+\delta}} m(t)\sqrt{X_t}\left(\rho(t)dW_t + \bar{\rho}(t)dZ_t\right) \\
m(t)^2 &= \left(\Delta B(t,T,\delta) + 2\epsilon\Delta D(t,T,\delta)\right)^2 + \bar{\rho}^2\Delta B(t,T,\delta)^2 \\
\rho(t) &= \left(\Delta B(t,T,\delta) + 2\epsilon\Delta D(t,T,\delta)\right)/m(t).
\end{aligned}
$$

The functions m and ρ determine the shape of the smile. Under the non-explosion condition 4.2. The function m is decreasing as a function of t, the higher the value of ϵ, the higher the value of $m(0)$ and the higher the slope of the function m. The function ρ increases to 1 and conversely the function $\bar{\rho}$ decreases to 0. The higher the value of ϵ, the higher the lower the initial value of $\rho(0)$. This means that as time goes, the forward ZCB is mainly driven by the factor X, and the model becomes a local volatility model. See Figure 4.11. When the condition 4.2 is not verified, and we can observe explosion of the solution of the Riccati o.d.e. giving the bond reconstruction formula 4.3, the behavior of m and ρ is different. The function m is decreasing and then increasing (in fact it explodes in finite time). As the value of ϵ increases, the function m initially decreases faster and then also increases faster to explode at a earlier time. The function ρ decreases to negative values and the function $\bar{\rho}$ slowly decreases. This behavior is magnified for higher values of the parameter ϵ.

4.3.5 Analysis of the implied volatility

Now that we have asserted the quality of the implied volatility expansion, we can analyze the main drivers of the smile. The order zero term in the expansion (4.43) only depends on the parameters K, b and Ω and on the initial value of the factors X. This term sets the overall level of the smile and numerical experiments have proven that it is very close to the level of the at-the-money volatility which is given by

$$\sigma^S_{ATM} = \sigma^S_0 + \epsilon \frac{c^S_1}{\sqrt{v^S(T-t)}} + \epsilon^2 \left(\frac{d^S_1 + e^S_1}{\sqrt{v^S(T-t)}} - \frac{d^S_3 + e^S_4 + e^S_3}{\sqrt{v^S(T-t)}} + \frac{3e^S_6}{(v^S)^2\sqrt{v^S(T-t)}} \right). \tag{4.46}$$

As mentioned before the zero order term of the volatility σ^S_0 is the implied volatility one would obtain in the LGM model with a deterministic variance-covariance of the factors given by $(X^0_s)_s$. The model parameters b and Ω play respectively the role of a "mean reversion" and "long term mean" for the volatility factors X. The expansion also allows to express the at-the-money skew as follows

$$\partial_K \sigma^S_{ATM} = \epsilon \frac{c^S_2}{\sqrt{v^S(T-t)}} + \epsilon^2 \left(\frac{d^S_2 + e^S_2}{\sqrt{v^S(T-t)}} - \frac{d^S_3 + e^S_4 + e^S_3}{(v^S)^2\sqrt{v^S(T-t)}} + \frac{3e^S_5}{(v^S)^2\sqrt{v^S(T-t)}} \right). \tag{4.47}$$

To the first order the at-the-money skew is given by c^S_1, which is a correlation term between the part in Y of the forward swap rate $B^S(t)^\top Y$ and the volatility of the forward swap rate $\mathrm{Tr}\left(\partial_x v^S(t,T,x)x\right)$. The second order term of the skew is given by

$$\frac{d^S_2 + e^S_5}{\sqrt{v^S(T-t)}} - \frac{d^S_1 + e^S_3 + e^S_4}{(v^S)^2\sqrt{v^S(T-t)}} + \frac{3e^S_2}{(v^S)^2\sqrt{v^S(T-t)}}.$$

The term d^S_2 is a correlation term between the part in X of the forward swap rate $\mathrm{Tr}\left(D^S(t)X\right)$ and the volatility of the forward swap rate $\mathrm{Tr}\left(\partial_x v^S(t,T,x)x\right)$. Let us note that due to this term (as well as the term in d^S_1) the model exhibits a non zero at-the-money skew even when the parameter ρ is zero. This comes from the fact the yield curve and in particular the forward swap rate depend on the volatility factors X, implying a "natural" correlation between the forward swap rate and the volatility even when the correlation parameter ρ is zero. The other terms involved in the expansion are correlation terms of higher order.

4.4 Hedging in the SCVATSM

This section is devoted to the hedging of IR derivatives in the model. The model naturally belongs to the class of HJM AOA models [HJM92] and satisfies the conditions for the model to have a uniquely determined equivalent martingale measure under which the discounted prices of ZCBs are martingales[2]. From classic results [HP81] and [HK79] we thus have that the model is complete and thus any derivative can be replicated with a set of $p + d(d + 1)/2$ ZCBs $(P_{t,T_1}, ..., P_{t,T_{p+d(d+1)/2}})$ with distinct maturities. Strictly speaking the model is rather a local-volatility model than a stochastic volatility model, however our specification is specifically designed for the model to exhibits some traits of a stochastic volatility model. This is done by limiting the dependence of the yield curve on the volatility factors X. Let us note that the notion of completeness does not depend on the chosen market numeraire. Thorough we hedge with strategies in the T-forward neutral measure, where T stands for the maturity of the derivative we want to hedge, i.e. we chose the $P_{t,T}$ as a numeraire. This choice is classic in the interest rates market.

4.4.1 Traits of stochastic volatility

From an hedging point of view stochastic volatility introduces a stochastic dimension in the market which is not captured by the underlying market asset, or more precisely which cannot be replicated by only trading on the underlying asset. Stochastic volatility leads to incomplete markets in the sense that options cannot be replicated by portfolios which only include the numeraire and the underlying asset. Though our model is -strictly speaking- not a stochastic volatility model, the model exhibits some traits of stochastic volatility. In particular incompleteness -to a certain extent- is partly verified in the model.

An incomplete caplets and swaptions market

The IR market vanilla instruments are IRS, FRA, Caplets and Swaptions. From standard change of numeraire arguments the prices admit the following expressions.

$$
\begin{aligned}
FRA_t(T,\delta,K) &= P_{t,T+\delta}(L_T(T,\delta) - K) && (4.48)\\
IRS_t(T,\delta,n,K) &= P_{t,T} - P_{t,T+n\delta} - \delta \sum_{i=1}^{n} P_{t,T+i\delta} && (4.49)\\
C_t(T,\delta,K) &= P_{t,T}\mathbb{E}^{Q^T}\left[P_{T,T+\delta}(L_T(T,\delta) - K)|\mathcal{F}_t\right] && (4.50)\\
S_t(T,n,\delta,K) &= P_{t,T}\mathbb{E}^{Q^T}\left[\left(P_{T,T} - P_{T,T+\delta} - \delta K \sum_{i=1}^{n} P_{T,T+i\delta}\right)^+ |\mathcal{F}_t\right] && (4.51)
\end{aligned}
$$

We then define $l_t(T,\delta)$ by

$$\frac{P_{t,T+\delta}}{P_{t,T}} = \exp\left(l_t(T,\delta)\right). \tag{4.52}$$

Under Q^T, $\exp\left(l_t(T,\delta)\right)$ is a martingale, implying the following dynamics for $l_t(T,\delta)$:

[2] In our model this conditions mainly translate in the assumption that the functional relationship between the factors and the ZCBs or forward swap rates is invertible.

$$dl_t(T,\delta) = -\frac{1}{2}\left(\Delta b^\top cXc^\top\Delta b + 4\epsilon^2\mathrm{Tr}(\Delta dI_d^n\Delta dX) + 2\epsilon(\Delta b^\top cX\Delta dI_d^n\rho)\right)dt \tag{4.53}$$

$$+\Delta b(t,T,\delta)^T c\sqrt{X}\,(dW\rho + \bar{\rho}dZ) + 2\mathrm{Tr}\left(\Delta d(t,T,\delta)\sqrt{X}dWI_d^n\right)$$

where we have defined

$$\Delta(b,d) = (B,D)(T+\delta-t) - (B,D)(T-t).$$

We thus deduce that $(l_t(T_1,\delta),...,l_t(T_p,\delta),X_t)$ is a Markov process under Q^T. Then we have

$$C_t(T,\delta,K) = P_{t,T}VFC\,(t,T,\delta,l_t(T,\delta),K) \tag{4.54}$$
$$S_t(T,n,\delta,K) = P_{t,T}VFS\,(t,T,\delta,l_t(T,\delta),...,l_t(T,n\delta),K)\,. \tag{4.55}$$

For ease of notations we will drop the notation (T,δ,K) and (T,n,δ,K) when there is no ambiguity, and denote simply FRA_t, IRS_t, C_t and S_t respectively the FRA, Caplet and Swaption price. The prices dynamics under Q^T is given by

$$d\frac{FRA_t}{P_{t,T}} = -\frac{1+\delta K}{\delta}\frac{P_{t,T+\delta}}{P_{t,T}}\left(\Delta b(t,T,\delta)c\sqrt{X}(dW\rho+\bar{\rho}dZ) + 2\mathrm{Tr}\left(\Delta d(t,T,\delta)\sqrt{X}dWI_d^n\right)\right) \tag{4.56}$$

$$d\frac{IRS_t}{P_{t,T}} = -\frac{P_{t,T+n\delta}}{P_{t,T}}\left(\Delta b(t,T,n\delta)c\sqrt{X}(dW\rho+\bar{\rho}dZ) + 2\mathrm{Tr}\left(\Delta d(t,T,n\delta)\sqrt{X}dWI_d^n\right)\right)$$
$$-\ \delta K\sum_{i=1}^{n}\frac{P_{t,T+i\delta}}{P_{t,T}}\left(\Delta b(t,T,i\delta)c\sqrt{X}(dW\rho+\bar{\rho}dZ) + 2\mathrm{Tr}\left(\Delta d(t,T,i\delta)\sqrt{X}dWI_d^n\right)\right) \tag{4.57}$$

$$d\frac{C_t}{P_{t,T}} = \partial_l VFC\left(\Delta b(t,T,\delta)c\sqrt{X}(dW\rho+\bar{\rho}dZ) + 2\mathrm{Tr}\left(\Delta d(t,T,\delta)\sqrt{X}dWI_d^n\right)\right)$$
$$+\ 2\mathrm{Tr}\left(\partial_x VFC\sqrt{X}dWI_d^n\right) \tag{4.58}$$

$$d\frac{S_t}{P_{t,T}} = \sum_{i=1}^{n}\partial_{l_i}VFS\left(\Delta b(t,T,i\delta)c\sqrt{X}(dW\rho+\bar{\rho}dZ) + 2\mathrm{Tr}\left(\Delta d(t,T,i\delta)\sqrt{X}dWI_d^n\right)\right)$$
$$+\ 2\mathrm{Tr}\left(\partial_x VFS\sqrt{X}dWI_d^n\right) \tag{4.59}$$

Proposition 38 — *Caplets cannot be hedged by a portfolio made of only FRAs of same tenor and maturity.*

Proof : Let C_t denote the price of a caplet of maturity T, strike K and tenor δ. Suppose there exists a set of FRAs $FRA^1,...,FRA^r$ with fixed rate $K_1,...,K_r$ and a self-financed portfolio $\phi=(\phi_1,...,\phi_r)$ that replicates C. Then we would have

$$d\left(\frac{C_t}{P_{t,T}}\right) = d\left(\sum_{i=1}^{r} \phi_t^i \frac{FRA_t^i}{P_{t,T}}\right).$$

From self-financing we then have

$$d\left(\frac{C_t}{P_{t,T}}\right) = \sum_{i=1}^{r} \phi_t^i d\left(\frac{FRA_t^i}{P_{t,T}}\right)$$

In particular this equality hods under the T-forward measure, using (4.56) and (4.58) this implies

$$\begin{aligned}
-\left(\sum_{i=1}^{r} \phi_t^i \frac{1+\delta K_i}{\delta}\right) \frac{P_{t,T+\delta}}{P_{t,T}} &= \partial_l VFC \\
-\left(\sum_{i=1}^{r} \phi_t^i \frac{1+\delta K_i}{\delta}\right) \frac{P_{t,T+\delta}}{P_{t,T}} \Delta d(t,T,\delta) &= \partial_l VFC \Delta d(t,T,\delta) + \partial_x VFC,
\end{aligned}$$

which leads to a contradiction since it would imply $\partial_x VFC = 0$. □

Proposition 39 — *Swaptions cannot be hedged by a portfolio made of only IRSs of same tenor and maturity.*

The proof of this proposition is identical to the proof of the preceding one.

Remark 40 — *Let us note that, in the model, both caplets and swaptions could be perfectly replicated by choosing a set of FRA and IRS of different maturities. This is the price to pay to have Markovian term structure model.*

Stated that swaptions cannot be hedged by trading the underlying IRS only, an interesting question is to know wether it is possible to hedge swaptions by trading in the underlying IRS and in a portfolio of caplets of the same expiry. As mentioned before the model exhibits a stochastic correlation between rates, therefore, since the swaptions prices depend on the correlation between the forward libor rates, it is not immediate to give a positive answer to this question. However, by looking at the equations (4.58) and (4.59), we see that, in theory, we can replicate the diffusion term of the dynamics of S_t with a well chosen set of caplets.

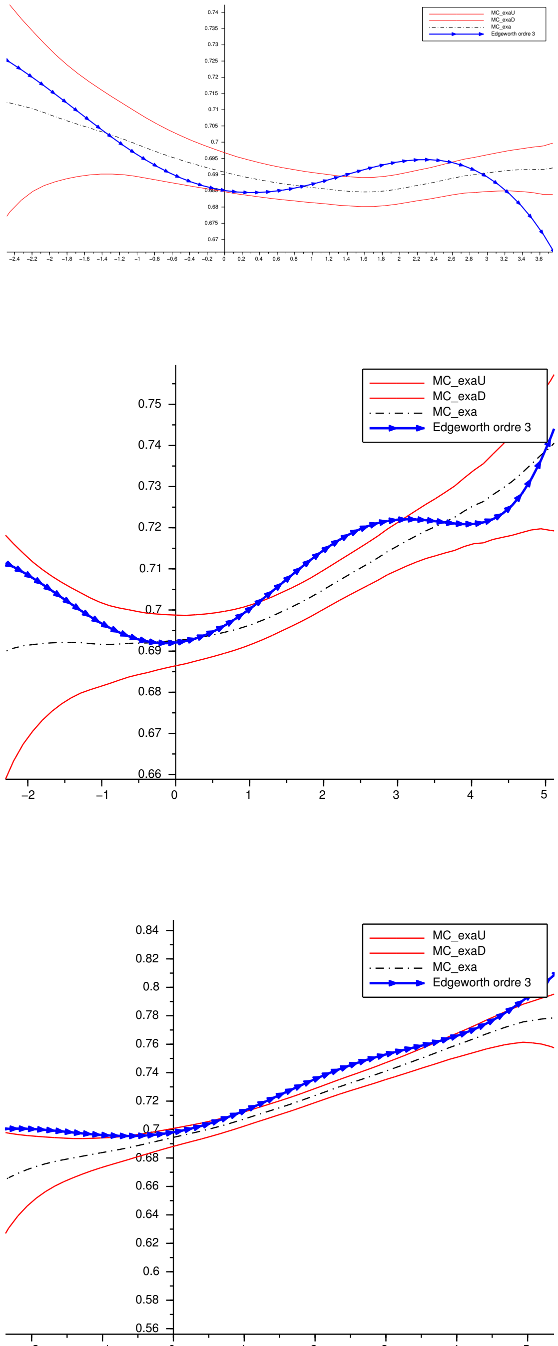

Figure 4.3: Normal implied volatility smile of a $3M \times 5Y$ caplet using Edgeworth series expansion of the cumulative distribution function of the ZCB $P_{T,T+\delta}$ with $T = 5Y$ and $\delta = 3M$. We have used the 3^{th} order expansion. The origin represent the ATM forward libor rate $L_t(T, \delta)$. The Monte Carlo price is obtained with 10^5 simulation paths and a $N = 30$ discretization steps. The figures represent different values of the correlation parameter ρ, starting from the left $\rho = -0.2\mathbf{1}_2, \rho = 0$ and $\rho = 0.2\mathbf{1}_2$. The higher order moments of the ZCB are computed using the numerical resolution of the MRDE described in 3.1.4, with a time discretization step of 0.01.

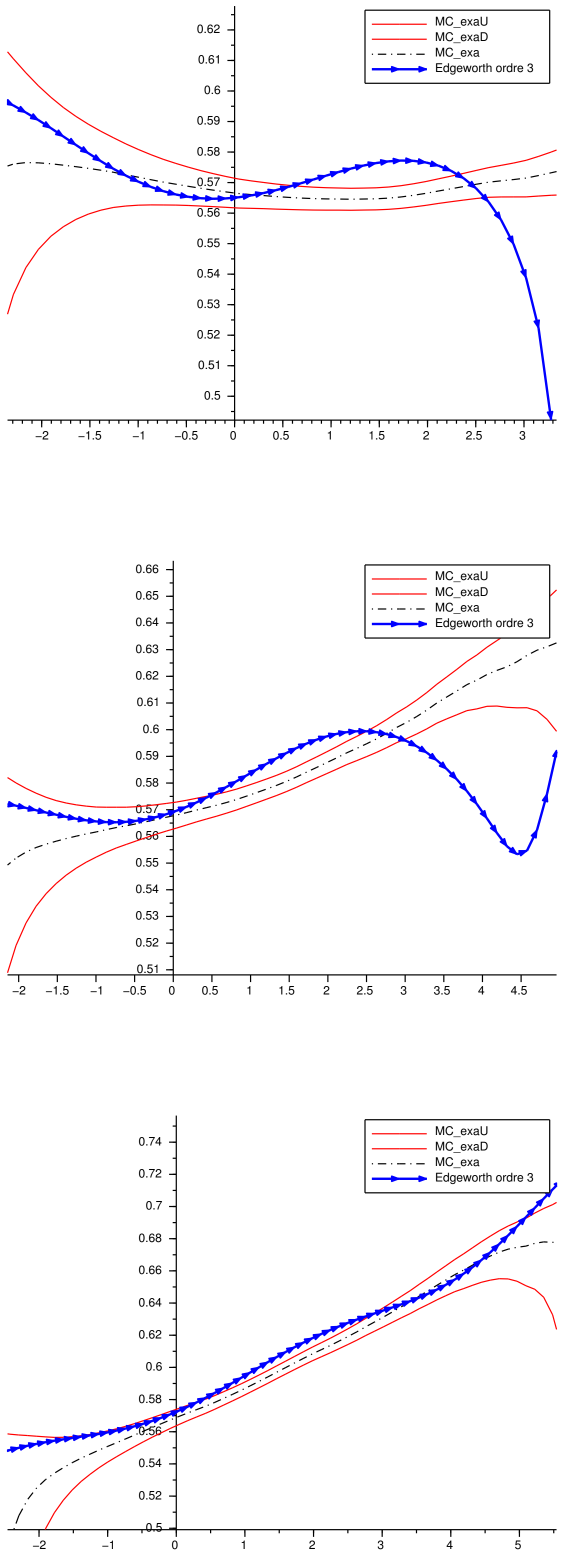

Figure 4.4: Normal implied volatility smile of a $2Y \times 5Y$ caplet using Edgeworth series expansion of the cumulative distribution function of the ZCB $P_{T,T+\delta}$ with $T = 5Y$ and $\delta = 2Y$. We have used the 3^{th} order expansion. The origin represent the ATM forward libor rate $L_t(T,\delta)$. The Monte Carlo price is obtained with 10^5 simulation paths and a $N = 30$ discretization steps. The figures represent different values of the correlation parameter ρ, starting from the left $\rho = -0.2\mathbf{1}_2, \rho = 0$ and $\rho = 0.2\mathbf{1}_2$. The higher order moments of the ZCB are computed using the numerical resolution of the MRDE described in 3.1.4, with a time discretization step of 0.01.

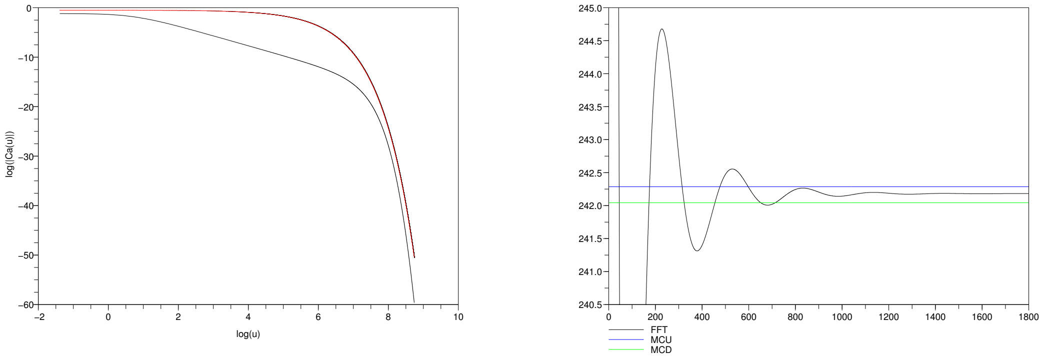

Figure 4.5: Analysis of the convergence of the FFT price of the 1Yx1Y Caplet in the 2Fx2F model at a given strike. Left: plot of the norm of the modified Fourier transform in log abscises (black curve), and of the upper bound $-0.5 - 0.008e^u$ (red curve). Right: Convergence of Σ_N towards the price. The higher order moments of the ZCB are computed using the numerical resolution of the MRDE described in 3.1.4, with a time discretization step of 0.01.

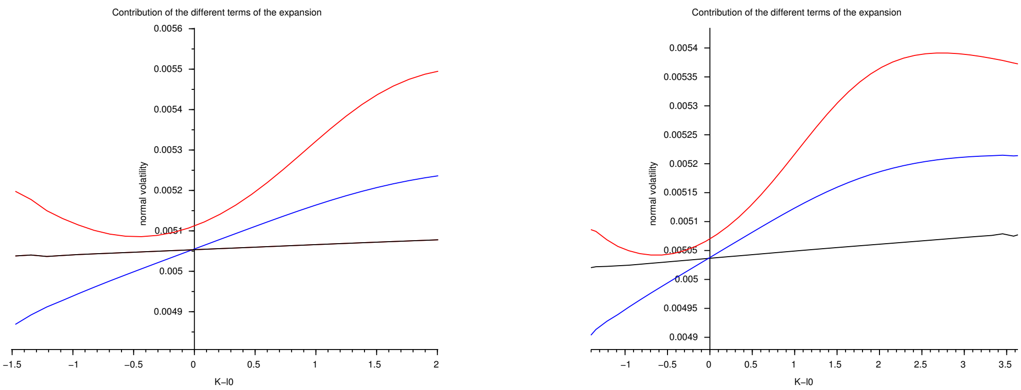

Figure 4.6: Parameter set 1 and setting $\rho^\top = (0.4, 0.2)$. Contribution of the expansion terms V_i on the expansion of the smile of the 1 year expiry 6 month maturity caplet for different values of the parameter ϵ, respectively $\epsilon = 0.002$ and $\epsilon = 0.0015$.

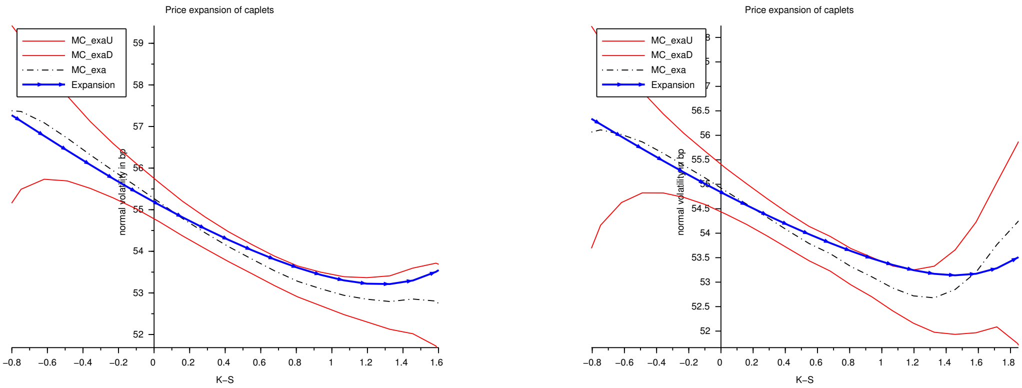

Figure 4.7: Parameter set 1 and $\rho^\top = (-0.4, -0.2)$. Plot of the expanded smile of a $1Y \times 1Y$ caplet against the Monte Carlo smile obtained with 100000 paths and a discretization grid of 4 points for different values of the parameter ϵ, respectively from left to right $\epsilon = 0.002$ and $\epsilon = 0.0015$. The forward Libor rate value is $L(0, 1Y, 1Y) = 1.02\%$.

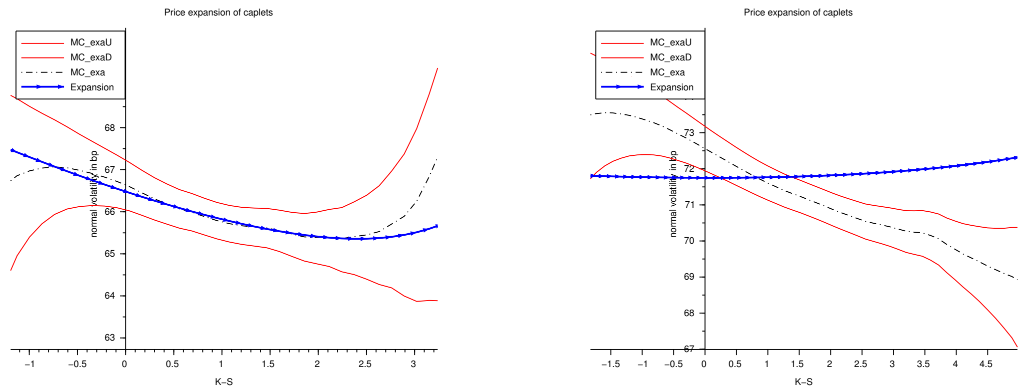

Figure 4.8: Parameter set 1 and setting the model parameters $\rho^\top = (-0.4, -0.2)$ and $\epsilon = 0.0015$. Left: plot of the expanded smile of a $6M \times 2Y$ caplet against the Monte Carlo smile. Right: plot of the expanded smile of a $6M \times 5Y$ caplet against the Monte Carlo smile. The Monte Carlo smile is obtained with 100000 paths and a discretization grid of 8 points. The forward Libor rates values are $L(0, 6M, 2Y) = 1.14\%$ and $L(0, 6M, 5Y) = 1.35\%$.

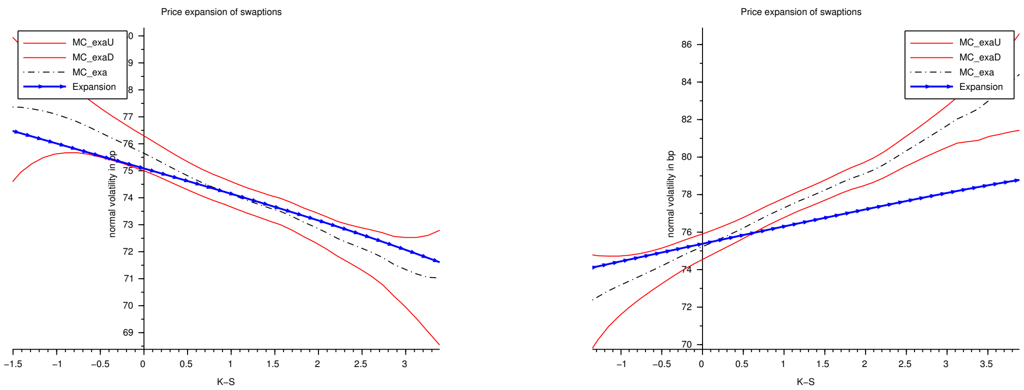

Figure 4.9: Parameter set 1 and setting $\epsilon = 0.0015$. Plot of the expanded smile of a $5Y \times 2Y$ swaption with coupon payment frequency of 6 months against the Monte Carlo smile obtained with 100000 paths and a discretization grid of 8 points for different values of the parameter ρ, from left to right $\rho^\top = (-0.4, -0.2)$ and $\rho^\top = (0.4, 0.2)$. The forward swap rate value is $S(0, 5Y, 2Y) = 1.3\%$.

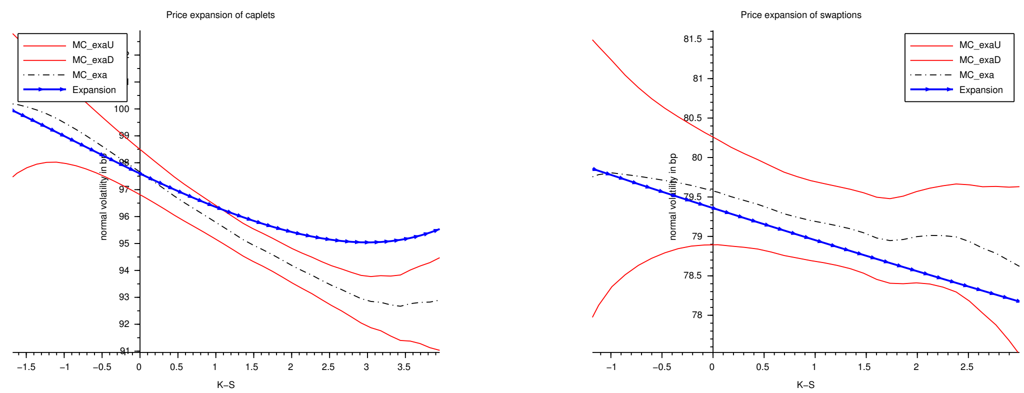

Figure 4.10: Parameter set 2 and setting $\rho^\top = -0.2\mathbf{1}_4$. Left: $\epsilon = 0.00125$, plot of the expanded smile of a $6M \times 2Y$ Caplet against the Monte Carlo smile obtained with 100000 paths and a discretization grid of 8 points. The forward Libor rate value is $L(0, 6M, 2Y) = 1.64\%$. Right: $\epsilon = 0.0005$, plot of the expanded smile of a $5Y \times 2Y$ swaption with coupon payment frequency of 6 months against the Monte Carlo smile obtained with 100000 paths and a discretization grid of grid of 8 points. The forward swap rate value is $S(0, 5Y, 2Y) = 2.04559\%$

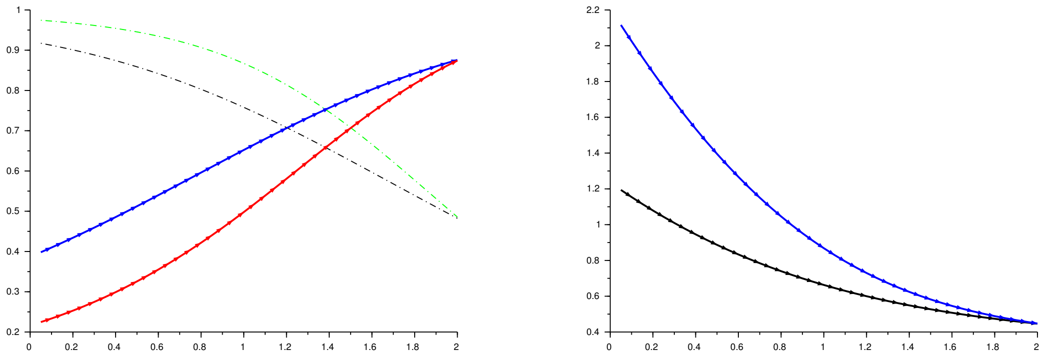

Figure 4.11: Left figure. Plot of the functions ρ (in blu $\epsilon = 0.025$, in red $\epsilon = 0.05$) and $\bar{\rho}$ (in black $\epsilon = 0.025$, in green $\epsilon = 0.05$). Right figure. Plot of the function m, in black $\epsilon = 0.025$ and in blu $\epsilon = 0.05$.

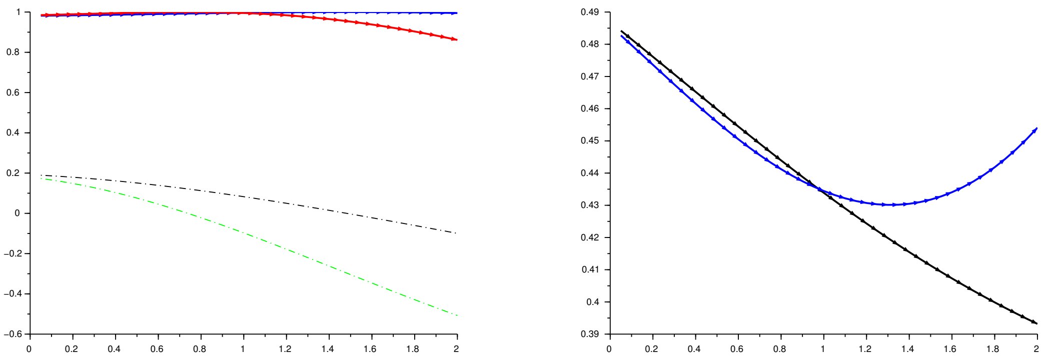

Figure 4.12: Left figure. Plot of the functions ρ (in blu $\epsilon = 0.025$, in red $\epsilon = 0.05$) and $\bar{\rho}$ (in black $\epsilon = 0.1$, in green $\epsilon = 0.05$). Right figure. Plot of the function m, in black $\epsilon = 0.025$ and in blu $\epsilon = 0.25$.

Chapter 5

Some applications

In the preceding chapter we have developed a theoretical and numerical framework to model multi-dimensional stochastic volatility in the interest rates market. In the present chapter we present some applications of the model. We further specify the model in order to fulfill our primary objective i.e. to keep the interpretation of the yield curve factors Y as in the LGM and to have a similar interpretation of the volatility factors X in terms of the implied volatility term structure.

Our approach to calibration is stepwise. We insist on the fact we want to fix the model parameters $\Psi = (K, \theta, c, b, \Omega, \epsilon, \rho)$ globally and calibrate the initial values of the underlying state variables (x, y) to fit observed market data. In particular we aim to calibrate the values of the matrix x to fit the swaption volatility cube (i.e. the cube formed by the implied volatilities of swaptions in the three dimensions: maturity, expiry and strike). Ideally, the dynamics of the interest rates options market should be explained by the dynamics of the factors X. The main idea of our calibration procedure is to exploit the expansion of the volatility obtained in section 4.3 to calibrate the value of x. We formulate the calibration problem as a positive semidefinite programming optimization problem (SDP hereafter) which allows to handle the positive semidefinite constraint very easily. This has previously been done in [d'A03] and [BW00] to calibrate the BGM model. We conduce a calibration exercise over US ATM volatilities between 2006 and 2011. The numerical results indicate that while the model fit of market prices is not always satisfactory compared to the standards used in the industry for pricing purposes, the model is able to globally capture the dynamics of the ATM volatility surface with a limited number of factors[1]. This makes it a good candidate as a benchmark for global hedging of a book of interest rates options.

The chapter is organized as follows, section 5.1 describes the numerical framework used to solve the calibration problem. We show how optimizing the sum of squared differences between the observed swaptions cumulated variances and the expansion to the order 1 is

[1]Baring in mind that the dynamics of the volatility cube in the model is homogeneous, and that the number of factors is quite limited, the calibration results prove to be quite satisfactory and robust across the changes in the market.

equivalent to solving a linear criteria under a positive semidefinite constraint, and which can be solved using SDP. Section 5.2 discusses the design of the specification of the model used for calibration. In particular we describe how the parameters b and c are specified in such a way that the role of the volatility factors X is well identified in terms of the swaptions implied volatilities dynamics. Finally section 5.3 discusses the numerical results of the calibration.

5.1 A semidefinite programming approach for model calibration

We present here the main optimization tool we have used in our calibration procedure. We merely followed the footsteps of [d'A03] and [BW00], and reformulated the calibration problem as a positive SDP problem which can be easily solved. SDP has many applications in different fields of applied mathematics and physics. The SDP programming is a mature research field and there are today very stable and efficient numerical methods to solve SDP problems. See [BV03] for a reference on SDP. Whilst optimization problems under positive semidefinite constraints arise naturally in financial mathematics, there are very few examples of applications of SDP. To the best of our knowledge [d'A03] and [BW00] are the only example of application for interest rates modeling.

5.1.1 The calibration problem

We aim to calibrate the values of the matrix x to fit the swaption volatility cube (i.e. the cube formed by the implied volatilities of swaptions in the three dimensions: maturity, expiry and strike). Typically the calibration problem will take the form

$$x_t = \text{argmin}_{x \in \mathcal{S}_d^+(\mathbb{R})} f_t(x),$$

where the function f_t is the criteria that describes the fitting error of the volatility cube at date t. Alternatively the calibration problem takes the form

$$\Delta x_t = \text{argmin}_{x_{t-1}+\delta x \in \mathcal{S}_d^+(\mathbb{R})} g_t(\delta x),$$

where the function g_t is the criteria that describes the fitting error of the variations of the volatility cube at date t. The constraints $x \in \mathcal{S}_d^+(\mathbb{R})$ and $x_{t-1} + \delta x \in \mathcal{S}_d^+(\mathbb{R})$ are convex constraints. In order to derive an SDP formulation of the calibration problem (or more precisely of a step of the calibration problem) it is important that the function f_t and g_t are linear or quadratic functions of the factors x and δx. Clearly this severely restricts the range of objective functions for our calibration procedure. In particular, to have a linear or quadratic criteria in x, we cannot include out-of-the-money options in the optimization, since these naturally introduce non-linearity. We thus restrict the first step of the calibration problem to the ATM swaptions. We notice that up to the first order the expansions of the ATM implied volatility we have obtained is a linear functions of x. Also, figure 4.6 shows that the contribution of the first order term to the ATM volatility is negligible. The idea of our calibration procedure is to define a criteria using the expansion of the implied volatility. This leads to the forms (5.7) and (5.9) respectively for f_t and g_t.

5.1.2 The SDP formulation

Both our calibration problems fit into the general optimization problem

$$\begin{aligned} \min \sum_{i=1}^{N} \left(\operatorname{Tr}\left(M_i \xi\right) - V_i\right)^2 \\ \xi + \xi_0 \in \mathcal{S}_d^+(\mathbb{R}) \quad . \end{aligned} \tag{5.1}$$

Where M_i are positive semidefinite matrices, and V_i are positive scalars. We will show that this problem can be formulated as an SDP problem. Let us introduce new auxiliary variables $t_1, ..., t_N$, the problem (5.1) is equivalent to the optimization problem

$$\begin{aligned} \min \sum_{i=1}^{N} t_i \\ t_i \geq \left(\operatorname{Tr}\left(M_i \xi\right) - V_i\right)^2 \\ \xi + \xi_0 \in \mathcal{S}_d^+(\mathbb{R}) \quad , \end{aligned}$$

which can also be written as

$$\begin{aligned} \min \sum_{i=1}^{N} t_i \\ \begin{pmatrix} t_i & \left(\operatorname{Tr}\left(M_i \xi\right) - V_i\right) \\ \left(\operatorname{Tr}\left(M_i \xi\right) - V_i\right) & 1 \end{pmatrix} \in \mathcal{S}_d^+(\mathbb{R}) \\ \xi + \xi_0 \in \mathcal{S}_d^+(\mathbb{R}) \quad . \end{aligned}$$

Let us define $z = (t_1, ..., t_N, \xi_{11}, ..., \xi_{1d}, ..., \xi_{dd})$. This problem fit in the general SDP formulation

$$\begin{aligned} \min_{z \in \mathbb{R}^p} c^T z \\ F_0 + \sum_{i=1}^{p} z_i F_i \in \mathcal{S}_d^+(\mathbb{R}) \quad , \end{aligned}$$

where $F_1, ..., F_p$ are symmetrical matrixes. As mentioned before there are several methods to solve the above SDP problem, and some stable and efficient numerical codes available such as CPLEX, MOSEK, SEDUMI...

The SDP framework is flexible and allows to insert more complex constraints. In particular we might be interested in block-diagonal solutions of the matrix ξ. It is straightforward to adapt the problem formulation in order to account for this constraint, we just need to define a vector z which only includes the elements of the matrix ξ which we want to be non-zero.

5.2 Choice of the constant parameters for model calibration

We have insisted on the fact we want to fix the model parameters $\Psi = (K, \theta, c, b, \Omega, \epsilon, \rho)$ globally and calibrate the initial values of the underlying state variables (x, y) to fit observed market data. In particular we aim to calibrate the values of the matrix x to fit the swaption volatility cube (i.e. the cube formed by the implied volatilities of swaptions in the three dimensions: maturity, expiry and strike). The parameters K and b determine respectively the term structure of the yield curve and of the at-the-money volatility surface and are fixed arbitrarily. This is a fairly common practice (see [Pit09]), and is justified by the fact these parameters are time scales that allow to identify the factors role in terms of yield curve and volatility surface dynamics. Once these parameters are fixed we exploit the price expansion we have developed in subsection 4.3.2. For a suitably chosen level of the parameter ϵ, the results of our numerical experiments showed that the expansion provides a good approximation of the at-the-money smile. Furthermore the ATM smile is fairly unsensitive to the choice of the parameter ρ and the instantaneous variance is a linear function of the matrix x. We thus approach the first step of the calibration problem by the problem of choosing x so that the formula (4.43) (which only depends upon the parameters K, b and c) fits the observed at-the-money volatilities. We suggest that the remaining degrees of freedom, namely ϵ and ρ can be set by optimizing a global criteria on the whole volatility cube (adding some relevant strikes).

The mean reversion parameters K and b can be fixed in order to identify the role of the model factors. Our choice of the parameters K is based on the yield curve dynamics in the LGM model as described in section 2.4. Once the parameters k_i are set, and the role of the associated factors is clearly identified, the role of the volatility factors x just follows. The instantaneous variance-covariance of the process Y is given by $cov_t(Y_t) = cX_tc^\top$. The interpretation of the factors X then follows from the interpretation we have given of the factors Y. The instantaneous variance of the factor Y_i is given by $\sum_{k,l=1}^d c_{ik}c_{il}X_{kl}$ and the instantaneous covariance between the factors Y_i and Y_j is given by $\sum_{k,l=1}^d c_{ik}c_{jl}X_{kl}$. Let us note that the model thus exhibits a stochastic correlation between the main movements of the curve and thus between rates. Neglecting the diffusion part of the process X and solving the o.d.e. of the drift term we have

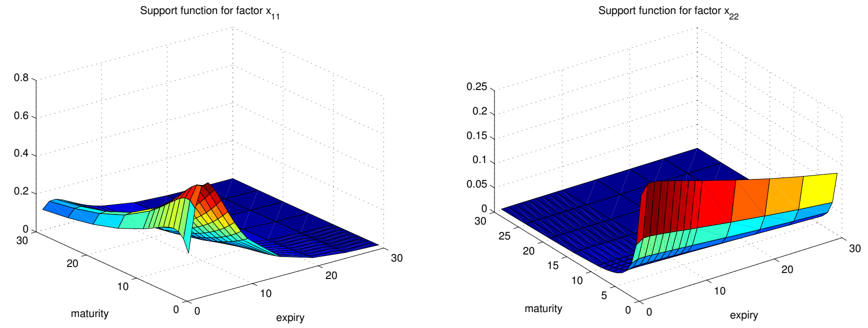

Figure 5.1: Plot of the support function of the order 0 implied cumulated variance (4.43) for the factors X_{11} and X_{22} in a model with 2 curve factors Y and a 2×2 volatility matrix X, with $k_1 = 0.1, k_2 = 1$, $b = \text{diag}(-0.41, -0.01)$ and the parameter c given by 5.3.

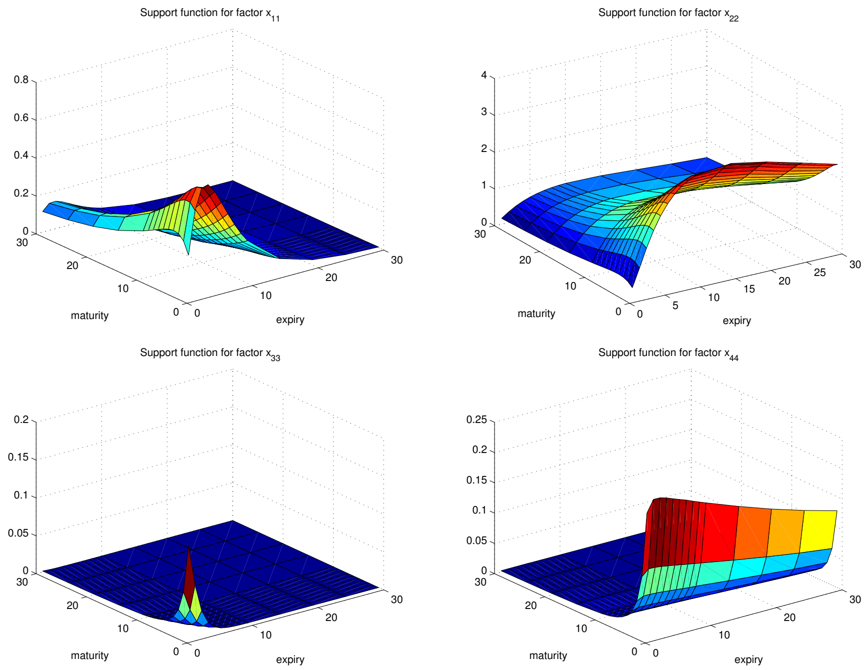

Figure 5.2: Plot of the support function of the order 0 implied cumulated variance (4.43) for the factors X_{11} and X_{22} in a model with 2 curve factors Y and a 4×4 volatility matrix X, with $k_1 = 0.1, k_2 = 1$, $b = \text{diag}(-0.41, -0.01, -0.41, -0.01)$ and the parameter c given by 5.3.

$$X(s) = e^{b(s-t)} \left(x + \int_t^s du e^{-b(u-t)} \Omega e^{-b^\top (u-t)} \right) e^{b^\top (s-t)}. \tag{5.2}$$

The first order term in the expansion of the volatility (4.43) is the implied volatility in the LGM with a deterministic covariance matrix given by (5.2). The parameter b appears as a time scale for the volatility and covariance of rates. Following our interpretation of the factors Y, a given tenor of the yield curve is identified with a particular factor Y_i, either with a linear combination of these factors. We want to have a term structure of implied volatility for the tenors of the curve, and possibly a clear interpretation of the factors X_{ij} in terms of the volatility cube movements. The dimension of X will typically be equal or higher than the dimension of Y, i.e. $d \geq p$. We will typically consider $d = mp$, where m depends on the degree of flexibility we need to model the volatility term structure. Let us illustrate how the parameter c can be chosen to achieve this: consider the simple model with $d = p = 2$. Setting the mean reversion parameters K to be diagonal and with values $k_1 = 1\%$ and $k_2 = 1$ we know that in the LGM the factor Y_1 drives the long term rates of the yield curve and the factor Y_2 drives the slope of the yield curve. Choosing c as follow:

$$c = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \tag{5.3}$$

we have

$$d < Y^1 >= X_{11} \quad \text{and} \quad d < Y^1 + Y^2 >= X_{22}.$$

Then following the identification of the factors Y, we can also identify the factors X as the main drivers of the volatility curve. Precisely, the factor X_{11} will drive the volatility of long maturities swap rates and the factor X_{22} will drive the volatility of short maturity swap rates. The graphics in Figure 5.1 represent the support function of the ATM volatility in a model with 2 curve factors Y and a 4×4 volatility matrix X. The figures confirm the intuition, the support function of the factor X_{22} is almost zero on long for long maturities swaps and has a high value for short maturity swaps. As a consequence, the factor X_{22} will drive the short maturity swaptions. Furthermore the correlation between Y_1 and $Y_1 + Y_2$ is given by X_{12}. As a consequence, the model exhibits a stochastic correlation between rates, and the non-diagonal factor X_{ij} will drive the prices of products which depend on this correlation such as CMS spread options.

When $d = mp$ we will typically choose c to be a block-diagonal matrix. For example for $d = 4$ and $p = 2$, we might consider the matrix c given by

$$c = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 1 \end{pmatrix}. \tag{5.4}$$

The matrix b is set to be a diagonal matrix. The diagonal elements are set to different time scales so that the factors $(X_{im+k,im+k})_{i=1...p,k=1...m}$ are associated to different maturities of the volatility. In particular, for a given swap maturity, we expect to have a factor which drives the long expiries implied volatility and a factor which drives the short expiry volatility. The graphics in Figure 5.2 represent the support function of the ATM volatility in a model with 2 curve factors Y and a 4×4 volatility matrix X. The graphics confirm that the parameter b can be used as a time scale on the expiry dimension of swaptions. The factors associated to the parameters $b(i,i)$ with higher value drive the short expiries volatilities, while the factors associated to parameters $b(i,i)$ with smaller value drive the long term expiries volatilities.

5.3 Calibration results

As a first step of the calibration procedure we will calibrate the initial value of the volatility factors x to the at-the-money volatility surface. We have shown that for ϵ small enough, the at-the-money volatility smile is very close to the order 0 term of the expansion (4.22) and (4.43). Let us note that these expressions are linear forms of the volatility factors x. We denote by $V_t^C(T,\delta)$ the date t market quote of the at-the-money implied cumulated variance of the Caplet of maturity T and tenor δ, and by $V_t^S(T,m)$ the date t market quote of the at-the-money implied cumulated variance of the swaption of expiry T and maturity m. We rewrite (4.22) and (4.43) as follows

$$\begin{aligned} v(t,T,\delta) &= \mathrm{Tr}\left(M_t^C(T,\delta)x_t\right) + \mathrm{Tr}\left(\bar{M}_t^C(T,\delta)\Omega\right) && (5.5)\\ \bar{v}(t,T,m) &= \mathrm{Tr}\left(M_t^S(T,m)x_t\right) + \mathrm{Tr}\left(\bar{M}_t^S(T,m)\Omega\right). && (5.6)\end{aligned}$$

The calibration problem is formulated as a minimization problem of a quadratic criteria. Given a set of maturities and tenors $(T_1,\delta_1),...,(T_{N_1},\delta_{N_1}),(T_1,m_1),...,(T_{N_2},m_{N_2})$ to which we want to fit our model at date t. Once again let us repeat that our model aims to capture the dynamics of both the yield curve and the volatility cube through the underlying state variable (X,Y). In particular the implied volatility cube dynamics is meant to be captured by the factors X. One of the main messages of Black-Scholes is that the only thing that matters when hedging a derivative contract is to be able to dynamically adjust our portfolio in response to the movements of the underlying asset. As a consequence it is important that the model is able to capture both the level of the volatility cube and its variations.

Though the analogy with the one-dimensional setting is not completely accurate, let us note that the parameter Ω can be interpreted as a parameter accounting for the long term mean of the stochastic variance-covariance process X. This parameter is meant to be characteristic of the option market and should not vary on a daily basis. The idea is to calibrate the values of x and Ω at a given date and then calibrate the variations of x in order to fit the variations of the implied volatility cube.

5.3.1 Calibrating x and Ω

Let us now describe the optimization problem we will solve in order to determine the values of x and Ω that fit the ATM volatility surface at a given date. Define f_t by

$$f_t(x,\Omega) = \sum_{i=1}^{N_1}\left(\mathrm{Tr}\left(M_t^C(T_i,\delta_i)x + \bar{M}_t^C(T,\delta)\Omega\right) - V_t^C(T_i,\delta_i)\right)^2$$

$$+\sum_{j=1}^{N_2}\left(\mathrm{Tr}\left(M_t^S(T_j,m_j\tau_j)x+\bar{M}_t^S(T_j,m_j\tau_j)\Omega\right)-V_t^S(T_j,m_j\delta_j)\right)^2. \quad (5.7)$$

The calibrated values (x_t,Ω_t) are defined as the solutions of the following optimization problem

$$(x_t,\Omega_t)=\underset{x\in\mathcal{S}_d^+(\mathbb{R}),\Omega\in\mathcal{S}_d^+(\mathbb{R})}{\mathrm{argmin}} f_t(x,\Omega). \quad (5.8)$$

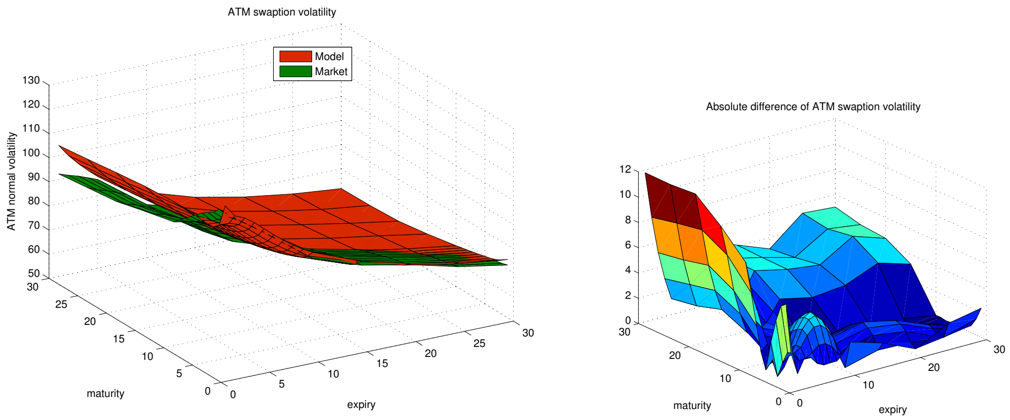

Figure 5.3: Plot of the swaptions' ATM normal volatility surface fitted model vs market at date 100 and absolute difference. We have used a model with 2 curve factors Y and a 4×4 volatility matrix X, with $k_1=0.01, k_2=1$, $b=\mathrm{diag}(-0.41,-0.01,-0.41,-0.01)$ and the parameter c is given by (5.4). Volatility and difference are in bp.

We have performed several calibration experiences for different dates. Overall, the model produces a good fit of the ATM volatility surface. We do not expect our model to fit the market prices perfectly, our calibration is global and we shall always keep in mind that the parameters b, c and K which determine the matrix functions $M^C(T,\delta)$ and $M^S(T,m\tau)$ are fixed and are not meant to change for different dates. Typically we will try to fit as much as 36 at-the-money volatilities (corresponding to the maturities 1 2 5 10 20 30 years and the expiries 3 months, 1 2 5 10 and 20 years) of the surface with 20 volatility factors, i.e. 4×4 matrix x_t and Ω. Note that once Ω is calibrated we consider it should be kept as fixed until we observe some major shifts in the market that impose to recalibrate this parameter.

The figures 5.3, 5.4 and 5.5 show the calibration results for three different dates. The absolute difference in the ATM normal volatility between the fitted and the observed implied normal volatility is lower than 5 bp on average except for very short and very long expiries for which the difference can be as much as 20 bp. Let us recall that as showed by the PCA analysis of section 1.5.1 the very short expiry swaption market tends to behave independently of the rest of the swaption market. In factorial term structure models the

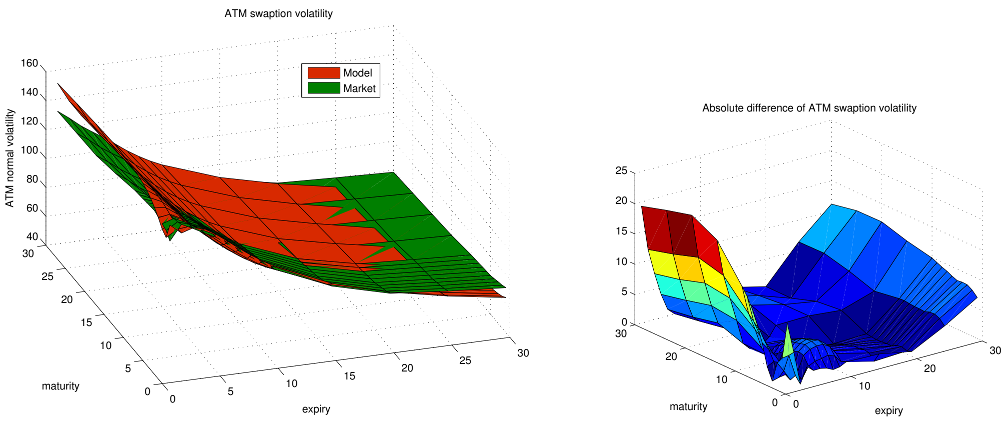

Figure 5.4: Plot of the swaptions' ATM normal volatility surface fitted model vs market at date 200 and absolute difference. We have used a model with 2 curve factors Y and a 4×4 volatility matrix X, with $k_1 = 0.01, k_2 = 1$, $b = \text{diag}(-0.41, -0.01, -0.41, -0.01)$ and the parameter c is given by (5.4). Volatility and difference are in bp.

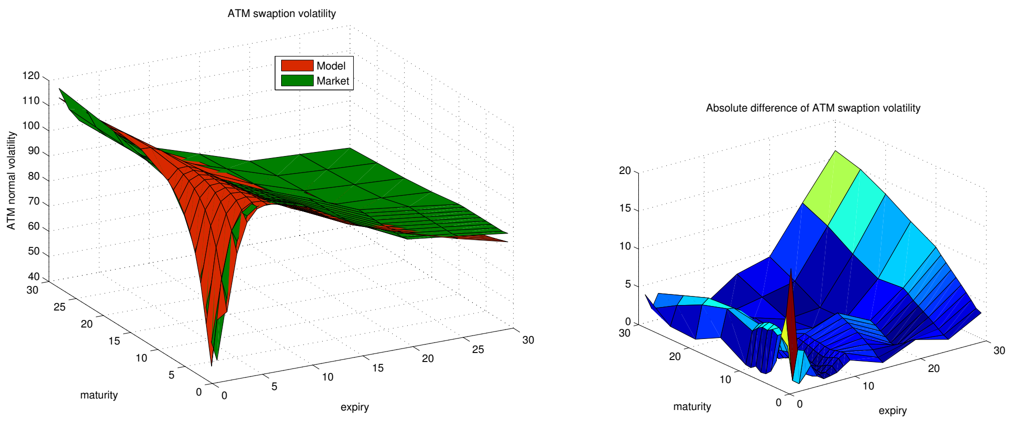

Figure 5.5: Plot of the swaptions' ATM normal volatility surface fitted model vs market at date 300 and absolute difference. We have used a model with 2 curve factors Y and a 4×4 volatility matrix X, with $k_1 = 0.01, k_2 = 1$, $b = \text{diag}(-0.41, -0.01, -0.41, -0.01)$ and the parameter c is given by (5.4). Volatility and difference are in bp.

yield curve and the volatility structure are constrained by no arbitrage conditions and this constraints the forms of the volatility surface that can be attained by the model. The shape of the swaptions ATM normal volatility surface has significantly changed in the period going from 2006 to 2011. In the first period of our dataset i.e. from 2006 to 2007 (see Figure 5.3) the market exhibits a very smooth volatility surface, the model fit to the market is good, with a difference smaller than 10bp across all expiries and maturities. The model is able to accurately fit the expiry term structure of volatility. The model fails to fit the maturity term structure of volatility, it tends to overestimate the short expiry volatility. In the mid period of our dataset i.e. 2008 to 2009 (see Figure 5.3) , which is the period of the credit crunch, the slope of the expiry term structure has increased and the shape of the short expiry volatilities has changed. The model fails to fit both the long and short expiries volatilities. In the last period 2010 and 2011 (see Figure 5.5), the form of the volatility surface has significantly changed: the short expiry volatility has a very pronounced humped shape as a function of the underlying swap maturity. The model is able to accurately reproduce the short expiries volatilities, and tends to overestimate the long term expiry volatilities.

5.3.2 Calibrating Δx

We now describe the optimization problem we solve to calibrate the values of the factors x that better fit the variations of the ATM volatility cube. Assuming that we have calibrated the model and in particular the values of x_t up to the date t_{k-1}, we calibrate the variation $\Delta x_k = x_{t_k} - x_{t_{k-1}}$ that better fit he variations of the ATM volatility surface between the dates t_{k-1} and t_k. We define the objective function between the dates t and s as follows

$$g_{t,s}(\Delta x) = \sum_{i=1}^{N_1} \left(\mathrm{Tr} \left(M_s^C(T_i,\delta_i)(x_t + \Delta x) - M_t^C(T_i,\delta_i)x_t + (\bar{M}_s^C(T,\delta) - \bar{M}_t^C(T,\delta))\Omega \right) - \Delta V^C(T_i,\delta_i) \right)^2$$

$$+ \sum_{j=1}^{N_2} \left(\mathrm{Tr} \left(M_s^S(T_j,m_j)(x_t + \Delta x) - M_t^S(T_j,m_j)x_t + (\bar{M}_s^S(T,m) - \bar{M}_t^S(T,m))\Omega \right) - \Delta V^S(T_j,m_j) \right)^2 . \tag{5.9}$$

The calibrated value Δx_k is defined as the solution of the following optimization problem

$$\Delta x_k = \underset{\Delta x + x_{t_k} \in \mathcal{S}_d^+(\mathbb{R})}{\mathrm{argmin}} \; g_{t_{k-1},t_k}(\Delta x). \tag{5.10}$$

Figure 5.6 shows the quality of the fit of the weekly variations of the ATM implied volatility surface. Clearly the fit is less satisfactory than the fit of the level of the ATM volatility surface obtained when calibrating the factors x and the parameter Ω. While the observed shapes of the ATM swaption volatility surface seem attainable by the model, its observed variations are not regular and they don't seem to be attainable by the model. As expected by iterating the calibration of the variations we accumulate the errors and the fit of the level of the implied volatility surface is not comparable to the one we obtain by calibrating both the factor x and the parameter Ω. Figure 5.7 shows the fit of the ATM volatility surface of swaptions at date 210 where we have calibrated the factor x and the parameter Ω by solving the optimization problem (5.8) at date 200 and the solving the optimization problem (5.10) for 10 successive weeks.

Figure 5.8 shows the market volatilities against the model volatilities when calibrating the weekly variations of the swaptions ATM volatility surfaces. As the calibration focuses on

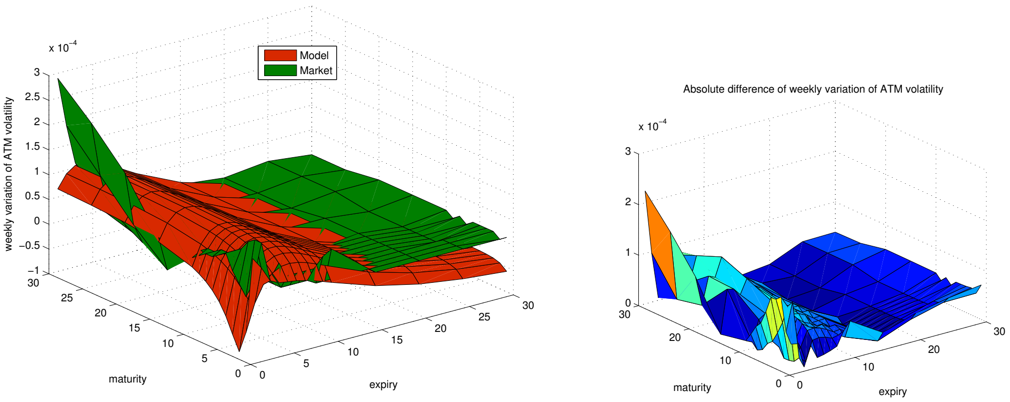

Figure 5.6: Plot of the weekly variations of the ATM volatility surface fitted model vs market between 209 and 210 and of the absolute difference. We have used a model with 2 curve factors Y and a $4x4$ volatility matrix X, with $k_1 = 0.01, k_2 = 1$, $b = \text{diag}(-0.41, -0.01, -0.41, -0.01)$ and the parameter c is given by (5.4).

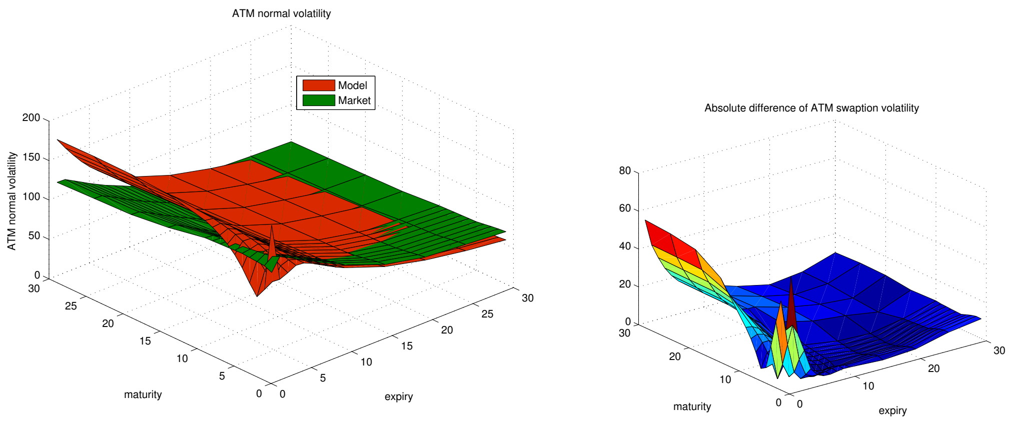

Figure 5.7: Plot of the ATM volatility surface fitted model vs market at date 210 and of the absolute difference. We have calibrated x and Ω at date 200 and calibrated the variations for 10 steps (i.e. 10 successive weeks). We have used a model with 2 curve factors Y and a $4x4$ volatility matrix X, with $k_1 = 0.01, k_2 = 1$, $b = \text{diag}(-0.41, -0.01, -0.41, -0.01)$ and the parameter c is given by (5.4).

the weekly variations of the instantaneous cumulated variance, it fails to calibrate the level of the swaption volatility. The calibration problem (5.10), focuses on optimizing the fit of the weekly variations of the ATM volatility surface rather than the level of the volatility itself. In terms of hedging the ability of the model to capture the market variations of the underlyings is as important as its ability to fit the observed market prices at a given date. Figure 5.9 shows the weekly variations of the volatilities, market against model. Whilst the model fails to calibrate the level of the volatilities, it correctly captures the weekly variations

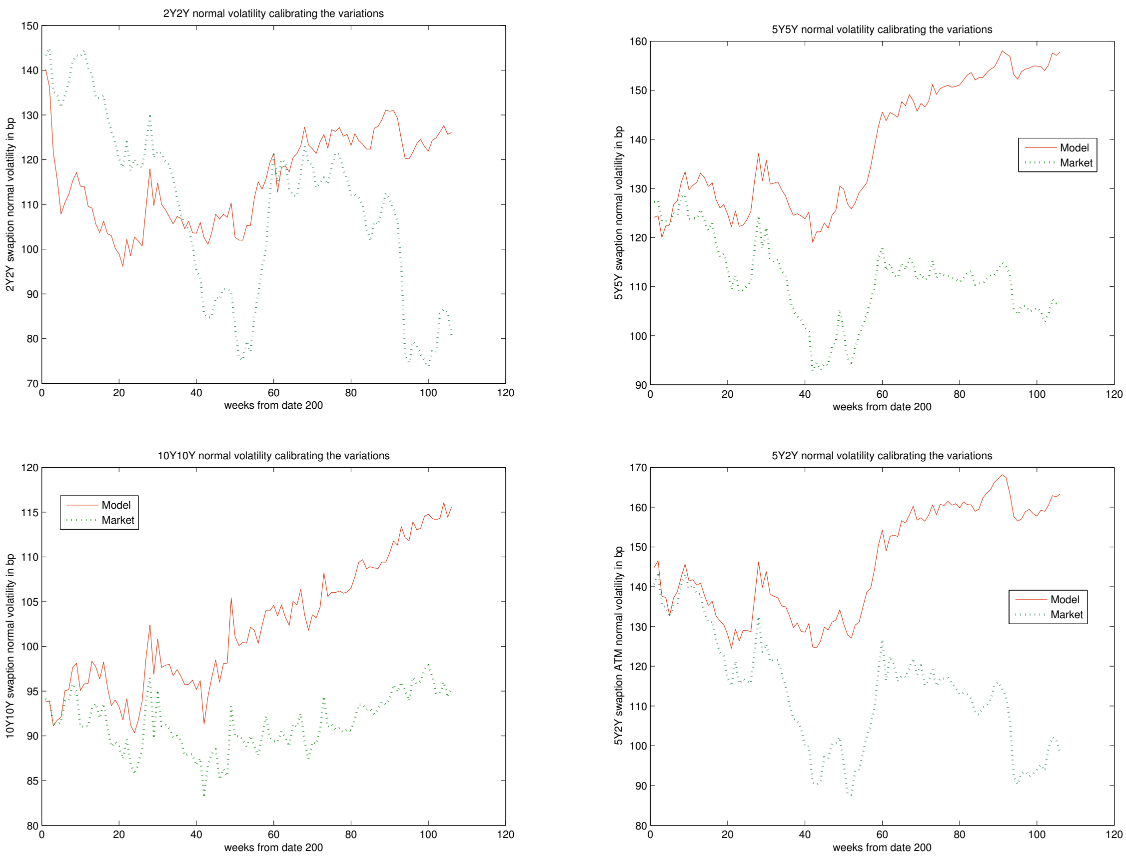

Figure 5.8: Plot of the historic swaption ATM normal volatility fitted model vs market. We have calibrated x and Ω at date 200 and calibrated the weekly variations of the volatility surface on successive weeks. We have used a model with 2 curve factors Y and a $4x4$ volatility matrix X, with $k_1 = 0.01, k_2 = 1$, $b = \text{diag}(-0.41, -0.01, -0.41, -0.01)$ and the parameter c is given by (5.4).

of the volatility surface. The table below gives the residuals of the R squared of the linear regression of the weekly variations of the volatilities against the volatility estimated by the model.

Swaption	2Y2Y	5Y5Y	10Y10Y	5Y2Y
R^2	0.488	0.901	0.914	0.79

To measure the contribution of the factors in explaining the variations of the volatility cube, we have regressed the weekly variations of the volatility directly on the calibrated factors, rather than the volatility estimated by the model. The regression of the weekly variations of the 10Y10Y swaption implied variance on the 3 factors which compose the 2x2 upper left submatrix of x gives an adjusted R squared of 0.92. Conversely the regression on the

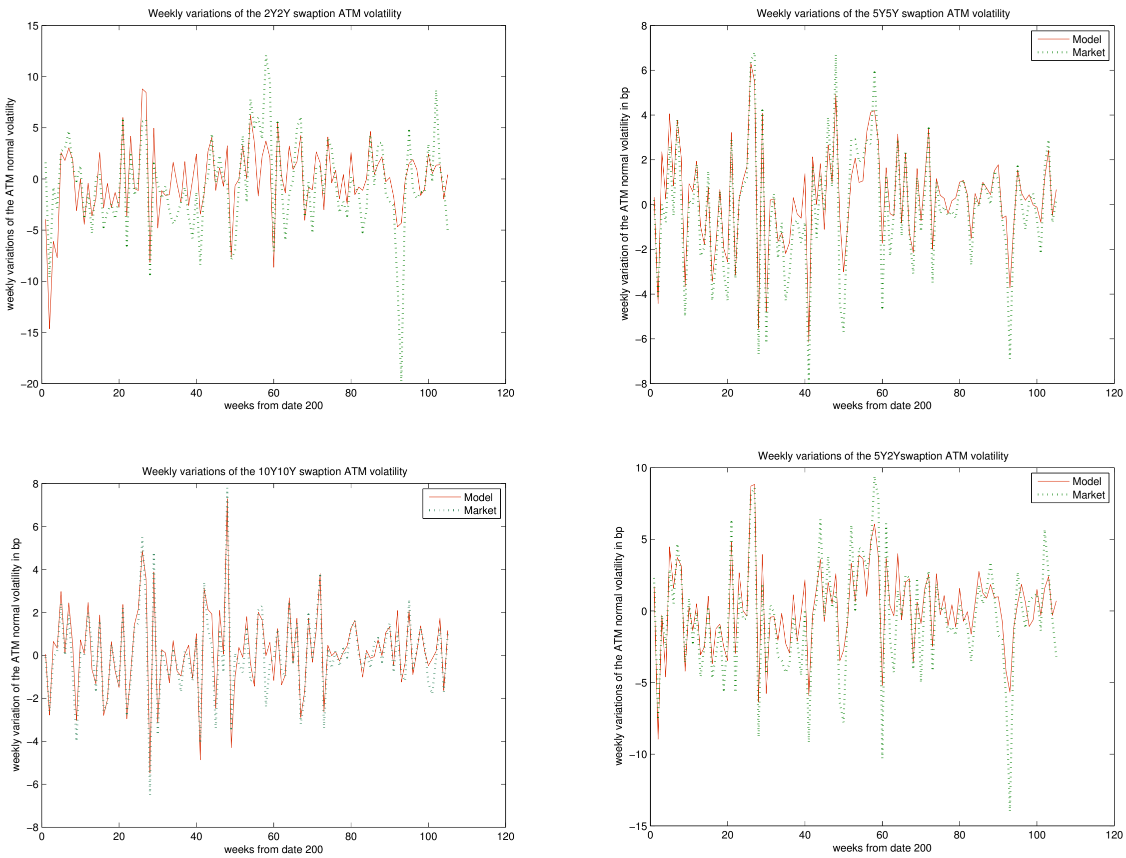

Figure 5.9: Plot of the historic variation of the swaption ATM normal volatility fitted model vs market. We have calibrated x and Ω at date 200 and calibrated the weekly variations of the volatility surface on successive weeks. We have used a model with 2 curve factors Y and a $4x4$ volatility matrix X, with $k_1 = 0.01, k_2 = 1$, $b = \mathrm{diag}(-0.41, -0.01, -0.41, -0.01)$ and the parameter c is given by (5.4).

3 factors which compose the 2x2 down right submatrix of x gives an adjusted R squared of 0.24. This is consistent with the interpretation of the factors x which are supposed to drive the long and short expiry of the long and short maturity swaps. Of course this mapping is not perfect, however it allows us to think in terms of principal component movements of the swaption volatility cube. In terms of hedging, this suggest how to map the "deltas" (i.e. the derivatives) w.r.t. x into the deltas of the products to swaptions.

Apart for short expiries, for which we have seen thorough the model fails to calibrate the market, the regression results show a high level of explanation of the weekly variations of the volatility surface by the calibrated model. Figure 5.10 gives the plot of the residuals against the normal distribution. The straight line represents the perfect adequacy of the distribution of the residuals to a centered normal distribution. The body of the distribution is in

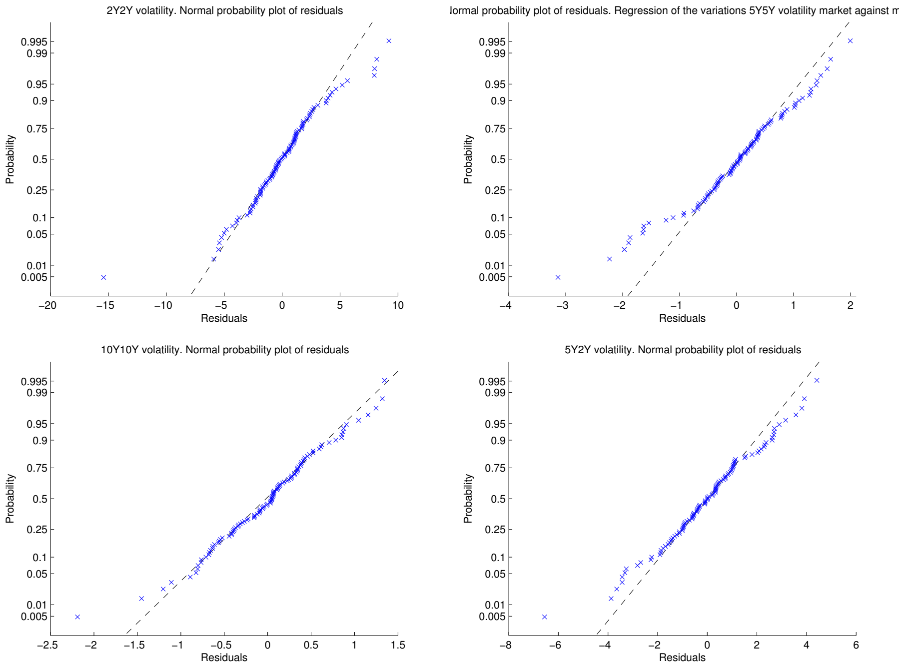

Figure 5.10: Plot of the residuals of the regression of the weekly variations of the swaption ATM normal volatility market vs fitted model.

line with the normal distribution, while the tail of the distribution of the residuals departs from the normal distribution. This indicates that the linear representation of the volatility we have tested in our first calibration step is not sufficient to explain the dynamics of the ATM volatility surface. An interesting question, which we leave to future investigation, is wether we can improve the explanatory power of the model by a better choice of the constant parameters b and c[2], or by adding the higher order terms of the expansion (and in particular the second order term) in the calibration problem[3].

Let us insist on the fact that, whilst the fit of the market data can be consider insufficient (when compared to parametric representation of the volatility such as SABR, which is the market standard to manage interest rates swaptions books), our model is a fully-homogeneous global model, which is able to capture the whole ATM volatility surface through a very limited number of factors. Most importantly, by choosing an adequate specification of the model, we managed to clearly identify the role of the factors. The role of the factors Y has been discussed in section 2.4 and is summarized by Figure 2.2. The main contribution of our model is to achieve a similar identification for the factors X in terms of implied volatility dynamics.

Remark 41 — ***Constrained calibration and robust hedging of exotic products:***

[2]However this approach will mine our attempt to have a clear interpretation of the model factors.
[3]In which case we cannot formulate the calibration problem as an SDP optimization problem.

We have described a robust numerical framework for model calibration. In particular we have showed how the model can be specified in such a way that the volatility factors X are clearly identified as main drivers of the ATM swaptions' volatility surface. As mentioned in section 5.1, the semidefinite programming optimization framework is flexible, and allows to introduce complementary constraints to the calibration problem. In particular, we could impose the constraint that X is a block diagonal matrix. Given a calibrated block-diagonal matrix X, we have additional degrees of freedom: the correlations between the factors that drive the off-diagonal terms outside the blocks. These parameters are the main drivers of exotic products depending on the correlation between rates. As a consequence the model allows to build boundary prices for these trades, inside which the models produces a robust hedging strategy for these products.

Part III

A quantitative view on ALM

This part of the dissertation is meant to be introductory to a field of finance on which I came across in my working experience at Crédit Agricole S.A.. On several occasions I have worked for the Financial division of the bank on quantitative subjects related to interest rate and inflation risk in asset and liability management (ALM).

The theory of hedging of interest rates risks in ALM has mainly been developed in France. The seminal book of Frachot, Roncalli et al. [DFR03] and the book of Adam [Ada08] describe the key notions to manage interest rate risk in ALM. These theories have been development to solve the concrete problems which are faced by the financial division of French retail banks: hedging fixed rate saving accounts. In most countries the retail banking saving accounts pay a floating interest which is indexed either on market rates, either on some bond or equity underlying. In France, some of the saving accounts pay a fixed interest. This creates an interest rate risk for the bank which needs to finance its activity at market rates, and is thus exposed to the fluctuations of such rates. This exposure is called the interest rate risk and one of the roles of the Financial division of the bank is to manage such risk.

The purpose of this part of the dissertation is twofold. On one hand we want to show how derivative contracts such as IRS, and IR options appear naturally in a business which is very close to the primary economic needs of society (retail banking). On the other hand we want to communicate on a field which is little known from the mathematical finance community and in our modest opinion presents some interesting quantitative challenges. It is worth noting that this activity represents huge amounts of capital, trading books linked to the ALM activity are usually much larger than a normal trading book of an average investment bank.

This second part of the dissertation is organized as follows. In chapter 6 we briefly introduced the key notions used in ALM. We recall the definitions of schedule and interest rate gap. We refer to [DFR03] and [Ada08] for an exhaustive and detailed description of these concepts. We also introduce a new concept, the concept of envelope which generalizes the notion of schedule. To the best of our knowledge this concept has not been introduced in the literature so far, or at least not in such an explicit way. This concept is very useful in practice to account for the uncertainty on the future evolution of assets and liabilities. Also, it is a natural way to introduce optionality in the assets and liabilities.In chapter 7, we present some two concrete examples of French state regulated saving accounts. For each example we propose a model to hedge the interest rate risk related to this particular product. Section 7.1 describes the hedging problems linked to the Livret A saving account. This account is indexed on inflation and on nominal interest rates. We propose three different modeling approaches to hedge the interest rate risk related to this product. We show that depending on the assumptions on the behavior of the clients, and on the hedging strategy, the perception of the risk is significantly different. In section 7.2 we study a model for hedging the interest rate risk linked to Plan d'Epargne Logement (PEL) saving accounts. The idea of the model was initially proposed in Roncalli, Demey at al. in [BDJ+00]. We have extended their model by introducing a behavioral dimension in the model which allows to consider different behavior of the clients. Finally in chapter 8 we study the hedging of interest rates options in ALM. We discuss the peculiarities of the activity, and in particular the differences with the derivatives business. We propose two alternative hedging strategies, which correspond to different choices of numeraire. We discuss the pros and cons of both strategies for hedging options in ALM, and compare them in terms of hedging performance in a toy model.

Chapter 6

Notions of ALM

We will focus exclusively on the notion of interest rate risk in ALM. In particular we will cover the quantitative modeling techniques that can be used to manage such risks. This chapter gives some introductory notions to ALM. Liabilities, as well as assets are not necessarily known in advance, their evolution depends on several factors such as the behavior of creditors or savers, the evolution of market rates. The objective of the Financial division is to lock the future margin no matter the evolution of both assets and liabilities of the bank.

The key concept in ALM is what is commonly known as the interest rate gap. Interest rate risk comes from a mismatching between the rate to which the bank is able to fund its liabilities and the rate at which this liabilities are effectively remunerated. The philosophy of ALM consists in transforming any fixed rate (either on the liability side, or on the assets side) to a floating rate which is meant to be as close as possible to the rate at which the bank is able to fund its activity. This is called the transformation business.

6.1 Assets and liabilities

For a retail bank the liabilities are mainly the bank saving accounts and bond issues. The assets are mainly the credits. It is worth mentioning that France has some of the more complex saving accounts contracts and these are state regulated. Some of them are true exotic derivative contract, as the Livret A and the Plan d'Epargne Logement (PEL). Unfortunately for the bank, neither the notional of the liabilities, neither the notional of the assets are known in advance. The first duty of the financial division is to build a model that enables to anticipate the future notional of assets and liabilities. Once a model has been build, the second fundamental duty of the financial division is to hedge the positions of the bank.

6.1.1 Some fundamental differences with a Mark to market approach

It is important to understand that there are some fundamental differences between the activity of a trading desk of an investment banking division and the Financial division desks that manage the exposure of a large retail bank.

The role of a classic investment banking trading desk is to sell derivative contracts, receive the premium of the contracts up front and build an hedging strategy to globally hedge the exposure of the desk. The performance of the desk, i.e. the P&L, is measured on a daily basis by marking to market the instruments in the desk's book[1]. The future (real) outcomes of the derivatives, and more precisely of the netted position (derivative minus hedges, i.e. the so called tracking error) of the desk, are represented by the mark-to-market of the book.

Conversely, the performance of the Financial department is measured in accrual, i.e. it is the result of the actual cash flows of the assets and liabilities of the bank. For example, a credit paying an interest K to the bank is, from an ALM perspective, not valued as the net present value (NPV) of the future cash flows of the contract, but it is valued as an interest rate of K on the notional of the credit for the period of the credit. The performance of such a contract will be determined as the difference between the rate K and the rate at which the bank is able to fund this credit. It is important to note that, in representing assets and liabilities through the NPV or mark to market (MtM), the information on the actual underlying notional is lost. Conversely, in the ALM view the relevant information is made of the couple notional and rate we pay or receive on that notional.

6.2 Modeling schedules

We introduce the notion of schedule which is extensively used in [DFR03] and [Ada08].

Definition 42 — *The schedule of an asset or a liability is a model that gives the future evolution of its notional in the future as a function of time as if the bank stopped its activity today. A schedule is associated to a interest rate received or payed depending if it is an asset or a liability.*

While the notion of schedule is a market standard in ALM management, we consider it is incomplete in describing the evolution of assets and liabilities. It is extremely ambitious to think that we are able to model the future evolution of the notional of assets and liabilities of a bank in a deterministic way, since it depends on a number of factors such as the behavior of clients (savers and creditors), the evolution of economic factors (rates, inflation, equity).

[1] In practice this is not exactly the case, since for some products we cannot always observe the price in the market. We thus price these products with a model, we say that these products are marked to model.

Modeling the future evolution of the notional as if the bank stopped its activity today is a questionable assumption. In theory, assuming that the model is correct, the model enables the bank to hedge the stratums of assets and liabilities successively. However, in practice, it is very important to consider that the bank has an ongoing activity, and to pay attention at the actual evolution of the the total notional of its assets and liabilities.

It is possible to extend the notion of schedule introducing the notion of envelope. As mentioned before, the future evolution of assets and liabilities depends on several factors, in particular, it is reasonable to assume that it depends on a number of market factors such as interest rates, inflation... An envelope is a function of time and of a set of underlying state variables.

Definition 43 — *The envelope of an asset or a liability is a model that gives the future evolution of its notional as a function of time and of a set of state variables as if the bank stopped its activity today.*

If the function is linear, the envelope can be decomposed in a series of schedules associated to each rate involved in the expression of the envelope.

Let us illustrate the notion of schedule and envelope via an example. Let us consider a simple example of asset, a pool of mortgages with fixed interest rate K. As simple as this example can look, depending on the assumptions, the schedule and envelope can be very different. For simplicity, let us assume that all these mortgages have the same maturity T. Let us denote by N the total notional of these mortgages issued by the bank for a given month.

- If we assume that the clients only pay interests during the life of the mortgage and refund the total notional of the credit at maturity, and assuming furthermore that the mortgages do not allow any prepayment, the schedule is simply a constant from the inception date t until the date of maturity T. In this case the envelope is identical to the schedule.

- If we assume that the clients pay a constant annuity up to the maturity of the mortgage and assuming furthermore that the mortgages do not allow any prepayment. The schedule is a concave function starting at N at date t and finishing at 0 at maturity T. In this case the envelope is identical to the schedule.

- If we assume that the clients pay a constant annuity up to the maturity of the mortgage. If we furthermore assume that the clients can make prepayments of their mortgage at any time before maturity. Then the future evolution of the notional of the mortgage is not deterministic anymore and will depend on a model for the clients prepayments. The simplest possible model would be to consider that the clients will massively prepay their mortgage if they are able to renegotiate their credit at a better rate. We can thus assume that a portion p_t of the notional is refunded if the rate K is higher than the

3 months euribor rate[2] plus a spread s at date t. The envelope is thus given by this function $f(t, r_t) = N(1 - t/T) - p_t 1_{r_t+s<K}$, where r_t denotes the 3 months euribor rate. In figure 6.1 the envelope is the region between the red curve and the green curve.

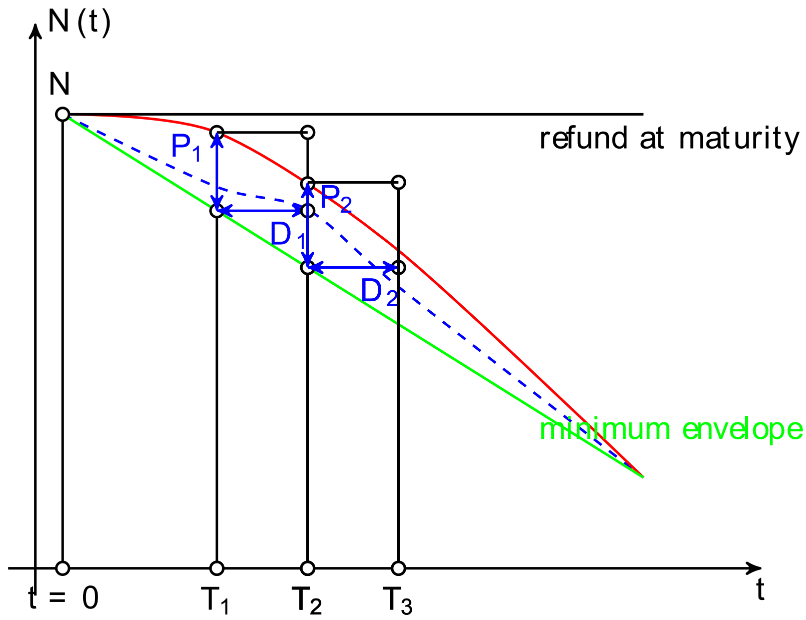

Figure 6.1: Example of schedule for a pool of mortgages. The constant black line is the schedule assuming constant interest, without prepayment and refunding of the full notional at maturity. The concave red curve is the schedule assuming constant annuities and without prepayment. The green straight line represents the minimum of the envelope, i.e. the estimated minimum notional that will remain in the mortgage pool. The envelope is the region between the red and the green curves and it can be approximated by a series of swaptions or caplets. In the figure we have represented two consecutive swaptions/caplets of expiry T_1 and T_2, maturity D_1 and D_2, and of notional N_1 and N_2. The dashed line represents the realized schedule.

The concept of schedule is the market standard in ALM (at least in France), and the concept of envelope is widely used by practitioners. Using these concepts we can represent the asset or the liability into an observable quantity which can be viewed as a proper derivative contract and thus hedged using the tools we have developed in the first part of this thesis. It precisely quantifies the amount of notional that is due or owned at a given future date. Modeling the future evolution of assets and liabilities using schedules is equivalent to assume that the notional evolves as linear functions of the reference rates, so that the schedule tells exactly the amount that should be invested in each one of these rates to hedge the position. The notion of envelope allows to include non-linear functions of the rates. In

[2] Where we have assumed that the 3 months euribor is the reference rate on which banks determine their mortgage rate.

the language of derivatives, the natural way to represent non-linearity is through options. Assets and liabilities naturally include optionality. The option can be either embedded (as it is the case for prepayments), or explicit (as it is the case for saving accounts such as Livret A, PEL and for callable bonds). In the previous example the envelope can be approximated as the combination of schedule 2 and of a set of Floorlets of strike $K - s$. Each Floorlet has notional p_t. In theory one would need a continuum of Floorlet options in order to replicate exactly the envelope. In practice, we consider a discrete set of options.

If the realized evolution of the asset or the liability stays inside the envelope, then the strategy given by the model will hedge the interest rate risk. Conversely, if the realized evolution of the notional of the asset or the liability is outside of the envelope , then the interest rate risk is not fully hedged.

6.3 Interest rate gap

Given a model for the schedule of a given liability or asset it is necessary to implement the hedging strategy for this asset or liability. The notion of interest gap is the key notion to represent risk in ALM. Consider the activity of a large retail bank, the bank faces a number of liabilities on which it pays different interest rates and has a number of assets which yield different interest rates. The margin of the bank is the difference between the interest rates payed on the liabilities and the interest rates received on the assets. Neglecting the pure commercial margin, the rest of the margin is the result of the work of the ALM division. The margin is realized by optimizing the difference between the interest rates payed on the liabilities and the interest rates received on the assets.

Definition 44 — *The interest rate gap is the amount of outstanding notional of either an asset or a liability which is not hedged and either yields or costs a fixed or floating rate different from the rate of funding of the bank for the same maturity.*

Whenever the bank has to pay or receives a fixed rate or a rate which is different from the rate at which it is able to fund, it is exposed to the evolution of the spread between these two rates. For this reason the standard practice consists in swapping the assets and the liabilities to the market interest rate swap rate. Doing so, the first part of the margin of an asset or a liability is immediately identified as the spread between the rate payed or received by the clients and the swap rate. Entering in an IRS which mirrors the cash flows of either an asset or a liability, implies that the interest payed or received are now indexed on the reference rate of the floating leg of the swap, which is the 3 or 6 months euribor rate. The IBOR is itself different from the actual funding rate of the bank and the second part of the margin is determined as the difference between the IBOR and the actual funding rate of the bank. Typically this rate is close to the OIS, simply put it is fair to say that the interest rate gap of a bank increases as the spread between the IBOR and the OIS rate.

Let us consider the example of the loan to illustrate how the interest gap is hedged. Assume

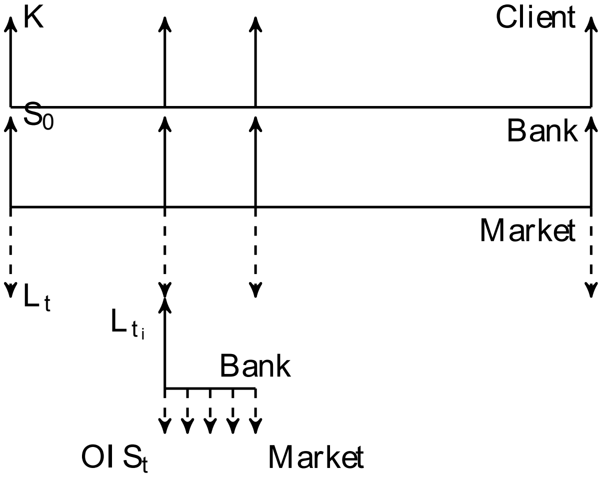

Figure 6.2: Illustration of the mechanism of hedging of the interest rate gap.

that the chosen schedule is 1. Assume that we can enter into a payer IRS of same notional, same payment dates and same maturity. Let us denote by S the par rate for this swap. Then at each coupon payment date T_i, the cash flows are the following:

- The bank receives $KN\delta(T_i - T_{i-1})$ from the creditors.
- The bank pays $SN\delta(T_i - T_{i-1})$ to the counterparty of the swap.
- The bank receives $L_{T_{i-1}}(T_i - T_{i-1})N\delta(T_i - T_{i-1})$ from the counterparty of the swap.

Where $\delta(T_i - T_{i-1})$ denotes the daycount fraction between two successive coupon payment dates. The netted flow at date T_i is given by

$$(K - S)N\delta(T_i - T_{i-1}) - L_{T_{i-1}}(T_i - T_{i-1})N\delta(T_i - T_{i-1}). \tag{6.1}$$

Let us recall that by definition a bank should be able to fund itself in the market at a cost which is exactly equal to the IBOR rate. From an ALM perspective, the position has been perfectly hedged by entering in the above payer swap. There is no uncertainty on the future outcome of the netted position, and the total margin is given by $K - S$. In practice, the funding is not done at the IBOR rate, the treasury desk uses a mix of different funding techniques to acheive the cheapest possible funding rate. For example, it can lend its assets (such as bonds or even ABS in some cases) using repo contracts, or it can swap again once the libor rate has fixed using OIS swaps. In the latter case, the OIS swap will have the length of the coupon payment period $T_i - T_{i-1}$ and the netted flow at date T_i is given by

$$(K-S)N\delta(T_i-T_{i-1})-(L_{T_{i-1}}(T_i-T_{i-1})-OIS_{T_{i-1}}(T_i-T_{i-1}))N\delta(T_i-T_{i-1})-N\int_{T_{i-1}}^{T_i} dsOIS_s. \tag{6.2}$$

Where $OIS_{T_{i-1}}(T_i - T_{i-1})$ denotes the par rate of the OIS swap between dates T_{i-1} and T_i and OIS_s denotes the overnight rate. While it is not realistic, especially in the current market environment, to assume that the bank is able to fund itself at the IBOR rate, it is fair to say that the very short term funding cost of banks is close to the OIS rate. Then, if the Financial division desk has implemented the above strategy, entering in two successive payer interest swaps, the first paying IBOR, and the second paying OIS, it has secured a margin which depends on the spread between the client rate and the at par swap rate for the same period and on the spread between IBOR and OIS.

Any asset or liability which is not hedged following the procedure we have described above represents a risk for the bank and is a directional position on the future evolution of interest rates, this is called an open gap, meaning that the asset or the liabilities is not matched by another asset or liability. As mentioned before, the amount of notional involved is enormous and usually Financial division are allowed to keep a part of the gap open.

6.4 Modeling and data

As mentioned before, the primary role of the ALM division of a retail bank is building a model to estimate the future evolution of assets and liabilities. For some of these the future evolution is straightforward, this is the case of bond issues for example. However, as we have seen, even for very simple liabilities such as mortgages, the modeling of schedules requires non trivial assumptions on the behavior of clients. The archetypal example of liability are the French regulated saving accounts such as the Livret A and the PEL, which we will describe in details later.

Let us follow up with our example on mortgages. We have defined a very simple model of prepayments based on a rational behavior of clients and on the fact that the 3 months IBOR rate plus a constant spread represents a good proxy for the mortgage rate. This modeling is full of unrealistic assumptions. Clearly the behavior of clients is not rational, secondly the mortgage rate depends on various factors such as the perception of credit risk and the overall funding capabilities of the bank, which are factors that cannot be reduced to a constant spread. Finally it is reasonable to think that the number of prepayments depends on other economic and market factors such as the GDP and inflation. Given that it is crucial for the bank to estimate the number of prepayments of mortgages, the ALM divisions of retail banks put a lot of effort to understand the behavior of clients and estimate the prepayment amounts. The main mathematical tool used for this purpose is statistical analysis of historical time series of observed prepayments. Classic models try to identify significant correlation relationships between the amount of prepayments and market observable quantities. Typically the quantities that are used are the ones the bank is able to hedge using interest rates options.

Let briefly discuss the data available to analyze the behavior of assets and liabilities and to built the schedule model. The banking industry is probably the industry which has the higher number of information on its own clients. Unfortunately the information system is not designed to allow to exploit this information. It is often very difficult to get granular information on the clients. The situation is similar to the situation in the Investment Banking business, the information is often available only on an aggregate basis. For example, for a given saving account we will have the time series of the total notional of the saving accounts but we will not necessarily have the information on the notional amount of the new saving accounts that have been opened at a given date, nor the information on the notional amount of the saving accounts that have been closed. A crucial part of the modeling teams of the ALM division consists in building a set of data that can be exploited to extract relevant information on the assets and liabilities.

Chapter 7

Hedging Interest rate risk in ALM: some examples

7.1 Hedging the Livret A saving accounts

This section briefly presents part of the work I have done on the Livret A saving accounts at Crédit Agricole S.A. in collaboration with Mzoughi Karim and Krzyzaniak Stephan.

The Livret A is a state regulated tax free saving account in France. The minimum notional to open a Livret A is 1.5 euros and the maximum notional is 23000 euros, above that notional amount the account pays the interest on the total notional but no instalment can be made. The total notional of the livret A saving accounts was 268,7 billion euros in April 2014. Around half (the rules to determine exact portion of the total notional are fairly complex) of the total notional is centralized and managed by the CDC (Caisse des dpots et consignations), and used to finance public housing in France. There are other state regulated saving accounts which yield the same rate and are used to finance different public works such as sustainable development works. The total notional amount of these saving accounts was 371,1 billion euros in April 2014.

One of the main characteristics of the Livret A saving account is that its coupons are indexed on inflation. The interest rate of the product is fixed every 6 month and holds for the 6 months following the fixing. The rate is given by the formula:

$$L_t = \max\left(\frac{1}{2}\left[\frac{\overline{\text{EONIA}}_{t-2m} + \overline{\text{EURIBOR3M}}_{t-2m}}{2} + \left(\frac{CPI_{t-2m}}{CPI_{t-14m}} - 1\right)\right], \left(\frac{CPI_{t-2m}}{CPI_{t-14m}} - 1\right) + 25\text{bp}\right). \tag{7.1}$$

where $\overline{\text{EONIA}}_t$ and $\overline{\text{EURIBOR3M}}_t$ are respectively the mean of the EONIA rate and the 3M EURIBOR rate over the month t. In the following we use the notations

$$
\begin{aligned}
I_t &= \frac{CPI_t}{CPI_{t-1y}} - 1 \\
n_t &= \frac{\overline{\text{EONIA}}_t + \overline{\text{EURIBOR3M}}_t}{2}
\end{aligned}
$$

We will ignore two further options which are used to determine the interest rate. The first is that the difference between two consecutive fixings is capped and floored, the second is that the state can decide not to apply the formula (this is often the case when we observe a significant variation between two consecutive fixings).

It is worth noting that, being state regulated, the Livret A contract has changed dramatically in the last years. For example, the maximum instalment has recently been raised from 15000 to 23000 euros. For this reason, the explanatory power of time series of the total notional amount in the Livret A saving accounts is not necessarily very high. We will analyze the livret A saving accounts from an investment banking perspective. In the following we propose three different exotic structures which can be used to hedge the product.

7.1.1 Different views of the Livret A

Depending on both, the assumptions we make and the management choices, the livret A can be modeled in different ways. For example, the product is (at least in theory) perpetual. In practice it is necessary to assume that the product as a fixed expiry. The choice of the expiry will be driven by two considerations. The first that it is important that this expiry is consistent we the observed behavior of client (this will be supported by an analysis of the historical data available, and the assumptions on the behavior of clients). The second that the expiry we chose determines the expiry of the hedging instrument we will use, and therefore will be driven by liquidity and cost considerations. In all the following we will assume that the livret A saving accounts have a maximum expiry of 15 years. Another important assumptions is on the ability of the clients to close the account before expiry. Taking this right into account will introduce path dependent optionality in the product. In practice the livret A saving accounts will be hedged by stratums. The interest of the livret A are capitalized in the saving account and they are then compounded for future interests. Two alternative choices are available, either to add the interests to the new stratums, or to add them to the notional amount of the saving accounts. Depending on the choice the corresponding exotic derivative structure will show different sensitivities.

We now describe three different derivative structures which model the livret A saving accounts.

Livret A swap

The first approach to hedge the Livret A is to guarantee the coupons w.r.t. a fixed notional amount N (which corresponds to the stock at a certain date). The structure of the product is then similar to an inflation swap plus a cap on the real rate.

The swap payment frequency is six months. At every payment date t_{i+1} the product pays the amount

$$0.5NL_{t_i}$$

where the rate L_{t_i} is the Livret A rate fixed at date t_i, which can be written as

$$\begin{aligned} L_{t_i} &= I_{t_i-2m} + \max\left(\frac{n_{t_i-2m} - I_{t_i-2m}}{2}, 25\text{bp}\right) \\ &= I_{t_i-2m} + 25\text{bp} + \left(\frac{n_{t_i-2m} - I_{t_i-2m}}{2} - 25\text{bp}\right)^+ . \end{aligned}$$

The main drawback of this approach is that this product does not truly hedges the Livret A saving accounts. In practice interests are not actually perceived as a coupon by the client but capitalized in the saving account until the client decides to close its account(i.e. until maturity). In the latter we present an exotic structure that exactly replicates the livret A saving account.

The Livret A product: capitalizing on interest

Instead of modeling the product as a swap which pays the interest on the notional in the form of coupons. We define a product with a unique terminal cash flow where we have capitalized all the coupons. At maturity T the product pays

$$N \sum_{i:t_i<T} \Pi_{k=i}^{n} L_{t_i}$$

where by convention $t_n = T$.

Cancellable Livret A product

In practice the owner of the livret A has the right to close the contract at any time. The owner of the contract is thus long an American option which is not modeled in the previous structures. We thus introduce this option in the livret A product. Though the client can exercise at any date, we will assume that the client has the right to exercise every 6 month, receiving the capitalized interests on the notional. At the exercise date T^{ex} the product pays

$$N \sum_{i:t_i<T^{ex}} \Pi_{k=i}^{n_e} L_{t_i}$$

where by convention $t_{n_e} = T^{ex}$.

7.1.2 Risk analysis of Livret A product

In this section we perform a Risk analysis of the three derivative structures: the Livret A swap (LAS), the Livret A product (ELA) and the cancellable livret A product (ALA). In particular we focus on the sensitivities of the product to vanilla market instruments. All tests are performed with a standard three factors Jarrow-Yildirim model, see [JY03] calibrated to the term structure volatilities of 3 months tenor caplets and 1y CPI ratio options. Market data of 31-03-2011 is used. The delta sensitivities are computed by finite difference when bumping the market prices of the linear products which are used to calibrate the model. The nominal deltas are defined as the sensitivity of the price to a change in the market value of the nominal interest rate swap. The nominal delta of maturity kY corresponds to the sensitivity of the price to a change in the kY maturity interest rates swap. The inflation deltas are defined as the sensitivity of the price to a change in the market value of the zero-coupon inflation swap. See [BM06] for a definition of zero-coupon inflation swaps products. The inflation delta of maturity kY corresponds to the sensitivity of the price to a change in the kY maturity zero-coupon inflation swap. The vegas of the product are also computed by finite difference when bumping the market prices of optional products which are used to calibrate the model. The nominal vegas are defined as the sensitivities of the price to a change in the market value of the 3 month tenor caplets. The nominal vega of maturity kY corresponds to the sensitivity of the price to a change in the market value of the 3 month tenor caplet of expiry kY. The inflation vegas are defined as the sensitivities of the price to a change in the market value of the 1 year CPI ratio options. See [BM06] for a definition of CPI ratio options. The inflation vega of maturity kY corresponds to the sensitivity of the price to a change in the market value of the 1 year CPI ratio option of expiry kY.

Sensitivity analysis of the Livret A swap

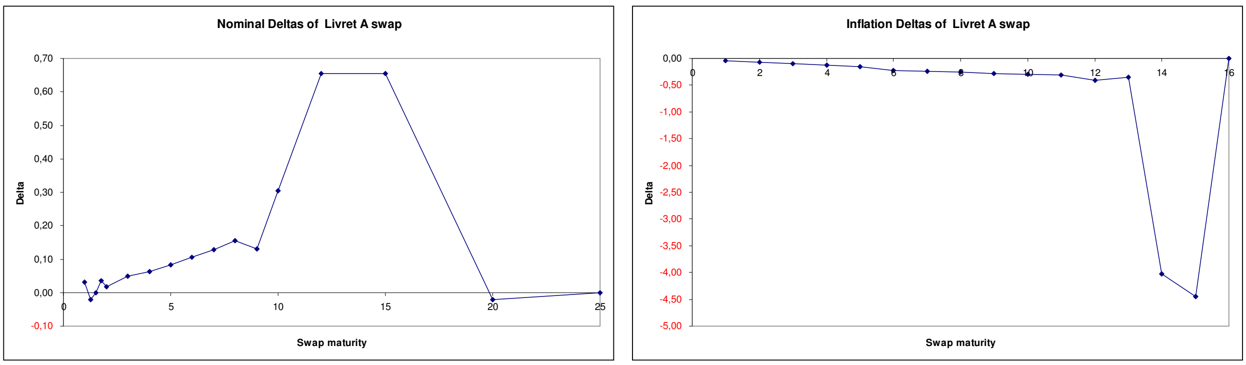

Figure 7.1: Deltas of Livret A swap.

Figure 7.1 shows the delta sensitivities of the LAS. The sensitivity is naturally concentrated around maturity (here 15 Y) for both the Nominal swap and Inflation swaps sensitivities. Note that the sensitivity to the breakeven Inflation curve movements is more important then the sensitivity to the nominal yield curve. In order to properly hedge the product we should take an important part of Inflation swaps.

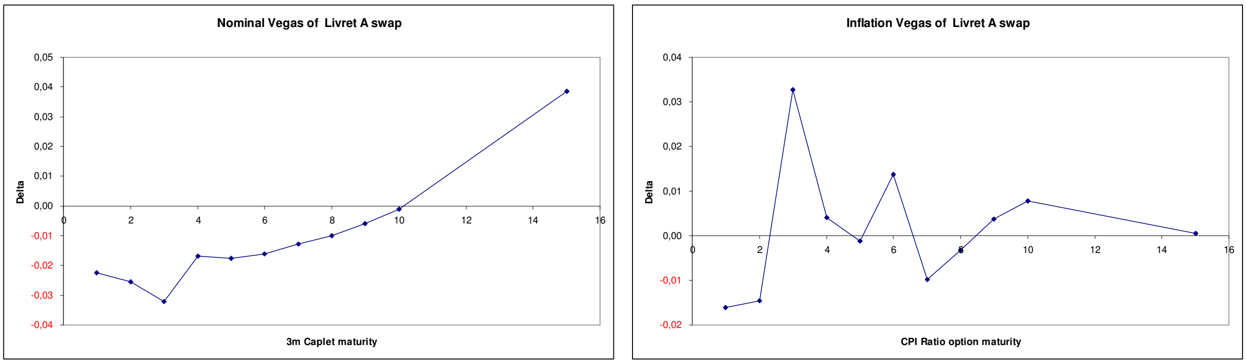

Figure 7.2: Vegas of Livret A swap.

Both Inflation and Nominal Vegas are negligible and don't seem to have a large impact on the product.

Sensitivity analysis of Livret A product

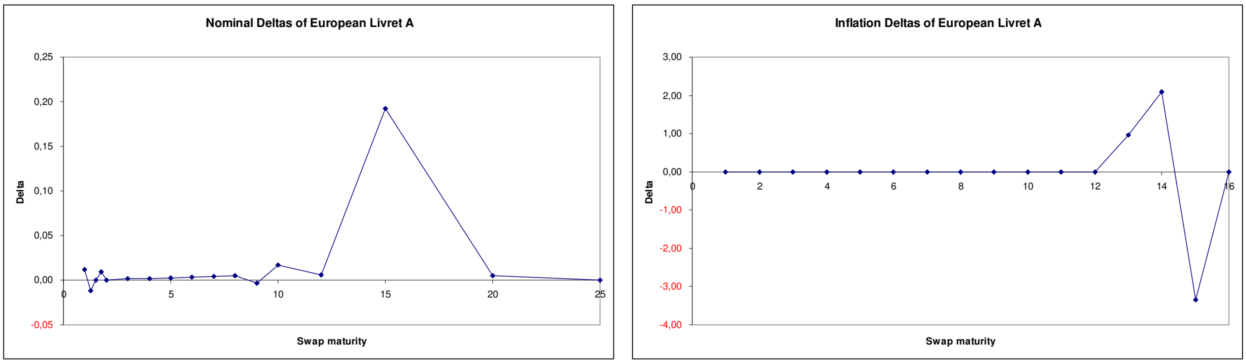

Figure 7.3: Deltas of European Livret A .

As for the LAS, figure 7.3 shows that the sensitivity is naturally concentrated around maturity (here 15 Y) for both the Nominal swap and Inflation swaps sensitivities. Again the sensitivity to the breakeven Inflation curve movements is more important then the sensitivity to the nominal yield curve.
The Nominal Vegas remain negligible and don't seem to have a large impact on the product. However the Inflation Vegas are much more important then in the case of the LAS, and in order to properly hedge we should take this volatility risk into account.

Sensitivity analysis The Delta sensitivities of the ALA (Figure 7.5) are identical to the Delta sensitivities of the ELA. The Nominal Vegas remain negligible and don't seem to have

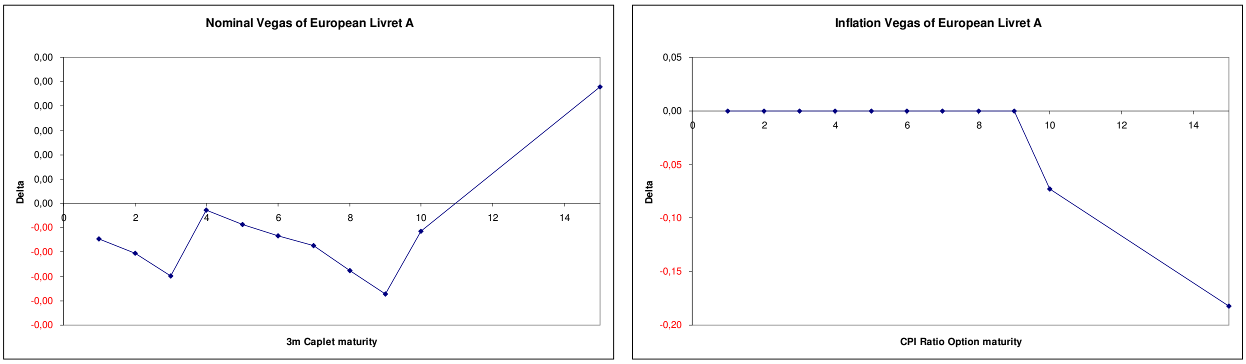

Figure 7.4: Vegas of European Livret A swap.

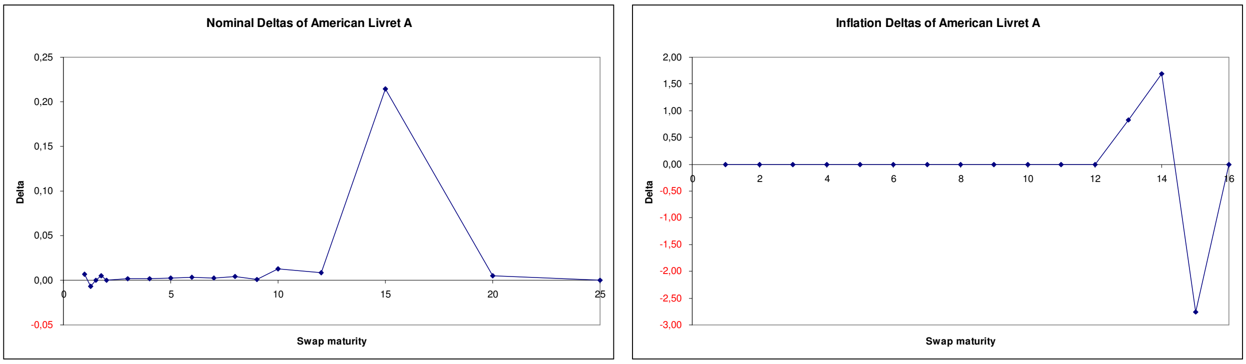

Figure 7.5: Deltas of American Livret A.

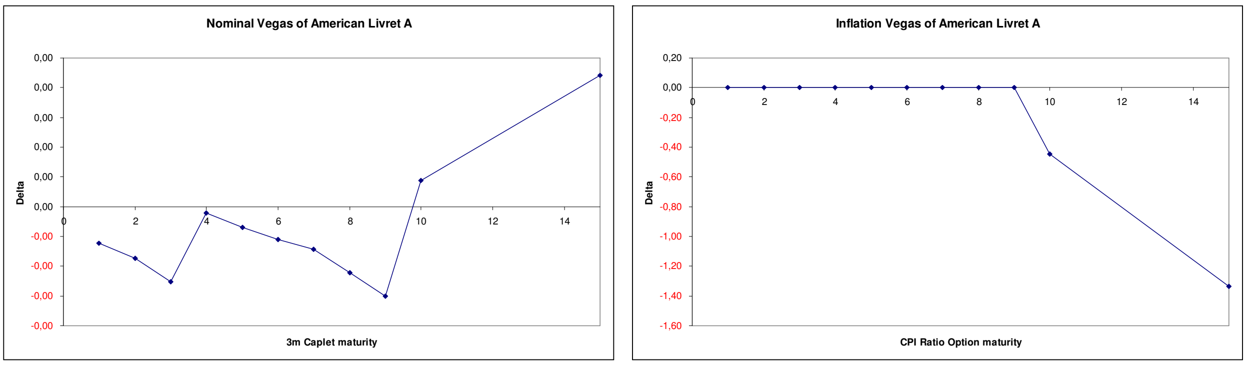

Figure 7.6: Vegas of American Livret A.

a large impact on the product. Inflation Vegas are almost ten times larger then in the case of the ELA. The Inflation volatility risk thus becomes an important one for the American structure. This volatility risk mainly comes from the exercise/cancel option (see figures 7.8 and 7.7).

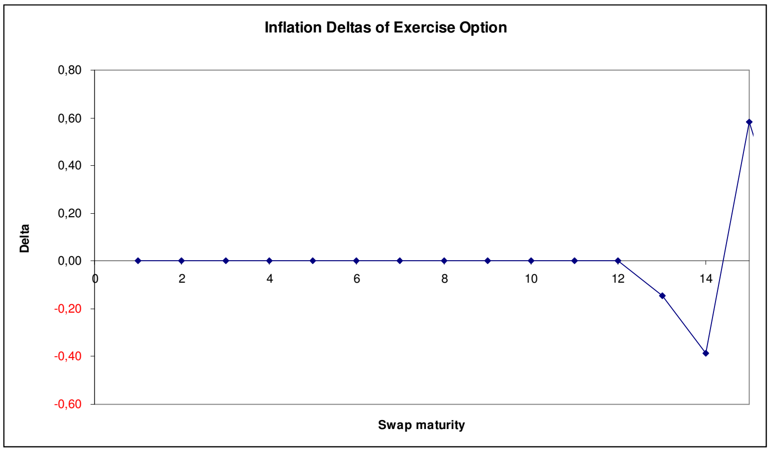

Figure 7.7: Vegas of the Option.

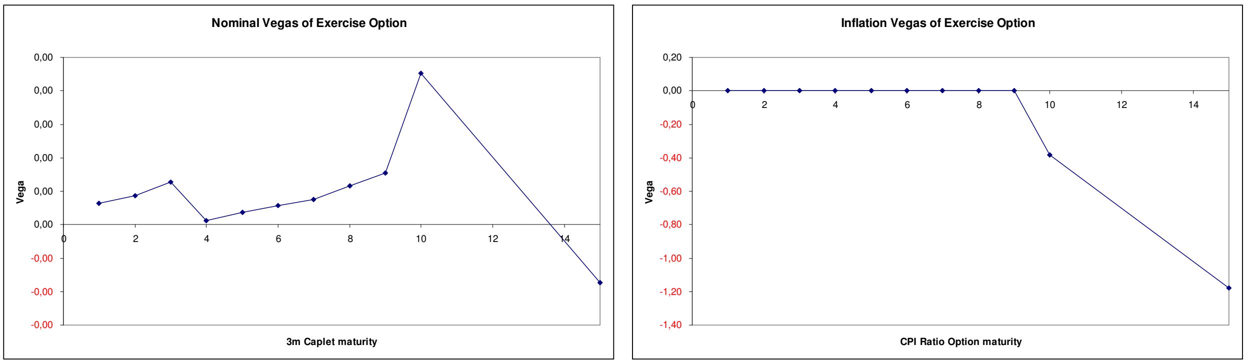

Figure 7.8: Vegas of the Option.

7.1.3 Hedging of Livret A in practice

Of the three exotic structures we have presented, the ALA is the product which is actually sold by the Bank to its clients. Thus in order to perfectly hedge the product one should take into account all the risks we have mentioned in the sensitivity analysis we have conduced for this product. The standard approach however consists in viewing the livret A saving accounts as livret A swaps. The notional is split between a nominal interest rates notional, which is swapped using the procedure described in section 6.3, and an inflation notional which is swapped with the same procedure, but using either year on year inflation swaps or zero-coupon inflation swaps[1] instead of standard IRS. These notional amounts are determined by the delta sensitivities of the livret A swap product. The option held by the clients of the product is supposed to be modeled by the deterministic schedule which incorporate all the option risks. A more accurate hedging could be achieved by defining an envelope which depends on the nominal interest rates (say the 3 months libor for example), and on the

[1]The choice between one or the other depends on the liquidity of these products in the market.

year-on-year inflation. The envelope could be determined by transforming the deltas and vegas of the ALA.

7.2 A model for PEL saving accounts

This section briefly presents part of the work I have done on the PEL saving accounts at Crédit Agricole S.A. in collaboration with Abdelmoula Omar. The Plan Épargne Logement

(PEL) is a state regulated saving account which has two separate phases, a saving phase, which can last between 4 and 10 years and includes regular instalments from the client, and a borrowing phase which allows the owner of the account to have a mortgage from the bank at a predetermined interest rate. If the owner decides to convert his saving account into a mortgage it will receive a bonus premium from the State. The amount of this premium depends on the notional of the saving account. The total notional of PEL saving accounts was estimated at 159 billions in 2009.

7.2.1 Product description

The PEL is a State regulated saving account of maximum duration 15 years. After a saving phase of minimum 4 years, the saving account can be converted into a mortgage, in this case the owner has the receives a bonus premium from the state which depends on the total notional amount of the saving account. The rules of the PEL contract have recently changed, the products we describe here are the PEL saving accounts that have been sold to clients in the period 01/08/2003 to 28/02/2011.

Below is the description of the instalement rules of the PEL saving account:

- The minimum initial instalment is 225 euro.
- The instalment frequency is either monthly, quarterly or semi-annual.
- The minimum instalment amount is 540 euros per year, 45 euros per month, 135 euros per quarter or 270 euros per semester.
- The total instalment amount cannot be higher than 61 200 euros.

The PEL contract pays a fixed interest rate on the total notional amount, we will denote this rate by γ_{PEL} in the following. For the generation of PEL contracts between 2003 and 2011, the rate was 2.5%. The minimum duration of the PEL contract is supposed to be 4 years. However, the subscriber has the right to cancel his contract at anytime before the 4^{th} anniversary of its contract. Below are the conditions to cancel the contract:

- If the subscriber closes the account before the 2^{nd} year of the contract, the interest are recalculated with a different rate, called the rate CEL (denoted γ_{CEL} hereafter) which is lower than the PEL rate.

- If the subscriber decides to close the contract between the 2^{nd} and 3^{rd} year, then the interests are calculated with γ_{PEL} but the owner loses the bonus premium and doesn't have the right to convert the contract into a mortgage.

- If the owner of the contract decides to close the contract between the 3^{rd} and 4^{th} year, the owner has the right to convert the contract into a mortgage, the notional of the mortgage is determined based on the amount the owner has saved before the 3^{rd} year. In this case the owner receives half of the bonus premium.

- If the owner of the contract decides to close the contract between the 4^{th} and 10^{th} year, the owner has the right to convert the contract into a mortgage, the notional of the mortgage is determined based on the amount the owner has saved before the last anniversary date prior the closing date. In this case the owner receives the full bonus premium.

7.2.2 Notations

Let $f \in \{2, 4, 12\}$ be the annual instalment frequency.

- **The PEL annualized interest rates**
  We denote by γ_{PEL} the annualized interest rate of the PEL contract. Let γ_f be the annualized interest rate for instalments occurring every $n = \frac{12}{f}$ months, we have

$$\gamma_f \quad = \quad (1 + \gamma_{\text{PEL}})^{\frac{1}{f}} - 1 \tag{7.2}$$

  Similarly, we define the annualized interest rate for an instalment frequency f equivalent to the CEL interest rate γ_{CEL}:

$$\gamma_f^- \quad = \quad (1 + \gamma_{\text{CEL}})^{\frac{1}{f}} - 1 \tag{7.3}$$

- **Instalments**
  We denote by v_0 the initial instalment, and M_t the total instalment amount at date t. Let us denote by $(t_k, k = 0, ..., N_I)$ the dates of the instalments.

- **The total notional amount of the PEL**
  For $t_k \leq t < t_{k+1})$ the total notional amount of the PEL contract at date t is given by:

$$N(t) = \begin{cases} v_0(1 + \gamma_{\text{CEL}})^t + \sum_{i=1}^k M_{t_i}(1 + \gamma_{\text{CEL}})^{t-t_i} & \text{if} \quad t \leq 2 \\ v_0(1 + \gamma_{\text{PEL}})^t + \sum_{i=1}^k M_{t_i}(1 + \gamma_{\text{PEL}})^{t-t_i} & \text{if} \quad 2 < t \leq 10 \\ (1 + \gamma_{\text{PEL}})^{t-10} N(10) & \text{if} \quad t > 10 \end{cases}$$

- **The notional of the mortgage**
  The owner of the PEL contract can convert the contract into a mortgage. The notional

amount of the mortgage depends on the interests perceived during the "saving phase". Let $I(t_k)$ denote the total amount of the interest perceived up to date t_k, we have:

$$I(t_k) \quad = \quad N(t_k) - v_0 - \sum_{i=1}^{k} M_{t_i} \tag{7.4}$$

The maximum notional of the mortgage is a function of $I(t_k)$

- **Bonus premium**
  If the owner of the PEL contract converts its contract into a mortgage, he receives a bonus premium. The bonus premium cannot be higher than 1525 euros, and is proportional to the maximum notional of the mortgage. The bonus premium $p_e(t)$ at date t ($t \geq 3$), is given by:

$$p_e(t) = \begin{cases} 0 & \text{ifsi} \quad t < 3 \\ \frac{1}{2}\min\left(\frac{2}{5}I(\lfloor t \rfloor), 1525\right) & \text{if} \quad 3 \leq t < 4 \\ \quad \min\left(\frac{2}{5}I(\lfloor t \rfloor), 1525\right) & \text{if} \quad 4 \leq t \leq t_{max} \\ \quad \min\left(\frac{2}{5}I(t_{max}), 1525\right) & \text{if} \quad t \geq t_{max} \end{cases}$$

  At date t_{max} the bonus premium is fixed. For the generation of PEL contracts issued between 2003 and 2011, $t_{max} = 10$.

- **Taxes**
  The bonus premium and the interests are not tax free. The taxes on are payed when the contract is closed or after the 10^{th} year. Let R_f be the tax rate (12.1% for the generation of PEL contract we have studied), the tax amount to be payed assuming we close the contract at date t is given by:

$$F(t) \quad = \quad R_f \times I(t) \tag{7.5}$$

7.2.3 The mortgage

After year 3, the PEL contract can be converted into a mortgage. The maximum notional amount of the mortgage $N_{\max}$ at date τ is given by:

$$\mathrm{N}_{\max} \quad = \quad \min\left(2.5 \times I(\lfloor \tau \rfloor)\frac{(1+\gamma_f)^T - 1}{1 + (1+\gamma_f)^T(\gamma_f T - 1)}, 92\mathrm{K}\right) \tag{7.6}$$

where T is the total number of monthly instalments of the mortgage. Assuming that the owner of the contract converts into a mortgage of notional $\mathrm{N}_{\min} \leq \mathrm{N} \leq \mathrm{N}^*$, the value of the monthly instalments is given by:

$$\overline{\mathrm{M}} \quad = \quad \frac{\rho_{\mathrm{PEL}}}{1 - \frac{1}{(1+\rho_{\mathrm{PEL}})^T}}\mathrm{N}, \tag{7.7}$$

where ρ_{PEL} is the prefixed interest rate of the mortgage.

7.2.4 A model for the evolution of the PEL contracts

Let us now build a model for the behavior of the clients of the PEL contract. Our goal is to build a model that describes the behavior of the clients of the PEL contract. We have merely followed the footsteps of [BDJ+00].

The owner of a PEL saving account holds two options which are embedded in the contract:

- The client can convert the contract into a mortgage.
- The client can close the PEL saving account.

The population of the owners of the PEL saving accounts is heterogeneous and it would be unrealistic to assume that all the clients share the same behavior. For example, some clients consider the PEL saving account as a pure saving account and are not interested in converting the contract into a mortgage[2]. A realistic model should be able to model all the different behaviors of the client population. We will introduce a utility function which will allow us to differentiate between the different client behaviors. Also, historical data shows that the clients do not follow a rational behavior, we will also introduce an uncontrolled source of randomness which models irrational behavior of the clients.

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ be a filtered probability space. Let θ be the first jump of a Poisson point process with a deterministic intensity λ, we have

$$P(\theta > t) = \exp\left(-\int_0^t ds\lambda(s)\right). \tag{7.8}$$

We assume that the intensity λ is given by

$$\lambda(t) = \begin{cases} C_2 & \text{if} \quad 3 < t \leq 4 \\ C_3 & \text{if} \quad t \geq 4 \\ C_1 & \text{otherwise} \end{cases}$$

where C_1, C_2 et $C_3 \in \mathbb{R}_+^*$. The random process θ models the time when the owner decides to close the PEL saving account for no particular reasons. We use θ to model the "irrational" behavior of the clients.

We now model the "rational" behavior of the clients by a random utility function. Let X be a random variable taking values in the interval $[0, 1]$, for example X can be taken to be uniformly distributed. We define the utility function of the client if he decides to exit at date tau by

$$f(\tau, p_e(\tau), \tau, T, K) = \left(p_e(\tau) + X\mathrm{N}_{min} + (1 - X)\mathrm{N}^* - \delta K\right)_+, \tag{7.9}$$

[2]For the PEL contracts issued between 2003 and 2011 this still gives the owner to right to benefit to the (minimum) of the state bonus premium.

where N_{min} is the minimal notional amount the client can borrow, and N^* is the maximum notional amount the client can borrow, δ represents the credit spread of the clients, and K is the net present value of the monthly payments of the mortgage, at date τ the net present value of the cost of a mortgage with T constant monthly payments is given by

$$K(\tau, T) = \overline{\mathrm{M}} \sum_{i=1}^{T} P_{\tau,\tau+\frac{i}{12}}. \tag{7.10}$$

The random variable X can be viewed as a cursor between the behavior of a client who is willing to use the PEL as a pure saving account, and a client who is willing to convert the PEL saving account into a mortgage. Given a model for the yield curve dynamics the client faces an optimal stopping problem. The value of the option at date t is given by

$$V_t = -\sum_{i=0}^{[f\times t]} M_{t_i} e^{\int_{t_i}^{t} r_s ds} +$$

$$\sup_{\tau\in\mathcal{T}_{t,t_{\max}}} \mathbb{E}\left[e^{-\int_t^{\tau\wedge\theta} r_s ds}\left(N(\tau\wedge\theta) - F(\tau\wedge\theta) + 1_{\{\theta>\tau\}} f\Big(p_e, \tau, T, K(\tau,T)\Big) + 1_{\{\theta\le\tau\}} p_e(\theta)\right)|\mathcal{F}_t\right]. \tag{7.11}$$

where $\mathcal{T}_{t,t_{max}}$ is the set of stopping times which take values in the interval $[t, t_{max}]$.

In order to derive a tractable resolution of the optimal stopping problem, we will make some assumptions. We assume that "irrational closing time" θ is independent of the yield curve dynamics. We also assume that the distribution of the clients between savers and mortgage borrowers X is independent of the yield curve dynamics. This allows us to study the problem as if both the variables θ and X where deterministic, and then compute the result of the problem as the expectation under the law of X and θ of the solution. We will also assume that the instalments of the client are deterministic as if the owner of the PEL saving accounts decided its saving strategy at inception of the contract.

7.2.5 A PDE resolution in the Vasicek model

We solve the optimal stopping problem (7.11) in the Vasicek model [Vas77]. The spot rate is assumed to be solution of the SDE,

$$dr_t = k(b - r_t)dt + \sigma dW_t. \tag{7.12}$$

We assume that X and θ are deterministic, therefore the utility function for $x \in [0,1]$ is given by

$$f_u(p_e, \tau, T, K(\cdot,\cdot)) = \Big(p_e(\tau) + x\mathrm{N}_{min} + (1-x)\mathrm{N}^* - \delta K(\tau, T)\Big)_+$$

.

In the Vasicek model the yield curve is a function of the spot rate only, therefore K is also a function of the spot rate only, the utility function f is also a function of the spot rate only and of the characteristics of the contract. We define

$$f(t, r_t) := f_u(p_e, \tau, T, K(t, T))$$

The value of the optimal stopping problem (7.11) can be viewed as the price of an American option held by the client. Ignoring the cost of the instalments up to date t, we define the value function as

$$V(t, r_t) := \sup_{\tau \in \mathcal{T}_{t,t_{\max}}} \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^{\tau \wedge \theta} r_s ds} \left(N(\tau) - F(\tau) + f(t, r_\tau) \right) - \sum_{i=[f\times t]+1}^{[f\times\tau]} M_{t_i} e^{-\int_t^{t_i} r_s ds} | \mathcal{F}_t \right]. \tag{7.13}$$

The underlyings (the total notional amount of the PEL saving account, the bonus premium and the cost of the mortgage) are stepwise constant functions. It is simpler to resolve a continuous version of the problem rather the original one. Let us define

$$V^c(t, r_t) = \sup_{\tau \in \mathcal{T}_{t,t_{\max}}} \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^{\tau \wedge \theta} r_s ds} \left(E(\tau) - F(\tau) + f(t, r_\tau) \right) - \int_t^{\tau} M(s) e^{-\int_t^s r_\xi d\xi} | \mathcal{F}_t \right].$$

The price is solution of the following parabolic partial differential inequality

$$\begin{cases} \frac{\partial V}{\partial t} + (b(t) - r)\frac{\partial V}{\partial r} + \frac{\sigma^2}{2}\frac{\partial^2 V}{r^2} - rV - M(s) \leq 0, V(t, r) \geq f(t, r) \quad \text{dans} \quad [0, t_{max}] \times \mathbb{R} \\ (\frac{\partial V}{\partial t} + (b(t) - r)\frac{\partial V}{\partial r} + \frac{\sigma^2}{2}\frac{\partial^2 V}{r^2} - rV - M(s))(f(t, r) - V(t, r)) = 0 \quad \text{dans} \quad [0, t_{max}] \times \mathbb{R} \\ V(t_{max}, r) = f(t_{max}, r) \end{cases}$$

There are no explicit solutions to the above differential inequality, but we can easily solve it with numerical schemes such as the implicit Euler finite difference scheme. The numerical solution of the above partial differential inequality provides a grid of prices in the time and space dimensions. It is then straightforward to approach the exercise boundary of the option. The continuation region is given by

$$\mathcal{C} = \{(t, r) \in [t, t_{max}] \times \mathbb{R} : V(t, r) > f(t, r)\}.$$

Defining τ^*

$$\tau^* = \inf\{t : (t, r) \in \mathcal{C}\},$$

we know that τ^* is the smallest stopping time which realizes the maximum of the value function 7.13. Using Monte Carlo simulation we can easily estimate the distribution of τ^* by simulating different trajectories of the spot rate in the Vasicek model.

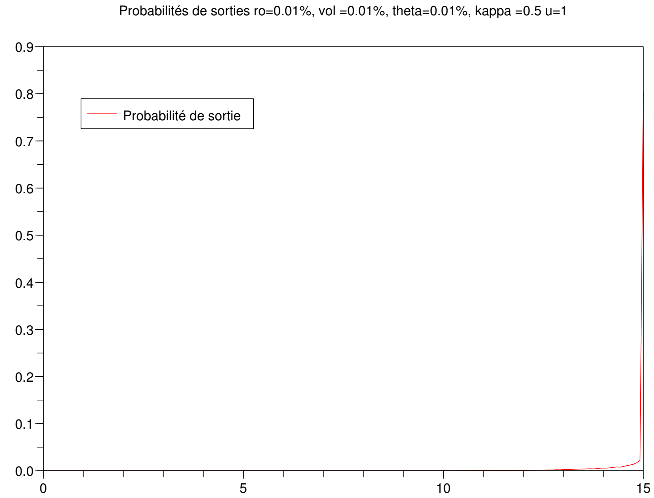

Figure 7.9: Behavior of the saver, i.e. $x = 1$. Plot of the distribution of the optimal stopping time. $\kappa = 0.5, \theta = 1\%, \sigma = 1\%, r_0 = 1\%$. The figure is obtained with 10^5 Monte Carlo simulations.

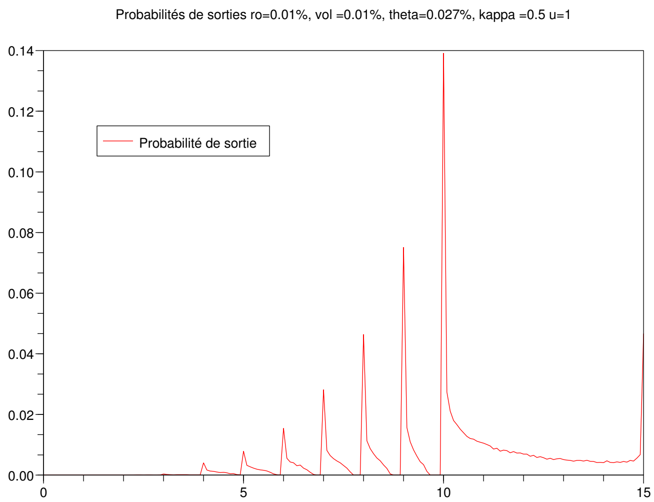

Figure 7.10: Behavior of the saver, i.e. $x = 1$. Plot of the distribution of the optimal stopping time. $\kappa = 0.5, b = 2.7\%, \sigma = 1\%, r_0 = 1\%$. The figure is obtained with 10^5 Monte Carlo simulations.

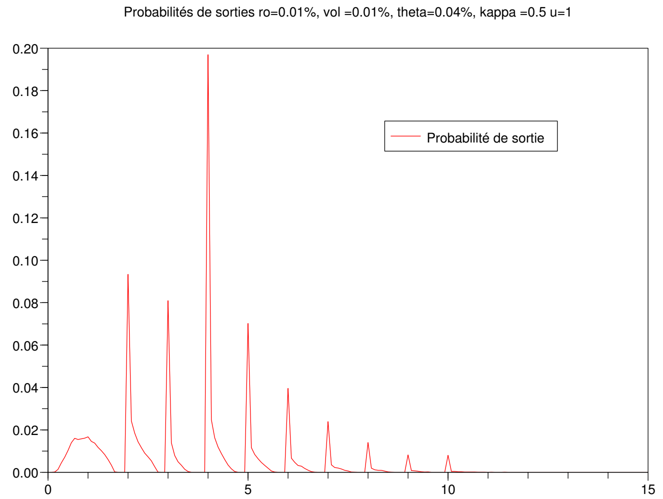

Figure 7.11: Behavior of the saver, i.e. $x = 1$. Plot of the distribution of the optimal stopping time. $\kappa = 0.5, b = 4\%, \sigma = 1\%, r_0 = 1\%$. The figure is obtained with 10^5 Monte Carlo simulations.

7.2.6 Numerical results

We have studied the distribution of the optimal stopping time, i.e. the time when the client decides either to close its account or to convert the contract into a mortgage. We observe that the behavior of the client is significantly different depending on the utility function, i.e. on the value of the parameter x.

Figures 7.9, 7.10 and 7.11 represent the distribution of the optimal stopping time of the saver, i.e. the time when the client decides to close its PEL saving account. The three figures represent three different interest rates scenarios, respectively a scenario where rates remain low (Figure 7.9), a scenario where rates increase to a level which is close to the interest rate of the PEL saving account γ_{PEL}, which is 2.5% for the considered PEL generation (Figure 7.10), and a scenario where rates will rapidly increase to a level around 4%. The different scenarios are determined by the value of the parameter b, which is the long term mean of interest rates. Depending on the scenario, the "saver" will close its PEL saving account at a different time. We observe that the saver tends to close its account at the dates of anniversary of the contract in order to optimize the state bonus premium (which is fixed at these dates). Also, even in an increasing interest rates scenario the saver tends to wait until the second anniversary before closing the account, because closing the account before that date would mean receiving the interests at the CEL interest rate which is significantly lower than the PEL interest rate.

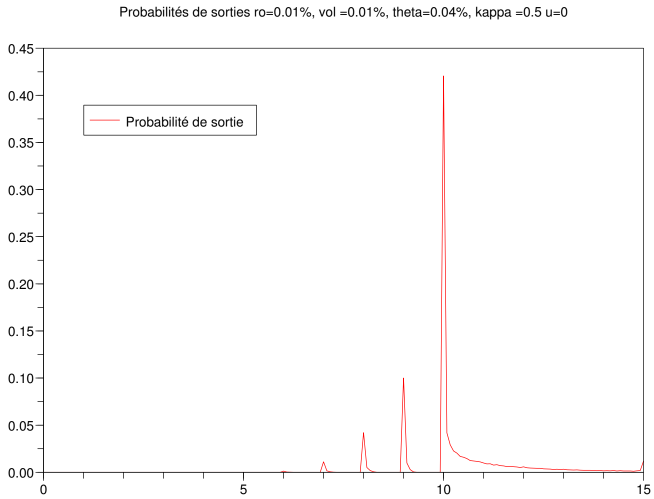

Figure 7.12: Behavior of the mortgage borrower, i.e. $x = 0$. $\kappa = 0.5, b = 4\%, \sigma = 1\%, r_0 = 1\%$. The figure is obtained with 10^5 Monte Carlo simulations.

The behavior of the mortgage borrower is significantly different. Figure 7.12 represents the distribution of the optimal stopping time of the mortgage borrower in an increasing interest rates scenario. For a mortgage borrower an increasing interest rates scenario will not necessarily be an incitation to convert the contract into a mortgage earlier. There are two major drivers of the mortgage borrower decision, first he wants to accumulate enough interest to maximize the notional amount he can borrow, secondly he wants to optimize the spread between the fixed rate of its mortgage and the market rate for borrowing (which in our modest framework is only dependent on the evolution of interest rates). Figure 7.12 shows that the highest number of conversion occurs between the 8^{th}, 9^{th} and 10^{th} anniversary of the contract, which is the moment when the notional amount of the mortgage becomes maximum.

Let us note that the results we have presented are based on an assumption of constant instalment. We have tested different amount for the instalment amounts during the saving phase, and we have noticed that this amount doesn't change the behavior of clients, it only anticipates (if the instalments are higher) or postpones (if the instalments are lower) the optimal stopping time of the client, since depending on the instalment amount the maximum mortgage notional, and the maximum state bonus premium will be reached sooner or later.

We have also performed a pseudo backtesting of the model checking what would have been the clients behaviors if they had followed the model. We have tested two different situation:

- We assume that the client calibrates the model at inception of the contract and then makes its decision based on the comparison between the spot rate and the barrier given

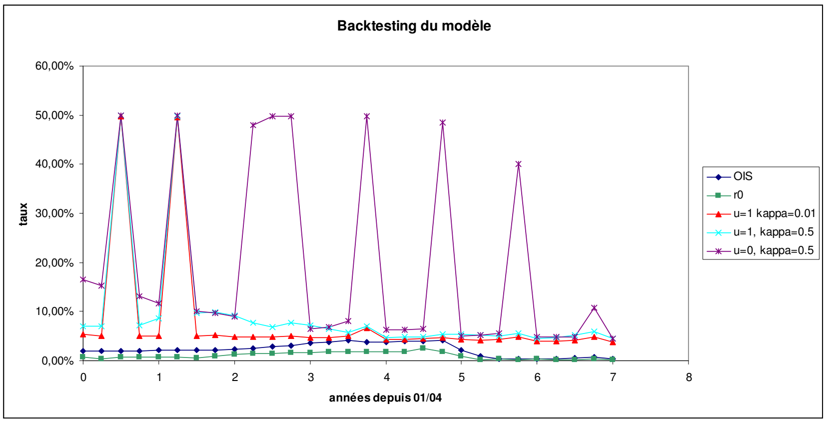

Figure 7.13: Backtesting of the model, calibration at inception. Plot of the barrier for different clients behaviors against the evolution of the extrapolated spot rate and the OIS rate.

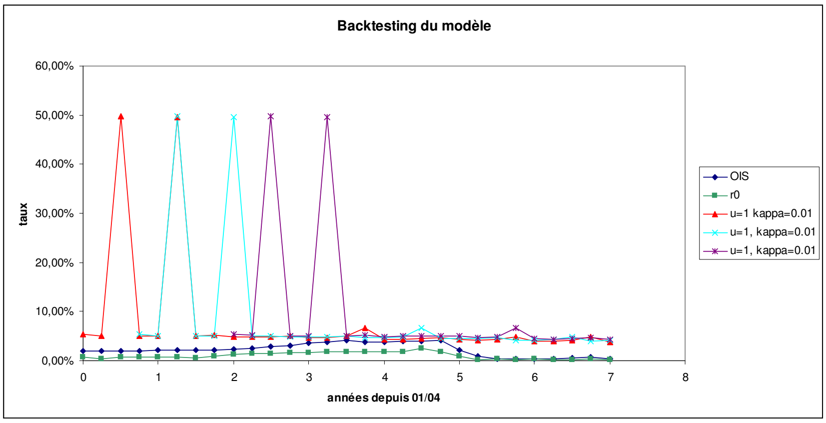

Figure 7.14: Backtesting of the model, calibration every day. Plot of the barrier for different clients behaviors against the evolution of the extrapolated spot rate and the OIS rate.

by the model at every future date. See Figure 7.13.

- We assume that the client calibrates the model every day and based on the calibrated model compares the spot rate with the barrier given by the model. See Figure 7.14.

The results of the backtesting indicate that in the period going from January 2004 to January 2011, the exercise barrier for different types of clients is higher than the observed spot rate. The barrier is significantly high for the mortgage borrowers and these clients should not have converted their contract if they had followed the model. On the other hand, the savers might have closed their account in the period going from January 2008 and January 2009. For the analyzed period, we have observed a behavior of clients which is significantly different to the one suggested by the model.

7.3 Conclusion of the chapter

In this chapter we have presented two archetypal examples of saving accounts. These saving accounts include embedded and explicit optionality, and require non-trivial modeling to be represent correctly. As we have seen, the modeling is very sensitive to the assumptions made on the behavior of the clients. For example, we have proposed three different exotic derivative structures on Inflation and on rates which are meant to hedge/mirror the livret A saving account exposure of the bank. The first order risks, namely the inflation and nominal delta are similar across the three structures (the difference comes from the hedging choice, i.e. swap as opposed to in-fine). The second order risks, namely the vegas, which come from the optionality of the product are very different across the three different structures. These three structures will imply significantly different hedging strategies.

Embedded and explicit options appear naturally in saving accounts and credit hedging and these must be included in the hedging strategy. Ignoring these options can imply a very important interest rate risk. For example most of the PEL saving accounts issued before 2003 were yielding an interest rate which is around 4-5%. While the mortgage borrowers have probably converted their account into a mortgage since then, the savers are clearly holding tide their account which yields a rate which is much higher than the rates available in the market. Clearly the owner of PEL contracts issued before 2003 are keeping their contract much longer than the previous generations, and this behavior could be hardly predicted from historical data. Hedging with an envelope and including an optimal exercise modeling could help to predict such a behavior.

Chapter 8

Hedging of options in ALM

This chapter presents part of the work I have done on the hedging of options in the interest rate gap at Crédit Agricole S.A.. The project was motivated by the needs of the Financial department and has been conduced with the supervision of the ALM methodology team. I am thankful to Yves Glaser (Head of the ALM methodology team at Crédit Agricole S.A.) and Stephan Krzyzaniak (Deputy head of the methodology team at the time of the study) for their comments.

8.1 Some questions about Delta hedging in ALM

In the previous chapters, we have seen how optionality plays a crucial role in representing assets and liabilities of a bank. Assumed that we have defined a model to determine the envelope of an asset or a liability, and that this envelope can be replicated by options, the question of how to hedge such optionality remains open. The most natural idea would be to buy options in the market, however this is not straightforward. As we have mentioned before, the economic result of the ALM activity is measured in accrual, while interest rates derivative options are traded in OTC markets and valued MtM, therefore options should be put in a trading book. As a result including options to hedge the assets and liabilities exposure often requires the creation of mirroring books (one on the trading side and the other on the accounting accrual side). Due to different rules of valuation and different capital charge treatments between the banking and the trading book, this leads to complex accounting treatments. In practice options are hedged by including the delta hedges of the options in the gap. Instead of buying the option itself, the total notional of the asset or the liability to be swapped to a floating rate is augmented by a notional amount which is equal to the delta of the option.

Ignoring the bias coming from the fact markets are naturally incomplete and continuous hedging strategies cannot be implemented, it is in theory possible to perfectly replicate any option by implementing a delta hedging strategy. However things are not so easy in practice, and there are differences coming from the accounting rules used in ALM. We have already pointed out a number of fundamental differences between the MtM approach and the accrual approach. In theory the hedging portfolio is made of a quantity of asset (the delta) and a quantity of cash (lent or borrowed). In the interest rates derivatives market the assets used for delta hedging are typically IRS, FRA and Futures, and these contracts are typically traded at par. Investment banks tend to separate their activity into very specialized desks. It is rare that the traders who perform the delta hedging of optional products directly manage the funding of their position. Furthermore the position is globally hedged so that it is very difficult to track the actual hedging portfolios of the individual trades. Simply put, we could say that the mechanics of hedging works in such a way that the trader sells its hedging portfolio every night (at its MtM value) and starts from zero the day after. Of course, the trader is charged or payed for the funding of its position, but he doesn't manage directly the problems coming from the fixings of the swaps he trades, his vision is a pure cash settlement vision, and all the contracts will usually be offset before they settle. Things are significantly different in the accrual world. The swaps cannot be offset before they settle, or continuously valued on a MtM basis. These swaps will be valued for the actual cash flow they produce. The management of the fixings is a crucial part of the ALM management (especially considered the notional amounts involved). For example, while from a MtM perspective there is no difference between trading a FRA and trading the static replicating portfolio made of IRS (assumed that both instruments are liquidly traded), from an ALM (in accrual) perspective the two contracts will generate completely different cash flows (the FRA will only generate a cash flow at one date, while the two swaps will generate cash flows at the payment dates of the swaps), and thus lead to different economic results. Last but not least, it is important to keep in mind that the option premium is not directly payed by the client. Options are often held by the clients, and the interest rates payed (resp. received) by the bank to (resp. from) the clients are adjusted to incorporate the price of the option premiums. In other words the bank option premium is spread over the cash flows of the asset or the liability.

8.2 Delta hedging in theory: chosing the numeraire

Let us first recall some fundamental results on the delta hedging interest rates option. For simplicity, we will assume a one dimensional Markovian interest rates dynamics. Let us denote by X the underlying state variable, the zero-coupon function is denoted by $P_{t,T} = P(t,T,X_t)$, the forward libor rate $L_t(T,\delta) = L(t,T,\delta,X_t)$, and the forward swap $S_t(T,n,\delta) = S(t,T,n,\delta,X_t)$.

8.2.1 The classic risk-neutral delta hedging portfolio

Let $H = f(X_T)$ the payoff of the option at expiry T. The time t price of the option paying H at expiry T is given by

$$M(t,X_t) = \mathbb{E}^Q\left[e^{-\int_t^T ds r_s} f(X_T)|\mathcal{F}_t\right],$$

where Q denotes the risk-neutral measure. The risk-neutral dynamics of the price is given by

$$dM(t,X_t) = r_t M(t,X_t)dt + \partial_x M(t,X_t)\sigma(X_t)dW_t,$$

where we have assume that the underlying state variable X follows the following dynamics under the risk-neutral measure

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t.$$

Let $S_t = S(t,X_t)$ be an hedging instrument (in the following S will be an interest rate swap). The portfolio $\phi_t^r = (\Delta_t^r, C_t^r)$, containing a quantity Δ_t^r of the asset S, and C_t^r cash, defined by

$$(\Delta_t^r, C_t) = \left(\frac{\partial_x M(t,X_t)}{\partial_x S(t,X_t)}, \tilde{M}(t,X_t) - \Delta_t^r \tilde{S}(t,X_t)\right),$$

where $\tilde{M}$ and $\tilde{S}$ are respectively the discounted value of the option M and the hedging instrument S, is self-finance and replicates the price of the option M.

Hedging of a caplet

Let us analyze the construction of the hedging portfolio of a caplet. Let $\text{Caplet}(t,T,\delta,K,X_t)$ be the time t price of a caplet of expiry T, maturity δ and strike K. It is natural to hedge this option with FRA expiring at T and maturing at $T+\delta$ of fixed rate K. The payoff of the FRA is given by

$$(K - L_T(T,\delta))\delta \quad \text{at} \quad \text{date} \quad T+\delta.$$

The time t price is given by

$$P(t,T+\delta,X_t)(K - L(t,T,\delta,X_t)\delta.$$

The hedging portfolio of time t is given by a quantity Δ_t^r of FRA:

$$\Delta_t^r = \frac{\partial_x \text{Caplet}(t,T,\delta,K,X_t)}{\delta\left(\partial_x P(t,T+\delta,X_t)\left(K - L(t,T,\delta,X_t)\right) - P(t,T+\delta,X_t)\partial_x L(t,T,\delta,X_t)\right)}. \tag{8.1}$$

This is the classic delta hedging strategy of the option with a portfolio consisting in the underlying asset and a cash amount held in the money market account.

8.2.2 Alternative hedging: hedging under the forward measure

We now consider an alternative hedging strategy which is proposed in [Ada08]. In the interest rates space, it is standard to perform a change of measure in order to derive an easier dynamics of the underlying of the option under this measure. These change of measures correspond to a choice of a different numeraire. For a given numeraire, we can always build a delta hedging portfolio consisting in the underlying asset and the numeraire which perfectly replicates the option. In the following, instead of hedging using the bank account/cash numeraire, we will hedge using the terminal zero-coupon numeraire. This is equivalent to hedge the forward value of the option. We define the T-forward value by

$$V(t, X_t) = \mathbb{E}^{Q^T}\left[f(X_T)|\mathcal{F}_t\right],$$

where Q^T denotes the forward neutral measure. The dynamics of the forward value under the measure Q^T is given by

$$dV(t, X_t) = \partial_x V(t, X_t)\sigma(X_t)dW_t.$$

Let S be an hedging instrument, denote $S_t^T = S^T(t, X_t)$ its T-forward value, defined by

$$S^T(t, X_t) = \mathbb{E}^{Q^T}\left[S_T|\mathcal{F}_t\right].$$

The portfolio $\phi_t^f = (\Delta_t^f, E_t^f)$, containing Δ_t^f of the asset S and a loan of maturity T of a notional E_t^f, i.e. containing a quantity E_t^f of the zero-coupon $P_{t,T}$

$$(\Delta_t^f, E_t^f) = \left(\frac{\partial_x V(t, X_t)}{\partial_x S^T(t, X_t)}, V(t, X_t) - \Delta_t^f S^T(t, X_t)\right),$$

is self-financed and replicates the option value.

Forward hedging of a caplet

Let us analyze the construction of the forward neutral hedging portfolio of a caplet. Let $\text{VCaplet}(t, T, \delta, K, X_t)$ be the time t $T+\delta$-forward value of a caplet of expiry T, maturity δ and K, i.e.

$$\text{VCaplet}(t, T, \delta, K, X_t) = \mathbb{E}^{Q^{T+\delta}}\left[(L_T(T, \delta) - K)^+ | X_t\right].$$

It is natural to hedge this option with FRA expiring at T and maturing at $T+\delta$ of fixed rate K. The payoff of the FRA is given by

$$(K - L_T(T, \delta))\delta \quad \text{at} \quad \text{date} \quad T + \delta.$$

Its $T+\delta$-forward value at time t is given by

$$(K - L(t, T, \delta, X_t)\delta.$$

The hedging portfolio of the $T+\delta$-forward value of the option contains a quantity Δ_t of the FRA given by:

$$\Delta_t^f = \frac{\partial_x \text{VCaplet}(t, T, \delta, K, X_t)}{-\delta \partial_x L(t, T, \delta, X_t)}. \tag{8.2}$$

8.2.3 Comparison of the risk neutral and forward neutral measure

The two hedging strategies we have presented above are equivalent in the sense these both perfectly replicate the option value.

A first striking difference between the two strategies is the expression of the delta. Comparing the expression of the risk-neutral delta (8.2) with the expression of the forward-neutral delta (8.2) we can see that the denominator of the forward-neutral delta does not depends on the strike. This implies that, in a one-dimensional Markovian term structure model, the forward-neutral delta is independent of the fixed rate of the FRA used to hedge the caplet.

In the risk-neutral hedging strategy, the cost of funding of the delta position (which is given by the interest payed on the money market account) will directly impact the value of the portfolio. The interest payed or received on the netted position $M_t - \Delta_t S_t$ will generate cash flows in the gap. On the other side, the forward-neutral strategy replaces the money market account (i.e. the funding costs) by a zero-coupon (or equivalently a loan). Instead of continuously paying the costs of funding of the adjustments of the delta position, we fund these adjustment with a loan of maturity the expiry of the trade.

Let ut analyze the cash flows generated by the two hedging strategies. At date $t = 0$, the values of the risk-neutral and forward-neutral hedging portfolios are given by

$$
\begin{aligned}
V_0(\phi) &= \Delta_0 S_0 + (M_0 - \Delta_0 S_0) \\
V_0(\phi^f) &= \Delta_0^f S_0 + \frac{M_0 - \Delta_0^f S_0}{P_{0,T}} P_{0.T}.
\end{aligned}
$$

Suppose we hedge only at discrete dates $t_0 = 0 < t_1 < \cdots < t_n < T$. At date t_1 we have to adjust the delta position in the underlying asset. In the risk-neutral strategy we fund the adjustment at the money market account rate, in the forward-neutral strategy, we fund the adjustment with a loan at maturity. The value of the self financed portfolios at date t_1 is given by

$$
\begin{aligned}
V_{t_1}(\phi) &= \Delta_1 S_{t_1} + (M_0 - \Delta_0 S_0) e^{\int_0^{t_1} ds r_s} + (\Delta_0 - \Delta_1) S_{t_1} \\
V_{t_1}(\phi^f) &= \Delta_1^f S_{t_1} + \frac{M_0 - \Delta_0^f S_0}{P_{0,T}} P_{t_1,T} + \frac{(\Delta_0^f - \Delta_1^f) S_{t_1}}{P_{t_1,T}} P_{t_1,T}.
\end{aligned}
$$

While the interest rates payed on the adjustment of the risk-neutral position depends on the evolution of interest rates between t_1 and the expiry T, the cost of the of the adjustment of the forward-neutral position is known and date t_1, and is equal to $\frac{(\Delta_0^f - \Delta_1^f) S_{t_1}}{P_{t_1,T}}$. At maturity, the P&L of the hedged position, i.e. the tracking error is given by

$$
\begin{aligned}
P\&L_T(\phi) &= f(S_T) - \left(\Delta_n S_T + (M_0 - \Delta_0 S_0) e^{\int_0^T ds r_s} + \sum_{i=1}^{n} (\Delta_{i-1} - \Delta_i) S_{t_i} e^{\int_{t_i}^T ds r_s} \right) \\
P\&L_T(\phi^f) &= f(S_T) - \left(\Delta_n^f S_T + V_0 - \Delta_0^f S_0^T + \sum_{i=1}^{n} (\Delta_{t_{i-1}}^f - \Delta_{t_i}^f) S_{t_i}^T \right).
\end{aligned}
$$

The interests payed on the successive adjustments of the delta positions of the risk-neutral hedging strategy depend on the evolution of the interest rates, and the actual cash flows are determined at maturity. On the other hand, the cost of the delta adjustments to adjust the delta position of the forward-neutral strategy is determined at the time of the adjustment and will be payed at the expiry of the option. To some extent, we can argue that the forward-neutral hedging strategy is more consistent with the philosophy of hedging in ALM, since it allows to predict the future cash flows.

8.2.4 Hedging in practice

Let us now recall that the hedging is performed with at par IRS and FRAs. Which means that, in practice the hedging instrument S_t slides. Let us assume that at date t_i we hedge with the underlying S^i, such that $S_{t_i}^i = 0$. Let us define $\tilde{\Delta}_i^f$ as follows:

$$
\begin{aligned}
\tilde{\Delta}_0^f &= \frac{\partial_x V(t_0, X_0)}{\partial_x S_0^T(t_0, X_0)} \\
\tilde{\Delta}_i^f &= \frac{\partial_x V(t_i, X_{t_i})}{\partial_x S_i^T(t_0, X_0)} - \sum_{k=0}^{i-1} \tilde{\Delta}_k^f \frac{\partial_x S_k^T(t_i, X_{t_i})}{\partial_x S_i^T(t_i, X_{t_i})}, \quad i = 1...n.
\end{aligned}
$$

Let us note that in the case of the caplet we have $\partial_x S_k^T(t_i, X_{t_i}) = \partial_x S_l^T(t_i, X_{t_i}), \quad \forall k, l, i.$ Then the expression of $\tilde{\Delta}_i^f$ simplified, and becomes

$$
\begin{aligned}
\tilde{\Delta}_0^f &= \Delta_0^f \\
\tilde{\Delta}_i^f &= \Delta_i^f - \Delta_{i-1}^f, \quad i = 1...n.
\end{aligned}
$$

From an ALM perspective the additive property of the forward-neutral deltas when rolling the underlying is very convenient. The quantity Δ_t^f gives the total notional invested in the hedging instruments, and the quantity $\Delta_i^f - \Delta_{i-1}^f$ gives the notional amount that has to be invested in the hedging at date t_i. This additive property is not verified by the risk-neutral hedges of caplets, this is because the derivative $\partial_x S_k^T(t_i, X_{t_i})$ depends on the fixed rate of the FRA, which is $L_{t_k}(T, \delta)$.

Assuming that the dates of rebalancing of the portfolio are close enough, the portfolio with a quantity $\tilde{\Delta}_i^f$ of asset S_i hedges the option. Under the T-forward measure Q^T we have

$$
\begin{aligned}
\tilde{\Delta}_0^f dS_0^T(t, X_t) &= \frac{\partial_x V(t_0, X_0)}{\partial_x S_0^T(t_0, X_0)} \partial_x S_0^T(t, X_t) \sigma(X_t) dW_t \sim dV_t, \quad t \in (t_0, t_1) \\
\sum_{k=0}^{i} \tilde{\Delta}_k^f dS_k^T(t, X_t) &= \frac{\partial_x V(t_i, X_{t_i})}{\partial_x S_i^T(t_0, X_0)} \partial_x S_i^T(t, X_t) \sigma(X_t) dW_t \sim dV_t \quad t \in (t_i, t_{i+1}).
\end{aligned}
$$

At maturity the P&L of the hedged position is given by

$$
P\&L_T(\tilde{\phi}^f) = f(X_T) - \left(\sum_{k=0}^{n} \tilde{\Delta}_k^f S_k^T(T, X_T) + V_0 \right). \tag{8.3}
$$

In this case the successive adjustment of the delta position are "free", since the underlying asset is "rolled" so that we always trade at par. The P&L at expiry of the hedging strategy $\tilde{\phi}^f$ should be compared with the cash flows of the risk-neutral hedging strategy $\tilde{\phi}$ with a rolling underlying, which is given by

$$
P\&L_T(\tilde{\phi}) = f(X_T) - \left(\sum_{k=0}^{n} \tilde{\Delta}_k S_k(T, X_T) + M_0 e^{\int_0^T ds r_s} \right). \tag{8.4}
$$

The term $M_0 e^{\int_0^T ds r_s}$ in the P&L of the risk-neutral strategy, which comes from the cash position, depends on the evolution of interest rates until the expiry of the option. Conversely the term V_0 in the P&L of the forward-neutral strategy, which comes from the zero-coupon position, is determined at inception of the trade. To some extent this is more consistent with the philosophy of ALM. For example it is possible to increase or decrease the interest rate received from or payed to the client exactly of this amount, and the margin is locked even if we don't take any cash or zero-coupon position.

8.3 Numerical implementation in the Hull-White model

We now perform a numerical analysis of the hedging strategies we have presented in the previous section. We consider a dramatically simplified problem setting: the hedging of a caplet using FRAs in the Hull-White model [HW94].

8.3.1 The derivations of the option and hedges values in the model

We assume that the spot rate follows an Ornstein-Uhlenbeck dynamics under the risk-neutral probability. Precisely, we define a state variable X solution of the following SDE under the risk-neutral measure

$$dX_t = -aX_t dt + \sigma dW_t.$$

We assume that the spot rate is given by

$$r_t = \alpha(t) + x_t,$$

where α is a deterministic function used to fit the initial yield curve. The time t price of a Caplet of expiry T, maturity δ and strike K is given by

$$\begin{aligned} \text{Caplet}_t(T,\delta,K) &= \frac{P_{t,T+\delta}}{\delta}\mathbb{E}^{Q^{T+\delta}}\left[\left(\frac{P_{T,T}}{P_{T,T+\delta}} - (1+\delta K)\right)^+ |\mathcal{F}_t\right] \\ &= \frac{P_{t,T+\delta}}{\delta} BS\left(1+\delta L_t(T,\delta), 1+\delta K, \sigma\frac{1-e^{-a\delta}}{a}\sqrt{\frac{1-e^{-2a(T-t)}}{2a}}\right), \end{aligned}$$

where

$$\begin{aligned} BS(S,K,\Sigma) &= S\phi(d_+) - K\phi(d_-) \\ d_\pm &= \frac{\log(S/K) \pm \Sigma^2/2}{\Sigma}. \end{aligned}$$

The $T+\delta$-forward value of the caplet is given by

$$\text{FCaplet}_{t,T+\delta}(T,\delta,K) = BS\left(1+\delta L_t(T,\delta), 1+\delta K, \sigma\frac{1-e^{-a\delta}}{a}\sqrt{\frac{1-e^{-2a(T-t)}}{2a}}\right)/\delta.$$

Remark 45 — *Let us note that the payoff of the caplet is known at the expiry date T. It is standard to cash-settle the option at this date. However in a pure physical settlement logic (which is the prevailing logic in ALM), the actual cash flow of the option occurs at date $T+\delta$. Therefore in all the following we consider the $T+\delta$-forward values of the caplet and of the FRA used to hedge the option.*

The forward libor rate function $L(t,T,\delta,x)$ and its partial derivative with respect to x is given by

$$\begin{aligned}
L(t,T,\delta,x) &= \frac{1}{\delta}\left(\frac{A(t,T)}{A(t,T+\delta)}\exp(-(B(t,T)-B(t,T+\delta))(\alpha(t)+x))-1\right) \\
\partial_x L(t,T,\delta,x) &= \frac{1}{\delta}\frac{A(t,T)}{A(t,T+\delta)}\exp(-(B(t,T)-B(t,T+\delta))(\alpha(t)+x))(B(t,T+\delta)-B(t,T)) \\
&= \frac{(B(t,T+\delta)-B(t,T))}{\delta}(1+\delta L_t(T,\delta))
\end{aligned}$$

We deduce the expressions of the risk-neutral and forward neutral hedging portfolios when using the at par FRA as an hedging instrument:

$$\begin{aligned}
\Delta_t^{f,ATM}(x) &= -\phi(d_+) \quad \text{forward neutral ATM} \\
\Delta_t^{r,ATM}(x) &= -\phi(d_+) + \frac{1-e^{-a(T-t)}}{e^{-a(T-t)}-e^{-a(T+\delta-t)}}\frac{BS}{1+\delta L_{t,T,T+\delta}} \quad \text{risk neutral ATM}
\end{aligned}$$

where we have used the abuse of notations

$$BS = BS\left(1+\delta L_t(T,\delta), 1+\delta K, \sigma\frac{1-e^{-a\delta}}{a}\sqrt{\frac{1-e^{-2a(T-t)}}{2a}}\right).$$

We also give the expression of the delta when hedging with the FRA of fixed rate K:

$$\Delta_t^{r,K}(x_t) = \frac{\phi(d+)\partial_x L - B(t,T+\delta)BS/\delta}{\partial_x L - B(t,T+\delta)(K-L_t(T,\delta))}.$$

As mentioned before the forward-neutral delta is independent of the fixed rate of the FRA used to hedge the option, i.e. we have $\Delta_t^{f,K}(x_t) = \Delta_t^{f,ATM}(x), \forall K$.

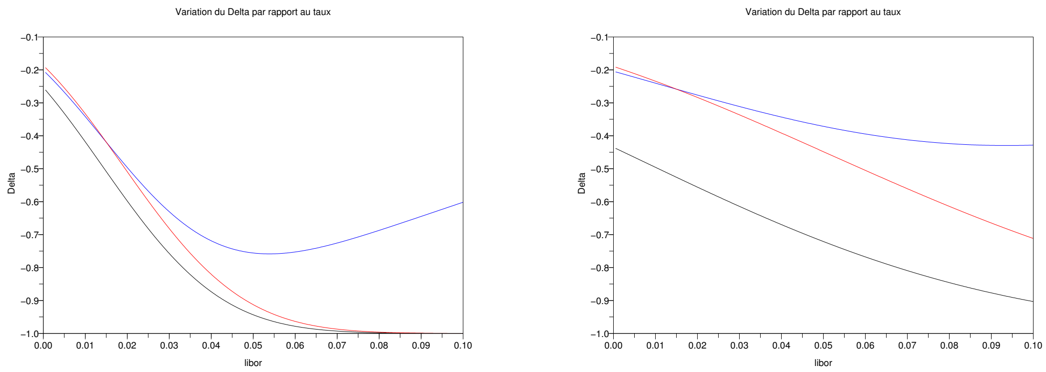

Figure 8.1: Variation of the delta w.r.t. the underlying state variable. Forward-neutral at-the-money in black, risk-neutral at-the-money in blu and risk-neutral hedging with a FRA of fixed rate K in red. The values of the model parameters are $a = 1\%, \quad K = 1.5\%, \quad T = 5, \quad \delta = 1$. The graphic in the left has been obtained with a volatility of 1%. The graphic in the right has been obtained with a volatility of 3%.

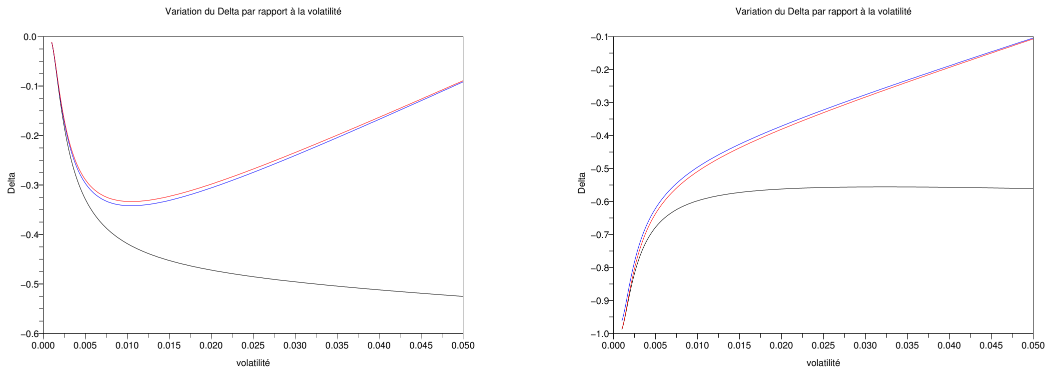

Figure 8.2: Variation of the delta w.r.t. the volatility. Forward-neutral at-the-money in black, risk-neutral at-the-money in blu and risk-neutral hedging with a FRA of fixed rate K in red. The values of the model parameters are $a = 1\%, \quad K = 1.5\%, \quad T = 5, \quad \delta = 1$. The graphic in the left has been obtained with a spot rate of 1%. The graphic in the left has been obtained with a spot rate of 2%.

8.3.2 Numerical comparison of the alternative delta strategies

The delta of the hedging strategies behave differently w.r.t. the movements of the yield curve and the volatility. Figure 8.1 shows the variation of the delta w.r.t. the movements of the underlying state variable of the model X. The values are negative because we defined the FRA as payer of the floating rate and receiver of the fixed rate. In our modeling framework X is the only driver of the movements of the curve. It can be viewed as a parallel shift/long term rates movement of the yield curve. See the discussion in section 2.4. The variation of the delta w.r.t. X is close to be the gamma of the option. We observe that the at-the-money risk-neutral and forward-neutral deltas are significantly different. When rates become very high, the risk-neutral delta tends to decrease. This is because the impact of rates on the risk-neutral delta come from two sources, which have opposite sensitivity to the yield curve. The value of the option is given by the discount factor times the forward value. The forward value of the option is an increasing function of rates, while the discounting factor is a decreasing function of rates. When rates are very high, the discounting factor sensitivity prevails. As a consequence, the risk-neutral hedges will be gamma negative when the option is very deep in the money. Conversely, the variation of the forward-neutral delta has the typical delta profile of a call option on rates, and the gamma is always positive. It is interesting to observe that the risk-neutral delta when hedging with a FRA of fixed rate K (i.e. the fixed rate of the FRA equal to the strike of the Caplet, which is the "natural" hedging instrument of the option) is very similar to the forward-neutral delta.

Figure 8.2 shows the variation of the delta with respect to the volatility. We observe that the forward-neutral and the risk-neutral deltas exhibit a very similar behavior.

The results of this numerical analysis show that the risk-neutral and forward-neutral deltas will be significantly different for deep in the money options.

8.3.3 Hedging performance and cash flows generation

As we have seen there is a significant difference between the forward-neutral delta and the risk-neutral delta. We now compare the hedging strategies in terms of hedging performance. We consider the following indicators:

- The mean and the variance of the "tracking error", i.e. the difference between the payoff of the option and the value of the delta position of the portfolio at expiry.
- The evolution of the MtM of the delta positions in different interest rates scenarios.

We assume that the current market is given by a Vasicek model [Vas77], with long term mean $\theta = \alpha'(t)/k + \alpha(t) = 3\%$ under the risk-neutral measure. We assume that the market risk premium is constant i.e. the change of measure between the objective measure P and the risk-neutral Q is given by $dW^P + \lambda dt = dW^Q$, so that the dynamics of the spot rate under the objective measure is given by the same Ornstein-Uhlenbeck dynamics with a different long term mean $\tilde{\theta}$. The rising scenario is defined as the yield curve dynamics with long term mean $\tilde{\theta} = 5\%$. The decreasing scenario is defined as the yield curve dynamics with long term mean $\tilde{\theta} = 1\%$.
Since the premium of the option is not received up-front by the bank, the hedging portfolio will not include the the cash or zero-coupon position in practice. For this reason, we only consider the delta part of the portfolio and analyze the MtM and the tracking error as if the hedging portfolio was only made of the delta position (which is the case in practice when we

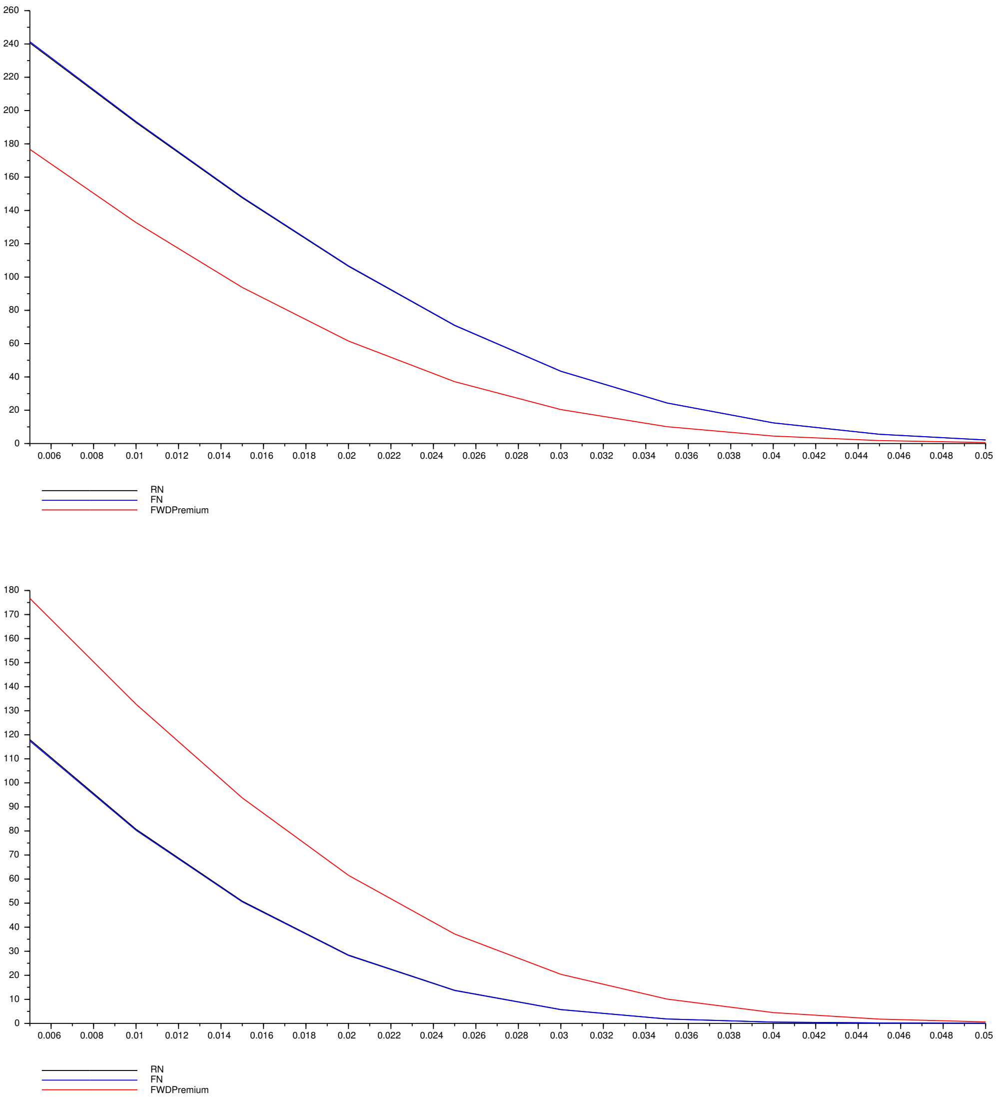

Figure 8.3: Expectation (in bp) of the tracking error i.e. payoff minus delta position as a function of the strike under the historical probability $a = 10\%, \quad T = 2, \quad \delta = 1., \quad r_0 = 2\%$. The graphic in the left corresponds to a scenario with rising interest rates, the graphic in the right corresponds to a scenario with decreasing interest rates. The hedging frequency is 1 month. The red curve represents the forward value of the option.

hedge an interest rate gap).

Figure 8.3 shows the expectation of the tracking error of the different hedging strategies under different interest rates scenarios. The difference between the two hedging strategies in terms of expectation of the tracking error is very small. The difference is slightly larger

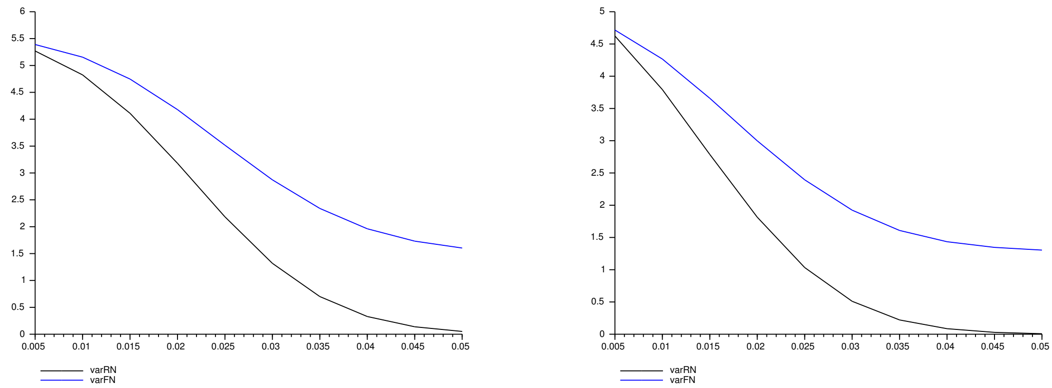

Figure 8.4: Variance (in bp) of the tracking error i.e. payoff minus delta position as a function of the strike under the historical probability $a = 10\%, \quad T = 2, \quad \delta = 1., \quad r_0 = 2\%$. The graphic in the left corresponds to a scenario with rising interest rates, the graphic in the right corresponds to a scenario with decreasing interest rates. The hedging frequency is 1 month.

for out-of-the-money options, but it remains small across all strikes. In the rising interest rates scenario, the expectation of the tracking error of the risk-neutral strategy is smaller than the tracking error of the forward neutral strategy. For both strategies the expectation of the tracking error is significantly higher than the forward premium. In the decreasing scenario, the expectation of the tracking error of the forward neutral strategy is smaller than the expectation of the tracking error of the risk-neutral strategy. For both strategies the expectation of the tracking error is significantly smaller than the forward premium. Looking at the expressions of the P&L of the strategy (8.4) we can see that these differences would be compensated by the term $M_0 e^{\int_0^T dsr_s}$ which will be higher than the term V_0 in (8.3) in the rising interest rates scenario, and smaller in the decreasing interest rates scenario.

Figure 8.4 shows the variance of the tracking error of the different hedging strategies under different interest rates scenarios. We observe that the variance of the tracking error of the forward-neutral strategy is always significantly higher than the variance of the risk-neutral strategy. This is because the zero-coupon position is more volatile than the cash position.

8.4 Conclusion of the chapter

In this chapter we have analyzed the hedging of options in ALM. This allowed us to put in practice the concepts we have introduced in section 2.7. As we have seen the practice of hedging is very different from its theory. In particular is the reality of self-financing portfolios. Hedging is usually performed by only implementing the delta part of the hedging portfolio. The peculiarities of ALM create even more differences from the theory. First the option premium is not perceived at inception of the option, but spread across the cash flows of the corresponding asset or liability. Secondly the only relevant measure in ALM is the actual cash flows generated by the option and its hedges.

We have compared two alternative hedging strategies. The first is the classic risk-neutral hedging strategy, the second is very common in the interest rates market, and consists in replacing the cash position in the risk-neutral strategy by a loan of maturity the expiry of the option (or equivalently a zero-coupon position). We call this the forward-neutral strategy. When we only implement the delta part of the strategy, the forward-neutral strategy allows, in theory, to anticipate the actual cash flows of the hedged position at expiry. This is a desirable property from an ALM perspective, since it allows to exactly predict the future margin generated by the hedged position. The results of theoretical and numerical analysis show that this advantages of the forward-neutral strategy come at the cost of a significantly higher variance of the tracking error of the delta adjusted position.

Bibliography

[AA13] A. Ahdida and A. Alfonsi. Exact and high-order discretization schemes for Wishart processes and their affine extensions. *Ann. Appl. Probab.*, 23(3):1025–1073, 2013.

[AAP14] A. Ahdida, A. Alfonsi, and E. Palidda. Smile with the gaussian term structure model. *preprint*, 2014.

[Ada08] A. Adam. *Handbook of Asset and Liability Management: From Models to Optimal Return Strategies.* The Wiley Finance Series. Wiley, 2008.

[Alf10] A. Alfonsi. High order discretization schemes for the CIR process: application to affine term structure and Heston models. *Math. Comp.*, 79(269):209–237, 2010.

[ALP95] M. Avellaneda, A. Levy, and A. Parás. Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance*, 2:73–88, 1995.

[BBEK08] A. Benabid, H. Bensusan, and N. El Karoui. Wishart Stochastic Volatility: Asymptotic Smile and Numerical Framework. 48 pages, June 2008.

[BCT08] A. Buraschi, A. Cieslak, and F. Trojani. Correlation risk and the term structure of interest rates., 2008.

[BDJ+00] N. Baud, P. Demey, D. Jacomy, G. Riboulet, and T. Roncalli. Plan epargne logement evaluation des options cachées d'un pel. *GRO working paper*, 2000.

[Ben10] H. Bensussan. *Interest rate and longevity risks : dynamic modelling and applications to derivative products and life insurance.* PhD thesis, Ecole Polytechnique, 2010.

[BG12] L. Bergomi and J. Guyon. Stochastic volatilities orderly smiles. *Risk Magazine*, 2012.

[BGH13] F. Biagini, A. Gnoatto, and M. Härtel. Affine HJM Framework on positive matrixes and Long-Term Yield. *ArXiv e-prints*, November 2013.

[BGM97] A. Brace, D. Gatarek, and M. Musiela. The market model of interest rate dynamics. *Math. Finance*, 7(2):127–155, 1997.

[BL13] N. Bauerle and Z. Li. Optimal portfolios for financial markets with wishart volatility. *J. Appl. Probab.*, 50(4):1025–1043, 12 2013.

[BM06] D. Brigo and F. Mercurio. *Interest Rate Models - Theory and Practice: With Smile, Inflation and Credit*. Springer, 2nd edition, August 2006.

[BP14] D. Brigo and A. Pallavicini. Nonlinear consistent valuation of ccp cleared or csa bilateral trades with initial margins under credit, funding and wrong-way risks. *Journal of Financial Engineering*, 01(01):1450001, 2014.

[BPT10] A. Buraschi, P. Porchia, and F. Trojani. Correlation risk and optimal portfolio choice. *The Journal of Finance*, 65(1):393–420, 2010.

[Bru91] M-F. Bru. Wishart processes. *Journal of Theoretical Probability*, 4(4):725–751, 1991.

[BSC+99] J-P. Bouchaud, N. Sagna, R. Cont, N. El-Karoui, and M. Potters. Phenomenology of the interest rate curve. *Applied Mathematical Finance*, 6(3):209–232, 1999.

[BV03] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2003.

[BW00] A. Brace and R. S. Womersley. Exact fit to the swaption volatility matrix using semidefinite programming. *Working paper, ICBI Global Derivatives Conference*, 2000.

[BZ88] E.O. Balitskaya and L.A. Zolotuhina. On the representation of a density by an edgeworth series. *Biometrika*, 75(1):185–187, 1988.

[CD02] R.S. Collin-Dufresne, P.and Goldstein. Do bonds span the fixed income markets? theory and evidence for unspanned stochastic volatility. *Journal of Finance*, 57:1685–1729, 2002.

[CDG02] P. Collin-Dufresne and R.S. Goldstein. Pricing swaptions within an affine framework. *The Journal of Derivatives*, 2002.

[CFMT11] C. Cuchiero, D. Filipović, E. Mayerhofer, and J. Teichmann. Affine processes on positive semidefinite matrices. *Ann. Appl. Probab.*, 21(2):397–463, 2011.

[Che92] O. Cheyette. Multifactor models of the term structure of interest rates. Working paper, Barra Inc., 1992.

[CIR85] J. C. Cox, J.E. Ingersoll, and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

[CM99] P. Carr and D. Madan. Option valuation using fast fourier transform. *Journal of Computational Finance*, 1999.

[Cra57] H. Cramér. *Mathematical Methods of Statistics*. Princeton University Press, 1957.

[Cuc11] C. Cuchiero. Affine and polynomial processes., 2011.

[CW02] P. Carr and L. Wu. Static hedging of standard options, 2002.

[CWZ15] M.C. Chiu, H.Y. Wong, and J. Zhao. Commodity derivatives pricing with cointegration and stochastic covariances. *European Journal of Operational Research*, (0), 2015.

[d'A03] A. d'Aspremont. Interest rate model calibration using semidefinite programming. *CoRR*, cs.CE/0302034, 2003.

[dBEG07] A. d'Aspremont, F. R. Bach, and L. El Ghaoui. Full regularization path for sparse principal component analysis. In *Proceedings of the 24th International Conference on Machine Learning*, ICML '07, pages 177–184, New York, NY, USA, 2007. ACM.

[DC13] R. Douady and S. Crépey. Lois: credit and liquidity. *Risk magazine*, 2013.

[DE94] L. Dieci and T. Eirola. Positive definiteness in the numerical solution of riccati differential equations, 1994.

[DFG11] J. Da Fonseca and M. Grasselli. Riding on the smiles. *Quantitative Finance*, 11(11):1609–1632, 2011.

[DFGI11] J. Da Fonseca, M. Grasselli, and F. Ielpo. Hedging (co) variance risk with variance swaps. *International Journal of Theoretical and Applied Finance*, 14(06):899–943, 2011.

[DFGI14] J. Da Fonseca, M. Grasselli, and F. Ielpo. Estimating the wishart affine stochastic correlation model using the empirical characteristic function. *Studies in Nonlinear Dynamics & Econometrics*, 18(3), 2014.

[DFGT08a] J. Da Fonseca, M. Grasselli, and C. Tebaldi. Option pricing when correlations are stochastic: an analytical framework. *Review of Derivatives Research*, 2008.

[DFGT08b] Jo. Da Fonseca, M. Grasselli, and C. Tebaldi. A multifactor volatility heston model. *Quantitative Finance*, 8(6):591–604, 2008.

[DFR03] P. Demey, A. Frachot, and G. Riboulet. *Introduction à la gestion actif-passif bancaire*. Economica, 2003.

[DFS03] D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. *Ann. Appl. Probab.*, 13(3):984–1053, 2003.

[DK96a] D. Duffie and R. Kan. Multi factor term structure models. *Philosophical Transactions of the Royal Society of London, Series A*, 1996.

[DK96b] D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6:379–406, 1996.

[DS00] Q. Dai and K. J. Singleton. Specification analysis of affine term structure models. *The Journal of Finance*, 55(5):1943–1978, 2000.

[Dup94] B. Dupire. Pricing with a smile. *Risk*, 1994.

[EKD98] N. El Karoui and P. Durand. Interest rates dynamics and option pricing with the quadratic gaussian model in several economies. working paper, 1998.

[EKJPS98] N. El Karoui, M. Jeanblanc-Picquè, and S. E. Shreve. Robustness of the black and scholes formula. *Mathematical finance*, 8(2):93–126, 1998.

[EKL92] N. El Karoui and V. Lacoste. Multifactor models of the term structure of interest rates. preprint University of Paris 6, 1992.

[EKLM+91] N. El Karoui, C. Lepage, R. Myneni, N. Roseau, and R. Wiswanathan. The valuation and hedging of contingent claims with markovian interest rates. preprint University of Paris 6, 1991.

[Fil09] D. Filipović. *Term-structure models*. Springer Finance. Springer-Verlag, Berlin, 2009. A graduate course.

[FPS00] J-P. Fouque, G. Papanicolaou, and K. R. Sircar. *Derivatives in financial markets with stochastic volatility.* Cambridge University Press, Cambridge, 2000.

[FT11] M. Fujii and A. Takahashi. Choice of collateral currency. *Risk Magazine*, 2011.

[Gat06] J. Gatheral. *The Volatility Surface (A practitioner's Guide).* Wiley Finance, 2006.

[GG14a] A. Gnoatto and M. Grasselli. An affine multicurrency model with stochastic volatility and stochastic interest rates. *SIAM J. Financial Math.*, 5(1):493–531, 2014.

[GG14b] A. Gnoatto and M. Grasselli. The explicit Laplace transform for the Wishart process. *J. Appl. Probab.*, 51(3):640–656, 2014.

[GMS10] C. Gourieroux, A. Monfort, and R. Sufana. International money and stock market contingent claims. *Journal of International Money and Finance*, 29(8):1727–1751, December 2010.

[Gno13] A. Gnoatto. The wishart short rate model. Working paper, 2013.

[Gou06] C. Gourieroux. Continuous time wishart process for stochastic risk. *Econometric Reviews*, 25(2-3):177–217, 2006.

[GREK95] H Geman, J-C. Rochet, and N. El Karoui. Changes of numeraire, changes of probability measure and option pricing. *Journal of Applied Probability*, 32:443–458, 1995.

[GS03] C. Gourieroux and R. Sufana. Wishart quadratic term structure models., 2003.

[GS07] C. Gourieroux and R. Sufana. Derivative pricing with multivariate stochastic volatility : Application to credit risk. *Working Paper CREF 04-09*, 2007.

[GT08] M. Grasselli and C. Tebaldi. Solvable affine term structure models. *Mathematical Finance*, 18(1):135–153, 2008.

[HJM92] D. Heath, R. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, 60(1):pp. 77–105, 1992.

[HK79] J. M. Harrison and D. Kreps. Martingales and Arbitrage in Multiperiod securities markets. *Journal of Economic Theory*, 20:381–408, 1979.

[HKLW02] P.S. Hagan, D. Kumar, A. S. Lesniewski, and D.E. Woodward. Managing smile risk. *Wilmott magazine*, pages 84–108, 2002.

[HKP00] P. Hunt, J. Kennedy, and A. Pelsser. Markov-functional interest rate models. *Finance Stoch.*, 4(4):391–408, 2000.

[HP81] J. M. Harrison and S. Pliska. Martingales and Stochastic Integrals in the Theory of Continuos Trading. *Stochastic Processes and Their Applications*, 15:313–316, 1981.

[HW94] J. Hull and A. White. Numerical Procedures for Implementing Term Structure Models I: Single-Factor Models. *Journal of Derivatives*, pages 7–16, 1994.

[Jam97] F. Jamshidian. Libor and swap market models and measures. *Finance and Stochastics*, 1(4):293–330, 1997.

[JY03] R. Jarrow and Y. Yildirim. Pricing treasury inflation protected securities and related derivatives using an hjm model. *Journal of Financial and Quantitative Analysis*, 38:337–358, 6 2003.

[Kno95] H-W. Knobloch. Differential inequalities and maximum principles for second order differential equations. *Nonlinear Anal.*, 25(9-10):1003–1016, 1995.

[Lee04] R. W. Lee. Option pricing by transform methods: Extensions, unification, and error control. *Journal of Computational Finance*, 2004.

[Lev59] J. J. Levin. On the matrix Riccati equation. *Proc. Amer. Math. Soc.*, 10:519–524, 1959.

[LS91] R. B. Litterman and J. Scheinkman. Common Factors Affecting Bond Returns. *The Journal of Fixed Income*, 1:54–61, 1991.

[Mor09] M. Morini. Solving the interest rates puzzle. *SSRN*, 2009.

[Mor11] M. Morini. Funding valuation adjustment: a consistent framework including cva, dva, collateral, netting rules and rehypothecation. *arXiv*, 2011.

[PA10] V. Piterbarg and L. Andersen. *Interest Rate Modeling*, volume Issue 2. Atlantic Financial Press, 2010.

[PB13] A. Pallavicini and D. Brigo. Interest-rate modelling in collateralized markets: Multiple curves, credit-liquidity effects, ccps. *ArXiv e-prints*, April 2013.

[Pit09] V. Piterbarg. Rates squared. *Risk Magazine*, 2009.

[Pit10] V. Piterbarg. Funding beyond discounting: collateral agreements and derivatives pricing. *Risk magazine*, 2010.

[PT11] N. Privault and T.R. Teng. Risk-neutral hedging of interest rate derivatives. Working paper, 2011.

[Rei70] W. T. Reid. Monotoneity properties of solutions of Hermitian Riccati matrix differential equations. *SIAM J. Math. Anal.*, 1(2):195–213, May 1970.

[Rou12] F. Roupin. Programmation semidéfinie: introduction et applications en optimisation combinatoire. Support de cours, 2012.

[Ryd97] T. H. Rydberg. A note on the existence of unique equivalent martingale measures in a markovian setting. *Finance and Stochastics*, 1(3):251–257, 1997.

[Sha86] M. A. Shayman. Phase portrait of the matrix riccati equation. *SIAM J. Control Optim.*, 24(1):1–65, January 1986.

[SP06] D. F. Schrager and A. A. J. Pelsser. Pricing swaptions and coupon bond options in affine term structure models. *Math. Finance*, 16(4):673–694, 2006.

[SU02] K. J. Singleton and L. Umantsev. Pricing coupon-bond options and swaptions in affine term structure models. *Math. Finance*, 12(4):427–446, 2002.

[TS10] A. B. Trolle and E. Schwartz. An empirical analysis of the swaption cube. Working Paper 16549, National Bureau of Economic Research, November 2010.

[TT90] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Analysis and Applications*, 8(4):483–509, 1990.

[TYW05] K. Tanaka, T. Yamada, and T. Watanabe. Approximation of interest rate derivatives prices by gram-charlier expansion and bond moments. *IMES discussion paper*, 2005.

[Vas77] O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177 – 188, 1977.

Appendix A

MF Calibration Algorithm

We describe the calibration algorithm proposed in [HKP00] in the one-dimensional case and taking the terminal ZCB bond as a numéraire. The algorithm strongly lies on hypotheses 1 and 2.

Let us assume we have a product (consider for instance a Bermudan swaption) with settlement dates $T_0, T_1, ..., T_{m-1}$ associated with a swap with fixed leg τ and $M_0, M_1, ..., M_{m-1}$ payment dates. Furthermore, let us assume that all the swaps have the same terminal date $T_m = T_{m-1} + \tau$. We chose the T_m ZCB as the numraire i.e. $N_t = P_{t,T_m}$. Let us denote by $\mathbb{Q}$ the martingale measure associated with the numraire N.

We chose to calibrate our model so that it is consistent with market prices of the swaptions corresponding to the swap involved in our product. We thus assume that for $i = 0, 1, ..., m-1$ we know a continuum of digital swaption prices $V_{0,i}^{mkt}(K)$ associated to the swap $S(t, T_i, M_i, \tau)$.

A.1 "Anti-diagonal" Calibration

step 0: determining $P(T_{m-1}, T_m, \cdot)$

We assumed we know a continuum of prices $V_{0,m-1}^{mkt}(K)$ of caplets on the forward libor rate $L_t(T_{m-1}, T_m - T_{m-1})$.

The arbitrage pricing theory gives

$$V_{0,m-1}^{mkt}(K) = P(0, T_m)\mathbb{E}^{\mathbb{Q}}\left[\mathbf{1}_{L_{T_{m-1}}(T_{m-1}, T_m - T_{m-1}) > K}\right]$$

Since we assumed that the IBOR rates are increasing monotonic functions of the factors there exists an unique $f(K)$ such that

$$L_{T_{m-1}}(T_{m-1}, T_m - T_{m-1}) > K \Leftrightarrow F_{T_m-1} > f(K),$$

were in fact $f(K) = L(T_{m-1}, T_{m-1}, T_m, \cdot)^{-1}(K)$, so that

$$\mathbb{E}^{\mathbb{Q}}\left[\mathbf{1}_{L_{T_{m-1}}(T_{m-1}, T_m - T_{m-1}) > K}\right] = \mathbb{E}^{\mathbb{Q}}\left[\mathbf{1}_{F_{T_{m-1}} > f(K)}\right].$$

Given that we know the distribution of $F_{T_{m-1}}$ under $\mathbb{Q}$ we can compute the right and side of the above equation as a function of $f(K)$. Let us denote by $K(s,t)$ the transition operator of F under $\mathbb{Q}$, i.e. $K(s,t) \circ g(y) = \mathbb{E}^{\mathbb{Q}}[g(F_t)|F_s = y]$. We can then write

$$\mathbb{E}^{\mathbb{Q}}\left[\mathbf{1}_{F_{T_{m-1}} > f(K)}\right] = K(0, T_{m-1}) \circ \mathbf{1}_{\cdot > f(K)}(f_0) = h(f(K)),$$

were f_0 is the initial value of the Markov process F.

We deduce an expression for $f(K)$,

$$f(K) = h^{-1} \circ \left(\frac{V^{mkt}_{0,m-1}(K)}{P(0, T_m)}\right).$$

Knowing $f(K)$ for each K is equivalent of knowing $L(T_{m-1}, T_{m-1}, T_m, \cdot)$, and since $P_{T_{m-1}, T_m} = \frac{1}{1 + \tau L_{T_{m-1}}(T_{m-1}, T_m - T_{m-1})}$ it is equivalent of knowing $P(T_{m-1}, T_m, \cdot)$, precisely we have,

$$P(T_{m-1}, T_m, \cdot) = \frac{1}{1 + \tau\left(\frac{V^{mkt}_{0,m-1}(\cdot)}{P(0,T_m)}\right)^{-1} \circ h}.$$

step 1: $m-1 \longrightarrow m-2$

Straightforward from the martingale property of the discounted asset prices we have,

$$\frac{P_{T_{m-2}, T_{m-1}}}{P_{T_{m-2}, T_m}} = \mathbb{E}^{\mathbb{Q}}\left[\frac{1}{P_{T_{m-1}, T_m}} | \mathcal{F}_{T_{m-2}}\right],$$

so that we can define

$$\frac{P_{T_{m-2}, T_{m-1}}}{P_{T_{m-2}, T_m}} = K(T_{m-2}, T_{m-1}) \circ \left(\frac{1}{P(T_{m-1}, T_m, \cdot)}\right)(F_{T_{m-2}}).$$

Let us now use the information given by the knowledge of the market prices of swaptions. We have

$$\begin{aligned} V^{mkt}_{0,m-2}(K) &= P(0, T_m)\mathbb{E}^{\mathbb{Q}}\left[\frac{A(T_{m-2}, T_{m-2}, M_{m-2}, \tau)}{P_{T_{m-2}, T_m}}\mathbf{1}_{S(T_{m-2}, T_{m-2}, M_{m-2}, \tau) > K}\right] \\ &= P(0, T_m)\tau\mathbb{E}^{\mathbb{Q}}\left[\left(1 + \frac{P_{T_{m-2}, T_{m-1}}}{P_{T_{m-2}, T_m}}\right)\mathbf{1}_{S(T_{m-2}, T_{m-2}, M_{m-2}, \tau) > K}\right] \end{aligned}$$

Since we assumed that the swap rates are increasing monotonic functions of the factors there exists an unique $f(K)$ such that

$$S(T_{m-2}, T_{m-2}, M_{m-2}, \tau) > K \Leftrightarrow F_{T_{m-2}} > f(K),$$

were $f(K) = S(T_{m-2}, T_{m-2}, M_{m-2}, \tau, \cdot)^{-1}(K)$, so that

$$\mathbb{E}^{\mathbb{Q}}\left[\left(1 + \frac{P_{T_{m-2},T_{m-1}}}{P_{T_{m-2},T_m}}\right)\mathbf{1}_{S(T_{m-2},T_{m-2},M_{m-2},\tau)>K}\right] = \mathbb{E}^{\mathbb{Q}}\left[\left(1 + \frac{P_{T_{m-2},T_{m-1}}}{P_{T_{m-2},T_m}}\right)\mathbf{1}_{F_{T_{m-2}}>f(K)}\right]$$

$$\begin{aligned}
&= \mathbb{E}^{\mathbb{Q}}\left[\left(1 + K(T_{m-2}, T_{m-1}) \circ \left(\frac{1}{P(T_{m-1}, T_m, \cdot)}\right)(F_{T_{m-2}})\right)\mathbf{1}_{F_{T_{m-2}}>f(K)}\right] \\
&= K(0, T_{m-2}) \circ \left(1 + K(T_{m-2}, T_{m-1}) \circ \left(\frac{1}{P(T_{m-1}, T_m, \cdot)}\right)(\cdot)\right)\mathbf{1}_{\cdot>f(K)}(f_0) \\
&= h(f(K)).
\end{aligned}$$

We then have

$$h^{-1} \circ \left(\frac{V^{mkt}_{0,m-2}(K)}{P(0, T_m)\tau}\right) = f(K),$$

we deduce

$$S(T_{m-2}, T_{m-2}, M_{m-2}, \tau, \cdot) = \left(\frac{V^{mkt}_{0,m-2}(\cdot)}{P(0, T_m)\tau}\right)^{-1} \circ h$$

and finally

$$\begin{aligned}
P(T_{m-2}, T_m, \cdot) &= \frac{1}{1 + \left(\frac{V^{mkt}_{0,m-2}(\cdot)}{P(0,T_m)\tau}\right)^{-1} \circ h * \frac{A(T_{m-2},T_{m-2},M_{m-2},\tau)}{P_{T_{m-2},T_m}}} \\
&= \frac{1}{1 + \left(\frac{V^{mkt}_{0,m-2}(\cdot)}{P(0,T_m)\tau}\right)^{-1} \circ h * \tau * \left(1 + K(T_{m-2}, T_{m-1}) \circ \left(\frac{1}{P(T_{m-1},T_m,\cdot)}\right)\right)}
\end{aligned}$$

step 2: $m - i \longrightarrow m - (i+1)$

Straightforward from the martingale property of the discounted asset prices we have,

$$\frac{P_{T_{m-(i+1)},T_{m-i}}}{P_{T_{m-(i+1)},T_m}} = \mathbb{E}^{\mathbb{Q}}\left[\frac{1}{P_{T_{m-i},T_m}}|\mathcal{F}_{T_{m-(i+1)}}\right],$$

$$\frac{P_{T_{m-(i+1)},T_{m-i+1}}}{P_{T_{m-(i+1)},T_m}} = \mathbb{E}^{\mathbb{Q}}\left[\frac{1}{P_{T_{m-i+1},T_m}}|\mathcal{F}_{T_{m-(i+1)}}\right],$$

...

$$\frac{P_{T_{m-(i+1)},T_{m-1}}}{P_{T_{m-(i+1)},T_m}} = \mathbb{E}^{\mathbb{Q}}\left[\frac{1}{P_{T_{m-1},T_m}}|\mathcal{F}_{T_{m-(i+1)}}\right].$$

so that we can define

$$\frac{P_{T_{m-(i+1)},T_{m-i}}}{P_{T_{m-(i+1)},T_m}} = K(T_{m-(i+1)},T_{m-i}) \circ \left(\frac{1}{P(T_{m-i},T_m,\cdot)}\right)(F_{T_{m-(i+1)}}),$$

$$\frac{P_{T_{m-(i+1)},T_{m-i+1}}}{P_{T_{m-(i+1)},T_m}} = K(T_{m-(i+1)},T_{m-i+1}) \circ \left(\frac{1}{P(T_{m-i+1},T_m,\cdot)}\right)(F_{T_{m-(i+1)}}),$$

...

$$\frac{P_{T_{m-(i+1)},T_{m-1}}}{P_{T_{m-(i+1)},T_m}} = K(T_{m-(i+1)},T_{m-1}) \circ \left(\frac{1}{P(T_{m-1},T_m,\cdot)}\right)(F_{T_{m-(i+1)}}).$$

Let us now use the information given by the knowledge of the market prices of swaptions. We have

$$\frac{V^{mkt}_{0,m-(i+1)}(K)}{P(0,T_m)} =$$

$$\begin{aligned} &= \mathbb{E}^{\mathbb{Q}}\left[\frac{A(T_{m-(i+1)},T_{m-(i+1)},M_{m-(i+1)},\tau)}{P_{T_{m-(i+1)},T_m}}\mathbf{1}_{S(T_{m-(i+1)},T_{m-(i+1)},M_{m-(i+1)},\tau)>K}\right] \\ &= \mathbb{E}^{\mathbb{Q}}\left[\tau\left(1+\frac{P_{T_{m-(i+1)},T_{m-i}}+\cdots+P_{T_{m-(i+1)},T_{m-1}}}{P_{T_{m-(i+1)},T_m}}\right)\mathbf{1}_{S(T_{m-(i+1)},T_{m-(i+1)},M_{m-(i+1)},\tau)>K}\right] \end{aligned}$$

Since we assumed that the swap rates are increasing monotonic functions of the factors there exists an unique $f(K)$ such that

$$S(T_{m-(i+1)},T_{m-(i+1)},M_{m-(i+1)},\tau) > K \Leftrightarrow F_{T_{m-(i+1)}} > f(K),$$

were $f(K) = S(T_{m-(i+1)},T_{m-(i+1)},M_{m-(i+1)},\tau,\cdot)^{-1}(K)$, so that

$$\mathbb{E}^{\mathbb{Q}}\left[\tau\left(1+\frac{P_{T_{m-(i+1)},T_{m-i}}+\cdots+P_{T_{m-(i+1)},T_{m-1}}}{P_{T_{m-(i+1)},T_m}}\right)\mathbf{1}_{S(T_{m-(i+1)},T_{m-(i+1)},M_{m-2},\tau)>K}\right] =$$

$$
\begin{aligned}
&= \mathbb{E}^{\mathbb{Q}}\left[\tau\left(1+\frac{P_{T_{m-(i+1)},T_{m-i}}+\cdots+P_{T_{m-(i+1)},T_{m-1}}}{P_{T_{m-(i+1)},T_m}}\right)\mathbf{1}_{F_{T_{m-(i+1)}}>f(K)}\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\tilde{A}(T_{m-(i+1)},T_{m-(i+1)},M_{m-(i+1)},\tau,X_{T_{m-(i+1)}})\mathbf{1}_{F_{T_{m-(i+1)}}>f(K)}\right] \\
&= K(0,T_{m-2})\circ\left(\tilde{A}(T_{m-(i+1)},T_{m-(i+1)},M_{m-(i+1)},\tau,\cdot)\mathbf{1}_{\cdot>f(K)}\right)(f_0) \\
&= h(f(K)),
\end{aligned}
$$

Where we have defined

$$
\tilde{A}(T_{m-(i+1)},T_{m-(i+1)},M_{m-(i+1)},\tau,\cdot)
$$

$$
=\tau\left(1+K(T_{m-(i+1)},T_{m-i})\circ\left(\frac{1}{P(T_{m-i},T_m,\cdot)}\right)+\cdots+K(T_{m-(i+1)},T_{m-1})\circ\left(\frac{1}{P(T_{m-1},T_m,\cdot)}\right)\right).
$$

We then have

$$
h^{-1}\circ\left(\frac{V^{mkt}_{0,m-(i+1)}(K)}{P(0,T_m)}\right)=f(K),
$$

we deduce

$$
S(T_{m-(i+1)},T_{m-(i+1)},M_{m-(i+1)},\tau,\cdot)=\left(\frac{V^{mkt}_{0,m-(i+1)}(\cdot)}{P(0,T_m)}\right)^{-1}\circ h
$$

and finally

$$
P(T_{m-(i+1)},T_m,\cdot)=\frac{1}{1+\left(\left(\frac{V^{mkt}_{0,m-(i+1)}(\cdot)}{P(0,T_m)}\right)^{-1}\circ h*\tilde{A}(T_{m-(i+1)},T_{m-(i+1)},M_{m-(i+1)},\tau,\cdot)\right)}. \tag{A.1}
$$

A.2 "Column" Calibration

In this section we show that, following a procedure similar to the one presented in the above section, the model can be calibrated to a column of caplets.

We take assumptions 1 and 2, we furthermore assume that we know a continuum of digital caplets prices $C^{mkt}_{0,i}(K)$ for $i=0,...,m-1$, associated to the period τ and of maturity T_i.

The algorithm

We determine the functional form $P(T_{m-1}, T_m, \cdot)$ following the exact procedure presented in the **step 0** of the above section and replacing $V^{mkt}_{0,m-1}$ by $C^{mkt}_{0,m-1}$.

We then have

$$P(T_{m-1}, T_m, \cdot) = \frac{1}{1 + \tau \left(\frac{C^{mkt}_{0,m-1}(\cdot)}{P(0,T_m)}\right)^{-1} \circ h}.$$

Let us now assume that we have determined the functional forms $P(T_k, T_m, \cdot)$ for $k = m-1, ..., i$ let us determine $P(T_{m-(i+1)}, T_m, \cdot)$.

step: $m - i \longrightarrow m - (i+1)$

Straightforward from the martingale property of the discounted asset prices we have,

$$\frac{P_{T_{m-(i+1)},T_{m-i}}}{P_{T_{m-(i+1)},T_m}} = \mathbb{E}^{\mathbb{Q}}\left[\frac{1}{P_{T_{m-i},T_m}} | \mathcal{F}_{T_{m-(i+1)}}\right],$$

so that we can define

$$\frac{P_{T_{m-(i+1)},T_{m-i}}}{P_{T_{m-(i+1)},T_m}} = K(T_{m-(i+1)}, T_{m-i}) \circ \left(\frac{1}{P(T_{m-i}, T_m, \cdot)}\right)(F_{T_{m-(i+1)}}).$$

We furthermore use the information on market prices,

$$\frac{C^{mkt}_{m-(i+1),0}(K)}{P(0,T_m)} =$$

$$\begin{aligned}
&= \mathbb{E}^{\mathbb{Q}}\left[\frac{P_{T_{m-(i+1)},T_{m-i}}}{P_{T_{m-(i+1)},T_m}} \mathbf{1}_{L_{T_{m-(i+1)}}(T_{m-(i+1)}, T_{m-i} - T_{m-(i+1)}) > K}\right], \\
&= \mathbb{E}^{\mathbb{Q}}\left[K(T_{m-(i+1)}, T_{m-i}) \circ \left(\frac{1}{P(T_{m-i}, T_m, \cdot)}\right)(F_{T_{m-(i+1)}}) \mathbf{1}_{L_{T_{m-(i+1)}}(T_{m-(i+1)}, T_{m-i} - T_{m-(i+1)}) > K}\right]
\end{aligned}$$

Since we assumed that IBOR rates are increasing monotonic functions of the factors there exists an unique $y(K)$ such that

$$L(T_{m-(i+1)}, T_{m-(i+1)}, T_{m-i}, X_{T_{m-(i+1)}}) > K \Leftrightarrow F_{T_{m-(i+1)}} > y(K),$$

were $y(K) = L(T_{m-(i+1)}, T_{m-(i+1)}, T_{m-i}, \cdot)^{-1}(K)$, so that

$$\mathbb{E}^{\mathbb{Q}}\left[K(T_{m-(i+1)}, T_{m-i}) \circ \left(\frac{1}{P(T_{m-i}, T_m, \cdot)}\right)(F_{T_{m-(i+1)}}) \mathbf{1}_{L_{T_{m-(i+1)}}(T_{m-(i+1)}, T_{m-i} - T_{m-(i+1)}) > K}\right] =$$

$$
\begin{aligned}
&= \mathbb{E}^{\mathbb{Q}}\left[K(T_{m-(i+1)}, T_{m-i}) \circ \left(\frac{1}{P(T_{m-i}, T_m, \cdot)}\right)(F_{T_{m-(i+1)}})\mathbf{1}_{F_{T_{m-(i+1)}}>y(K)}\right] \\
&= h(y(K))
\end{aligned}
$$

We then have

$$
h^{-1} \circ \left(\frac{C^{mkt}_{m-(i+1),0}(K)}{P(0,T_m)}\right) = y(K),
$$

we deduce

$$
L(T_{m-(i+1)}, T_{m-(i+1)}, T_{m-i}, \cdot) = \left(\frac{C^{mkt}_{m-(i+1),0}(\cdot)}{P(0,T_m)}\right)^{-1} \circ h,
$$

and finally we have

$$
P(T_{m-(i+1)}, T_m, \cdot) = \frac{1}{K(T_{m-(i+1)}, T_{m-i}) \circ \left(\frac{1}{P(T_{m-i},T_m,\cdot)}\right)(\cdot)\left(1 + \tau * \left(\frac{C^{mkt}_{m-(i+1),0}(\cdot)}{P(0,T_m)}\right)^{-1} \circ h\right)}. \tag{A.2}
$$

Appendix B

Proof of lemma 46

Lemma 46 — *Let $(\mathcal{F}_t)_{t\geq 0}$ denote the filtration generated by $((W_t, Z_t), t \geq 0)$. We consider a process $(Y_t, X_t)_{t\geq 0}$ valued in $\mathbb{R}^p \times \mathcal{S}_d^+(\mathbb{R})$, and we assume that there exist continuous $(\mathcal{F}_t)$-adapted processes $(A_t)_{t\geq 0}, (B_t)_{t\geq 0}, (C_t)_{t\geq 0}, (D_t)_{t\geq 0}$ and $(E_t)_{t\geq 0}$ that are respectively valued in $\mathbb{R}^p$, $\mathcal{M}_{p\times d}(\mathbb{R})$, $\mathcal{S}_d^+(\mathbb{R}), \mathcal{M}_d(\mathbb{R})$ and $\mathcal{M}_d(\mathbb{R})$ so that (Y_t, X_t) admits the following semimartingale decomposition:*

$$\begin{aligned} dY_t &= A_t dt + B_t\left(dW_t\rho + \bar{\rho} dZ_t\right) \\ dX_t &= C_t dt + D_t dW_t E_t + E_t^T dW_t^T D_t^T. \end{aligned} \tag{B.1}$$

Then, for $k, l \in \{1, ..., p\}, i, j, r, s \in \{1, ..., d\}$, the quadratic covariation of $(Y_t)_k, (Y_t)_l$, $(X_t)_{ij}$ and $(X_t)_{rs}$ is given by:

$$d < Y_k, Y_l >_t \;=\; (B_t B_t^T)_{kl} \tag{B.2}$$

$$\begin{aligned} d < X_{ij}, X_{rs} >_t &= (D_t D_t^T)_{is}(E_t E_t^T)_{jr} + (D_t D_t^T)_{ir}(E_t E_t^T)_{js} \\ &+ (D_t D_t^T)_{js}(E_t E_t^T)_{ir} + (D_t D_t^T)_{jr}(E_t E_t^T)_{is} \end{aligned} \tag{B.3}$$

$$d < Y_k, X_{rs} >_t \;=\; (DB^T)_{rk}(E^T\rho)_s + (DB^T)_{sk}(E^T\rho)_r. \tag{B.4}$$

Proof : The proof is a simple application of the results of Da Fonseca, Grasselli and Tebaldi [DFGT08a]. Since $W_t\rho + \bar{\rho} Z_t$ is a standard Brownian motion, we easily get that

$$d < Y_k, Y_l >_t = (BB^T)_{kl}.$$

On the other hand, we have

$$\begin{aligned} dY_k &= \sum_{l,m=1}^{d} B_{kl} dW_{lm}\rho_m + \bar{\rho}\sum_{l=1}^{d} B_{kl} dZ_l \\ dX_{ij} &= \sum_{l,m=1}^{d} D_{il} dW_{lm} E_{mj} + E_{li} dW_{ml} D_{jm} \end{aligned}$$

we thus get the following quadratic variation

$$\begin{aligned}
d < Y_k, X_{ij} > &= < \sum_{l,m=1}^{d} B_{kl} dW_{lm} \rho_m, \sum_{l,m=1}^{d} D_{il} dW_{lm} E_{mj} + E_{li} dW_{ml} D_{jm} > \\
&= \sum_{l,m=1}^{d} B_{kl} \rho_m < dW_{lm}, \sum_{l,m=1}^{d} D_{il} dW_{lm} E_{mj} + E_{li} dW_{ml} D_{jm} > \\
&= \sum_{l,m=1}^{d} B_{kl} \rho_m (D_{il} E_{mj} + E_{mi} D_{jl}) \\
&= \left(\sum_{l=1}^{d} B_{kl} D_{il} \sum_{m=1}^{d} E_{mj} \rho_m \right) + \left(\sum_{l=1}^{d} B_{kl} D_{jl} \sum_{m=1}^{d} E_{mi} \rho_m \right) \\
&= (DB^T)_{ik} (E^T \rho)_j + (DB^T)_{jk} (E^T \rho)_i
\end{aligned}$$

□

Appendix C

Black-Scholes and Bachelier prices and greeks

In this section we provide the formulas for the prices of call options when assuming that the underlying asset follows a log-normal and a Normal dynamics. We also provide the expression of the derivatives of the price (greeks) w.r.t. the log price in the log-normal case and the price in the Normal case.

C.1 Log-normal model

We provide here the expression of the Black-Scholes call price as a function of the underlying log-price and of the cumulative variance between the pricing and the maturity dates. We consider a fixed strike K and want to calculate $\partial_l^k \mathrm{BS}(h,v)$, where

$$\begin{aligned}\mathrm{BS}(h,v) &= \mathbb{E}\left[\left(\exp\left(h-\frac{1}{2}v+\sqrt{v}G\right)-K\right)^+\right], \quad G\sim N(0,1)\\ &= e^h\Phi(d_+)-K\Phi(d_-).\end{aligned}$$

Here, Φ denotes the cumulative distribution function of the normal distribution and $d_\pm = \frac{h-\log(K)\pm v/2}{\sqrt{v}}$. We also denote $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ and have $\varphi'(x)=-x\varphi(x)$. Thus, we have

$$k\in\mathbb{N}^*,\ \partial_h^k\mathrm{BS}(h,v,K)=e^h\left(\Phi(d_+)+P_k(d_+)\varphi(d_+)\right),$$

with $P_1(x)=0$ and $P_{k+1}(d_+)=P_k(d_+)+\frac{1}{\sqrt{v}}+\frac{1}{\sqrt{v}}P_k'(d_+)-\frac{d_+}{\sqrt{v}}P_k(d_+)$. In particular, we have $P_2(d_+)=\frac{1}{\sqrt{v}}$, $P_3(d_+)=\frac{1}{\sqrt{v}}\left(2-\frac{d_+}{\sqrt{v}}\right)$, $P_4(d_+)=\frac{3}{\sqrt{v}}-3\frac{d_+}{v}-\frac{1-d_+^2}{v\sqrt{v}}$.

C.2 Normal model

We provide here the expression of the Normal or Bachelier call price as a function of the underlying asset price and of the cumulative variance between the pricing and the maturity dates.

$$\begin{aligned} \mathrm{BH}(s,v) &= \mathbb{E}\left[\left(s+\sqrt{v}G-K\right)^+\right], \quad G \sim N(0,1) \\ &= (s-K)\Phi(d) + v\varphi(d), \end{aligned} \tag{C.1}$$

where Φ and φ are respectively the cumulative distribution function and density of the Normal distribution and $d = \frac{s-K}{\sqrt{v}}$.We provide the derivatives up to the order which appear in the expression of the price expansion:

$$\begin{aligned} &\partial_s \mathrm{BH}(s,v) = \Phi(d), && \partial_s^2 \mathrm{BH}(s,v) = \varphi(d)/\sqrt{v}, \\ &\partial_s^3 \mathrm{BH}(s,v) = -d\varphi(d)/v, && \partial_s^4 \mathrm{BH}(s,v) = -(1-d^2)\varphi(d)/v\sqrt{v}, \\ &\partial_s^5 \mathrm{BH}(s,v) = d(3-d^2)\varphi(d)/v^2, && \partial_s^6 \mathrm{BH}(s,v) = (3-6d^2+d^4)\varphi(d)/v^2\sqrt{v}. \end{aligned}$$

Appendix D

Gamma-Vega relationships

In this section we provide some useful relationships between the gamma and the vega of standard european options when assuming that the underlying asset price follows a log-normal process or a normal process. This relationship plays a key role in the derivation of the asymptotic expansions for caplets and swaptions provided in paragraphs 4.3.1 and 4.3.2. In the following we consider an abstract market with an asset S. We assume for simplicity that interest rates are 0.

D.1 Log-normal model

Let us assume that the asset S follows a Black-Scholes dynamics, under the risk-neutral measure we have

$$\frac{dS}{S} = \sigma dW. \tag{D.1}$$

Consider an european option with payoff $f(S_T)$ at maturity T. The price of the option $P(t, L_t, \sigma^2)$, as a function of the log price $L_t = \log(S_t)$ is solution of the following PDE

$$\partial_t P(t, l, \sigma^2) + \sigma^2(\partial_l^2 - \partial_l)P(t, l, \sigma^2) = 0, \quad P(T, l, \sigma^2) = f(e^l). \tag{D.2}$$

By the time-homogeneity property of the Black-Scholes dynamics, we notice that the price depends on t only through the time to maturity $\tau = T - t$. By a slight abuse of notations we will thus denote equivalently $P(t, l, \sigma^2)$ or $P(\tau, l, \sigma^2)$. From this equation, we see that the price verifies a certain scale invariance property, we have that $P(\tau/\alpha, l, \sigma^2\alpha) = P(\tau, l, \sigma^2), \alpha > 0$. By deriving the expression $P(\tau/\alpha, l, \sigma^2\alpha)$ as a function of α we must have:

$$\frac{-\tau}{\alpha^2}\partial_\tau P + \sigma^2 \partial_{\sigma^2} P = 0. \tag{D.3}$$

Now noticing that $\partial_t P = -\partial_\tau P$ and plugging (D.2) into (D.7) we obtain the gamma-vega relationship:

$$\partial_{\sigma^2} P = \frac{\tau}{2}\left(\partial_l^2 - \partial_l\right) P. \tag{D.4}$$

D.2 Normal model

We now follow the same path as in the previous paragraph. We assume here that the underlying asset follows a normal dynamics, under the risk-neutral measure we have

$$dS = \sigma dW \tag{D.5}$$

Consider an european option with payoff $f(S_T)$ at maturity T. The price of the option $P(t, S_t, \sigma^2)$ is solution of the following PDE

$$\partial_t P(t, S, \sigma^2) + \sigma^2 \partial_S^2 P(t, s, \sigma^2) = 0, \quad P(T, s, \sigma^2) = f(s). \tag{D.6}$$

By the time-homogeneity property of the normal dynamics, we notice that the price depends on t only through the time to maturity $\tau = T - t$. By a slight abuse of notations we will thus denote equivalently $P(t, s, \sigma^2)$ or $P(\tau, s, \sigma^2)$. From this equation, we see that the price verifies a certain scale invariance property, we have that $P(\tau/\alpha, s, \sigma^2\alpha) = P(\tau, s, \sigma^2), \alpha > 0$. By deriving the expression $P(\tau/\alpha, s, \sigma^2\alpha)$ as a function of α we must have:

$$\frac{-\tau}{\alpha^2}\partial_\tau P + \sigma^2 \partial_{\sigma^2} P = 0. \tag{D.7}$$

Now noticing that $\partial_t P = -\partial_\tau P$ and plugging (D.6) into (D.7) we obtain the gamma-vega relationship:

$$\partial_{\sigma^2} P = \frac{\tau}{2}\partial_s^2 P. \tag{D.8}$$

Appendix E

Expansion of the price and volatility: details of calculations

E.1 Caplets price expansion

Let us calculate $\tilde{L}_1(s)P_0(s,x,h)$. We denote by $\partial_x P_0$ the symmetric matrix $(\partial_x P_0)_{i,j} = \partial_{x_{i,j}} P_0$. From (4.19), we get

$$\tilde{L}_1(s)P_0 = 2(\Delta B^\top cx\Delta D_0 I_d^n \rho)(\partial_h^2 - \partial_h)P_0 + 2B^\top cx\partial_x P_0 I_d^n \rho + 2\Delta B^\top cx\partial_h\partial_x P_0 I_d^n \rho.$$

We now observe from (4.21) that

$$\partial_{x_{i,j}} P_0(s,x,h) = \frac{1}{2}\partial_{x_{i,j}} v(s,T,\delta,x)(\partial_h^2 - \partial_h)P_0(s,x,h). \tag{E.1}$$

We denote by $\partial_x v$ the symmetric matrix $(\partial_x v)_{i,j} = \partial_{x_{i,j}} v$. Since v is linear with respect to x, $\partial_x v$ doesnot depend on x and we have

$$\begin{aligned}(\partial_x v(s,T,\delta))_{i,j} &= \int_s^T \Delta B^\top(u,T,\delta)ce^{b(u-s)}e_{i,j}e^{b^\top(u-s)}c^\top \Delta B(u,T,\delta)du \\ &= \int_s^T [e^{b^\top(u-s)}c^\top \Delta B(u,T,\delta)]_i[e^{b^\top(u-s)}c^\top \Delta B(u,T,\delta)]_j du,\end{aligned}$$

where $e_{i,j}$ is the matrix defined by $(e_{i,j})_{k,l} = \mathbf{1}_{i=k,j=l}$. We now use that ∂_h and $\tilde{L}_0(t)$ commute, which gives that for any $k \in \mathbb{N}$, $\partial_t\partial_h^k P_0 + \tilde{L}_0(t)\partial_h^k P_0 = 0$. Thus, $P_0(s,X_s^0,H_t^0)$ is a martingale and we have

$$\mathbb{E}\left[\partial_h^k P_0(s,X_s^0,H_s^0)|H_t = h, X_t^0 = x\right] = \partial_h^k P_0(t,x,h). \tag{E.2}$$

From (4.23), we obtain (4.25) with

$$c_1(t,T,\delta,x) = \int_t^T \Delta B^\top(s,T,\delta)cX_{s-t}^0(x)\partial_x v(s,T,\delta)I_d^n \rho ds, \tag{E.3}$$

$$c_2(t,T,\delta,x) = \int_t^T \Delta B^\top(s,T,\delta)cX_{s-t}^0(x)\Delta D_0(s,T,\delta)I_d^n\rho + B^\top(T+\delta-s)cX_{s-t}^0(x)\partial_x v(s,T,\delta)I_d^n\rho ds, \tag{E.4}$$

where

$$X_s^0(x) = e^{bs}\left(x + \int_0^s e^{-bu}\Omega e^{-b^\top u} du\right) e^{b^\top s} \tag{E.5}$$

is the solution of $dX_t = (\Omega + bX_t + X_t b^\top)dt$ starting from x.

We now calculate similarly $\tilde{L}_1(s)P_1(s,x,h)$. To do so, we first calculate from (4.25)

$$\begin{aligned}\partial_{x_{i,j}} P_1(s,x,h) = &\big[\partial_{x_{i,j}} c_1(s,T,\delta,x)(\partial_h^3 - \partial_h^2) + \partial_{x_{i,j}} c_2(s,T,\delta,x)(\partial_h^2 - \partial_h)\\ &+ \frac{1}{2}\partial_{x_{i,j}} v(s,T,\delta) c_1(s,T,\delta,x)(\partial_h^3 - \partial_h^2)(\partial_h^2 - \partial_h)\\ &+ \frac{1}{2}\partial_{x_{i,j}} v(s,T,\delta) c_2(s,T,\delta,x)(\partial_h^2 - \partial_h)^2\big] P_0(s,x,h).\end{aligned}$$

We then have

$$\begin{aligned}\tilde{L}_1(s)P_1 &= (\Delta B^\top cx \Delta D_0 I_d^n \rho)(\partial_h^2 - \partial_h)P_1 + 2B^\top cx \partial_x P_1 I_d^n \rho + 2\Delta B^\top cx \partial_h \partial_x P_1 I_d^n \rho\\ &= \Bigg[(\Delta B^\top cx \Delta D_0 I_d^n \rho)[c_1(\partial_h^2 - \partial_h)^2 \partial_h + c_2(\partial_h^2 - \partial_h)^2] + 2B^\top cx \partial_x c_1 I_d^n \rho(\partial_h^2 - \partial_h)\partial_h\\ &+ 2B^\top cx \partial_x c_2 I_d^n \rho(\partial_h^2 - \partial_h) + c_1 B^\top cx \partial_x v I_d^n \rho(\partial_h^2 - \partial_h)^2 \partial_h + c_2 B^\top cx \partial_x v I_d^n \rho(\partial_h^2 - \partial_h)^2\\ &+ 2\Delta B^\top cx \partial_x c_1 I_d^n \rho(\partial_h^2 - \partial_h)\partial_h^2 + 2\Delta B^\top cx \partial_x c_2 I_d^n \rho(\partial_h^2 - \partial_h)\partial_h\\ &+ c_1 \Delta B^\top cx \partial_x v I_d^n \rho(\partial_h^2 - \partial_h)^2 \partial_h^2 + c_2 \Delta B^\top cx \partial_x v I_d^n \rho(\partial_h^2 - \partial_h)^2 \partial_h\Bigg] P_0(s,x,h).\end{aligned}$$

Let us observe that c_1 and c_2 are again linear with respect to x, and their derivatives with respect to x do not depend on x. More precisely, we have:

$$\begin{aligned}\partial_{x_{i,j}} c_1(t,T,\delta,x) &= \int_t^T \Delta B^\top(s,T,\delta) c e^{-b(s-t)} e_{i,j} e^{-b^\top(s-t)} \partial_x v(s,T,\delta) I_d^n \rho ds,\\ \partial_{x_{i,j}} c_2(t,T,\delta,x) &= \int_t^T \Delta B^\top(s,T,\delta) c e^{-b(s-t)} e_{i,j} e^{-b^\top(s-t)} \Delta D_0(s,T,\delta) I_d^n \rho\\ &\quad + B^\top(T+\delta-s) c e^{-b(s-t)} e_{i,j} e^{-b^\top(s-t)} \partial_x v(s,T,\delta) I_d^n \rho ds,\end{aligned}$$

Then, by using again (E.2), we obtain

$$\begin{aligned}&\mathbb{E}\left[\int_t^T \tilde{L}_1(s)P_1(s,X_s^0,H_s^0)ds | H_t = h, X_t^0 = x\right] = \Bigg[e_1(t,T,\delta,x)(\partial_h^2 - \partial_h)^2\partial_h^2 + e_2(t,T,\delta,x)(\partial_h^2 - \partial_h)^2\partial_h\\ &+ e_3(t,T,\delta,x)(\partial_h^2 - \partial_h)^2 + e_4(t,T,\delta,x)(\partial_h^2 - \partial_h)\partial_h^2 + e_5(t,T,\delta,x)(\partial_h^2 - \partial_h)\partial_h\\ &+ e_6(t,T,\delta,x)(\partial_h^2 - \partial_h)\Bigg] P_0(t,x,h),\end{aligned}$$

with

$$
\begin{aligned}
e_1(t,T,\delta,x) &= \int_t^T c_1(s,T,\delta,x)\Delta B^\top(s,T,\delta)cX^0_{s-t}(x)\partial_x v(s,T,\delta)I^n_d\rho ds,\\
e_2(t,T,\delta,x) &= \int_t^T c_1(s,T,\delta,x)[(\Delta B^\top(s,T,\delta)cX^0_{s-t}(x)\Delta D_0(s,T,\delta)I^n_d\rho)\\
&\quad + B^\top(T+\delta-s)cX^0_{s-t}(x)\partial_x v(s,T,\delta)I^n_d\rho]\\
&\quad + c_2(s,T,\delta,x)\Delta B^\top(s,T,\delta)cX^0_{s-t}(x)\partial_x v(s,T,\delta)I^n_d\rho ds,\\
e_3(t,T,\delta,x) &= \int_t^T c_2(s,T,\delta,x)[(\Delta B^\top(s,T,\delta)cX^0_{s-t}(x)\Delta D_0(s,T,\delta)I^n_d\rho)\\
&\quad + B^\top(T+\delta-s)cX^0_{s-t}(x)\partial_x v(s,T,\delta)I^n_d\rho]ds,\\
e_4(t,T,\delta,x) &= \int_t^T 2\Delta B^\top(s,T,\delta)cX^0_{s-t}(x)\partial_x c_1(s,T,\delta)I^n_d\rho ds,\\
e_5(t,T,\delta,x) &= \int_t^T 2B^\top(T+\delta-s)cX^0_{s-t}(x)\partial_x c_1(s,T,\delta)I^n_d\rho + 2\Delta B^\top(s,T,\delta)cX^0_{s-t}(x)\partial_x c_2(s,T,\delta)I^n_d\rho ds,\\
e_6(t,T,\delta,x) &= \int_t^T 2B^\top(T+\delta-s)cX^0_{s-t}(x)\partial_x c_2(s,T,\delta)I^n_d\rho ds.
\end{aligned}
$$

Last, we calculate $\tilde{L}_2(s)P_0(s,x,h)$ and get

$$
\begin{aligned}
\tilde{L}_2(s)P_0 = \Bigg[&\left(2\mathrm{Tr}(\Delta D_0 I^n_d \Delta D_0 x) + (\Delta B^\top cx\Delta D_1 I^n_d\rho)\right)(\partial_h^2 - \partial_h)\\
&+ \frac{1}{2}\mathrm{Tr}\left[((d-1)I^n_d + 4xD_0 I^n_d)\partial_x v\right](\partial_h^2 - \partial_h)\\
&+ \frac{1}{2}\mathrm{Tr}\left[I^n_d \partial_x v x \partial_x v\right](\partial_h^2 - \partial_h)^2 + 2\mathrm{Tr}\left[\Delta D_0 x\partial_x v I^n_d\right](\partial_h^2 - \partial_h)\partial_h\Bigg] P_0.
\end{aligned}
$$

We obtain by using (E.2)

$$
\begin{aligned}
\mathbb{E}\left[\int_t^T \tilde{L}_2(s)P_0(s,X^0_s,H^0_s)ds | H_t = h, X^0_t = x\right] = \Bigg[&d_1(t,T,\delta,x)(\partial_h^2 - \partial_h)^2\\
&+ d_2(t,T,\delta,x)(\partial_h^2 - \partial_h)\partial_h\\
&+ d_3(t,T,\delta,x)(\partial_h^2 - \partial_h)\Bigg] P_0(t,x,h),
\end{aligned}
$$

with

$$d_1(t,T,\delta,x) = \int_t^T \frac{1}{2}\mathrm{Tr}\left[I_d^n \partial_x v(s,T,\delta) X_{s-t}^0(x) \partial_x v(s,T,\delta)\right] ds$$

$$d_2(t,T,\delta,x) = \int_t^T 2\mathrm{Tr}\left[\Delta D_0(s,T,\delta) X_{s-t}^0(x) \partial_x v(s,T,\delta) I_d^n\right] ds$$

$$\begin{aligned} d_3(t,T,\delta,x) = \int_t^T & \big(2\mathrm{Tr}(\Delta D_0(s,T,\delta) I_d^n \Delta D_0(s,T,\delta) X_{s-t}^0(x)) + (\Delta B^\top(s,T,\delta) c X_{s-t}^0(x) \Delta D_1(s,T,\delta) I_d^n \rho)\big) \\ & + \frac{1}{2}\mathrm{Tr}\left[((d-1)I_d^n + 4X_{s-t}^0(x) D_0(T+\delta-s) I_d^n)\partial_x v(s,T,\delta)\right] ds. \end{aligned}$$

E.2 Swaption price expansion

Let us first calculate $\tilde{L}_1^S(u) P_0^S(u,x,s)$. We denote by $\partial_x P_0^S$ the symmetric matrix $(\partial_x P_0^S)_{i,j} = \partial_{x_{i,j}} P_0^S$, and we get from (4.39):

$$\tilde{L}_1^S(u) P_0^S = B^S(u)^\top c x D_0^S(u) I_d^n \rho \partial_s^2 P_0^S + 2B^A(u)^\top c x \partial_x P_0^S I_d^n \rho + 2B^S(u)^\top c x \partial_x \partial_s P_0^S I_d^n \rho.$$

From (D.8), we have

$$\partial_{x_{i,j}} P_0^S(u,x,s) = \frac{1}{2}\partial_{x_{i,j}} v^S(u,T,x) \partial_s^2 P_0^S(u,x,s).$$

Similarly, we denote by $\partial_x v^S$ the symmetric matrix made with the partial derivatives of v^S. Since $v^S(s,T,x)$ is linear with respect to x, $\partial_x v^S$ does not depend on x and we have

$$\partial_x v^S(u,T) = \int_u^T B^S(r)^\top c e^{b(r-u)} e_{i,j} e^{b^\top(r-u)} c^\top B^S(r) dr.$$

Since $\tilde{L}_0^S$ and ∂_s commute, we have for any $k \in \mathbb{N}$, $\partial_t \partial_s^k P_0^S + \tilde{L}_0^S(u) \partial_s^k P_0^S = 0$ for $u \in (0,T)$ and thus

$$t \le u \le T,\ \mathbb{E}[\partial_s^k P_0^S(u, X_u^0, S_u^0) | S_t^0 = s, X_t^0 = x] = \partial_s^k P_0^S(u,x,x).$$

We therefore obtain (4.44) with

$$c_1^S(t,T,x) = \int_t^T B^S(u)^\top c X_{u-t}^0(x) \partial_x v^S(u,T) I_d^n \rho du,$$

$$c_2^S(t,T,x) = \int_t^T B^S(u)^\top c X_{u-t}^0(x) D_0^S(u) I_d^n \rho + B^A(u)^\top c X_{u-t}^0(x) \partial_x v^S(u,T) I_d^n \rho du,$$

and $X_{u-t}^0(x)$ defined by (E.5).

Similarly, we calculate $\tilde{L}_1^S(u) P_1^S(u,x,s)$. We have

$$\begin{aligned} \partial_{x_{i,j}} P_1^S(u,x,s) = & \Big[\partial_{x_{i,j}} c_1^S(u,T,x) \partial_s^3 + \partial_{x_{i,j}} c_2^S(u,T,x) \partial_s^2 \\ & + \frac{1}{2}\partial_{x_{i,j}} v^S(u,T) c_1^S(u,T,x) \partial_s^5 + \frac{1}{2}\partial_{x_{i,j}} v^S(u,T) c_2^S(u,T,x) \partial_s^4 \Big] P_0^S(u,x,s). \end{aligned}$$

We note that the derivatives

$$\partial_{x_{i,j}} c_1^S(t,T,x) = \int_t^T B^S(u)^\top c e^{b(u-t)} e_{i,j} e^{b^\top(u-t)} \partial_x v^S(u,T) I_d^n \rho du,$$
$$\partial_{x_{i,j}} c_2^S(t,T,x) = \int_t^T B^S(u)^\top c e^{b(u-t)} e_{i,j} e^{b^\top(u-t)} D_0^S(u) I_d^n \rho + B^A(u)^\top c e^{b(u-t)} e_{i,j} e^{b^\top(u-t)} \partial_x v^S(u,T) I_d^n \rho du$$

do not depend on x. We now obtain

$$\begin{aligned}\tilde{L}_1^S(u)P_1^S(u,x,s) = \Big[&(B^S(u)^\top c x D_0^S(u) I_d^n \rho)[c_1^S(u,T,x)\partial_s^5 + c_2^S(u,T,x)\partial_s^4] + 2B^A(u)^\top c x \partial_x c_1^S(u,T) I_d^n \rho \partial_s^3 \\ &+ 2B^A(u)^\top c x \partial_x c_2^S(u,T) I_d^n \rho \partial_s^2 + c_1^S(u,T,x) B^A(u)^\top c x \partial_x v^S(u,T) I_d^n \rho \partial_s^5 \\ &+ c_2^S(u,T,x) B^A(u)^\top c x \partial_x v^S(u,T) I_d^n \rho \partial_s^4 + 2B^S(u)^\top c x \partial_x c_1^S(u,T) I_d^n \rho \partial_s^4 \\ &+ 2B^S(u)^\top c x \partial_x c_2^S(u,T) I_d^n \rho \partial_s^3 + c_1^S(u,T,x) B^S(u)^\top c x \partial_x v^S(u,T) I_d^n \rho \partial_s^6 \\ &+ c_2^S(u,T,x) B^S(u)^\top c x \partial_x v^S(u,T) I_d^n \rho \partial_s^5 \Big] P_0^S(u,x,s),\end{aligned}$$

and thus

$$\begin{aligned}\mathbb{E}\left[\int_t^T \tilde{L}_1^S(u) P_1^S(u, X_u^0, S_u^0) du \middle| S_t^0 = s, X_t^0 = x\right] = \Big[&e_1^S(t,T,x)\partial_h^6 + e_2^S(t,T,x)\partial_h^5 + e_3^S(t,T,x)\partial_h^4 \\ &+ e_4^S(t,T,x)\partial_h^3 + e_5^S(t,T,x)\partial_h^2 \Big] P_0^S(t,x,s),\end{aligned}$$

with

$$\begin{aligned} e_1^S(t,T,x) &= \int_t^T c_1^S(u,T,x)(B^S(u))^\top c X_{u-t}^0(x) \partial_x v^S(u,T) I_d^n \rho du, \\ e_2^S(t,T,x) &= \int_t^T c_1^S(u,T,x)\left[B^S(u)^\top c X_{u-t}^0(x) D_0^S(u) I_d^n \rho + B^A(u)^\top c X_{u-t}^0(x) \partial_x v^S(u,T) I_d^n \rho\right] \\ &\quad + c_2^S(u,T,x) B^S(u)^\top c X_{u-t}^0(x) \partial_x v^S(u,T) I_d^n \rho du, \\ e_3^S(t,T,x) &= \int_t^T c_2^S(u,T,x)\left[B^S(u)^\top c X_{u-t}^0(x) D_0^S(u) I_d^n \rho + B^A(u)^\top c X_{u-t}^0(x) \partial_x v^S(u,T) I_d^n \rho\right] \\ &\quad + 2B^S(u)^\top c X_{u-t}^0(x) \partial_x c_1^S(u,T) I_d^n \rho du, \\ e_4^S(t,T,x) &= \int_t^T 2B^A(u)^\top c X_{u-t}^0(x) \partial_x c_1^S(u,T) I_d^n \rho + 2B^S(u)^\top c X_{u-t}^0(x) \partial_x c_2^S(u,T) I_d^n \rho du, \\ e_5^S(t,T,x) &= \int_t^T 2B^A(u)^\top c X_{u-t}^0(x) \partial_x c_2^S(u,T) I_d^n \rho du. \end{aligned}$$

We finally calculate from (4.40)

$$\tilde{L}_2^S(u)P_0^S(u,x,s) = \Bigg[[2\mathrm{Tr}(D_0^S(u)I_d^n D_0^S(u)x) + (B^S(u))^\top cxD_1^S(u)I_d^n\rho]\partial_s^2 + \mathrm{Tr}\left([2xD_0^A(u)I_d^n + \frac{1}{2}(d-1)I_d^n]\partial_x v^S(u,T)\right)\partial_s^2 + \frac{1}{2}\mathrm{Tr}(I_d^n\partial_x v^S(u,T)x\partial_x v^S(u,T))\partial_s^4 + 2\mathrm{Tr}(D_0^S(u)x\partial_x v^S(u,T)I_s^n)\partial_s^3 \Bigg] P_0^S(u,x,s),$$

and get

$$\mathbb{E}\left[\int_t^T \tilde{L}_2^S(u)P_0^S(u,X_u^0,S_u^0)du \middle| S_t^0 = s, X_t^0 = x\right] = \Big[d_1^S(t,T,x)\partial_h^4 + d_2^S(t,T,x)\partial_h^3 + d_3^S(t,T,x)\partial_h^2 \Big] P_0^S(t,x,s),$$

with

$$\begin{aligned}
d_1^S(t,T,x) &= \int_t^T \frac{1}{2}\mathrm{Tr}(I_d^n\partial_x v^S(u,T)X_{u-t}^0(x)\partial_x v^S(u,T))du,\\
d_2^S(t,T,x) &= \int_t^T 2\mathrm{Tr}(D_0^S(u)X_{u-t}^0(x)\partial_x v^S(u,T)I_s^n)du,\\
d_3^S(t,T,x) &= \int_t^T [2\mathrm{Tr}(D_0^S(u)I_d^n D_0^S(u)X_{u-t}^0(x)) + (B^S(u))^\top cX_{u-t}^0(x)D_1^S(u)I_d^n\rho]\\
&\quad + \mathrm{Tr}\left([2X_{u-t}^0(x)D_0^A(u)I_d^n + \frac{1}{2}(d-1)I_d^n]\partial_x v^S(u,T)\right)du.
\end{aligned}$$