Polynomial models in finance

THÈSE N^{O} 7902 (2017)

PRÉSENTÉE LE 15 DÉCEMBRE 2017
AU COLLÈGE DU MANAGEMENT DE LA TECHNOLOGIE
CHAIRE SWISSQUOTE EN FINANCE QUANTITATIVE
PROGRAMME DOCTORAL EN FINANCE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Damien ACKERER

acceptée sur proposition du jury:

Prof. P. Collin Dufresne, président du jury
Prof. D. Filipovic, directeur de thèse
Prof. M. Larsson, rapporteur
Dr A. Pallavicini, rapporteur
Prof. J. Hugonnier, rapporteur

EPFL
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2017

Acknowledgements

I express my deepest gratitude to my supervisor, Damir Filipović, for his guidance and support. This thesis and my doctoral education greatly benefited from his insights and research experience. I thank the members of my thesis committee: Martin Larsson for his regular advice and feedback; Andrea Pallavicini for his comments and sharing his industry experience; Julien Hugonnier for the many helpful discussions and his mentoring; and Erwan Morellec for his encouragement to pursue a PhD degree and accepting to be the president of my jury. The research leading to this thesis has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 307465-POLYTE.

My gratitude is extended to my co-authors: Sergio Pulido for his positive attitude and perseverance in solving problems; and Thibault Vatter for his fresh mindset and the many breaks at Satellite. I thank Agostino Capponi for the opportunity to visit Columbia University and Ioannis Karatzas for his kind welcome. I am indebted to Lorraine and Sophie for all the administrative help and holiday advice.

The completion of a doctorate is a long journey during which a lot happens. I was fortunate to be surrounded by smart colleagues and good friends who made this experience truly enriching and joyful, thank you all. I would like in particular to express my appreciation to Benjamin, Boris, David, Jérémy, Julien, Romain, Thomas, Vincent, Yassine, and Yuan with whom I shared memorable moments in and out of the office.

I am thankful to my family and in-laws for their unconditional support and love. Last, my heartfelt appreciation goes to Alevtina for filling every single day with color and melody.

Lausanne, August 2017 D. A.

Abstract

This thesis presents new flexible dynamic stochastic models for the evolution of market prices and new methods for the valuation of derivatives. These models and methods build on the recently characterized class of polynomial jump-diffusion processes for which the conditional moments are analytic.

The first half of this thesis is concerned with modelling the fluctuations in the volatility of stock prices, and with the valuation of options on the stock. A new stochastic volatility model for which the squared volatility follows a Jacobi process is presented in the first chapter. The stock price volatility is allowed to continuously fluctuate between a lower and an upper bound, and option prices have closed-form series representations when their payoff functions depend on the stock price at finitely many dates. Truncating these series at some finite order entails accurate option price approximations. This method builds on the series expansion of the ratio between the log price density and an auxiliary density, with respect to an orthonormal basis of polynomials in a weighted Lebesgue space. When the payoff functions can be similarly expanded, the method is particularly efficient computationally. In the second chapter, more flexible choices of weighted spaces are studied in order to obtain new series representations for option prices with faster convergence rates. The option price approximation method can then be applied to various stochastic volatility models.

The second half of this thesis is concerned with modelling the default times of firms, and with the pricing of credit risk securities. A new class of credit risk models in which the firm default probability is linear in the factors is presented in the third chapter. The prices of defaultable bonds and credit default swaps have explicit linear-rational expressions in the factors. A polynomial model with compact support and bounded default intensities is developed. This property is exploited to approximate credit derivatives prices by interpolating their payoff functions with polynomials. In the fourth chapter, the joint term structure of default probabilities is flexibly modelled using factor copulas. A generic static framework is developed in which the prices of high dimensional and complex credit securities can be efficiently and exactly computed. Dynamic credit risk models with significant default dependence can in turn be constructed by combining polynomial factor copulas and linear credit risk models.

Key words: polynomial model, stochastic volatility, option pricing, credit risk, correlated defaults, credit derivatives

Résumé

Cette thèse présente des nouveaux modèles stochastiques, dynamiques, et flexibles pour l'évolution des prix du marché et des nouvelles méthodes de valorisation des produits dérivés. Ces modèles et méthodes s'appuient sur la classe récemment caractérisée de processus de diffusion avec sauts polynomial pour lesquels les moments conditionnels sont analytiques.

La première moitié de cette thèse se porte sur la modélisation des fluctuations de la volatilité des cours des actions et sur l'évaluation des prix des options sur actions. Un nouveau modèle de volatilité stochastique où le carré de la volatilité suit un processus de type Jacobi est présenté dans le premier chapitre. La volatilité des cours d'une action est autorisée à fluctuer continuellement entre une limite inférieure et une limite supérieure, et les prix des options ont des représentations en série de forme fermée lorsque leurs résultats à échéance dépendent du cours de l'action à un nombre fini de dates. En tronquant ces séries des approximations précises des prix des options sont obtenues. Cette méthode s'appuie sur le développement en série du rapport de vraisemblance entre la densité de probabilité du logarithm du cours de l'action et une mesure auxiliaire, par rapport à une base orthonormée de polynômes dans un espace pondéré par cette mesure auxiliaire. Quand le developpement en série de la fonction de résultat à échance peut aussi être effectuer, cette méthode est particulièrement efficace computationnellement. Dans le deuxième chapitre, des choix plus souple pour cet espace pondéré sont étudiés et de nouvelles représentations en série sont obtenues avec des taux de convergence plus rapides. Cette méthode d'approximation des prix des options peut être appliquée à divers modèles de volatilité stochastique.

La deuxième moitié de cette thèse se porte sur la modélisation des défaillances des entreprises et sur l'évaluation des dérivés sur évenements de crédit. Une nouvelle classe de modèles de risque de crédit pour laquelle la probabilité de défaut d'une entreprise est linéaire en les facteurs est présentée dans le troisième chapitre. Les prix des obligations et des couvertures de défaillance ont des expressions explicites données par des fonctions linéaire-rationelles en les facteurs. Un modèle polynomial plus particulier est étudié pour lequel les facteurs prennent leurs valeurs dans un hypercube et les taux de défaillance sont bornés. Cette propriété est exploitée pour approximer les prix des dérivés de crédit en interpolant leurs fonctions de paiement à échéance avec des polynômes. Dans le quatrième chapitre, les copules à facteurs sont proposés pour modéliser de manière flexible la structure par échéance des probabilités de défaillance commune. Un cadre statique générique est développé dans lequel les prix de titres de crédit complexes sur de nombreuses entreprises peuvent être calculés efficacement et exactement. Des modèles dynamiques avec une dépendance importante entre les défaults peuvent être construits en combinant des copules à facteurs polynomiales avec des modèles de crédit linéaires.

Mots clefs : modèle polynomial, volatilité stochastique, évaluation du prix des options, risque de crédit, corrélation entre défauts de paiement, produits de crédit dérivés

Contents

Acknowledgements i

Abstract (English/Français) iii

List of Figures ix

List of Tables xi

Introduction 1

1 The Jacobi Stochastic Volatility Model 5
- 1.1 Introduction . . . 5
- 1.2 Model Specification . . . 8
- 1.3 European Option Pricing . . . 11
- 1.4 Exotic Option Pricing . . . 17
- 1.5 Numerical Analysis . . . 21
- 1.6 Hermite Moments . . . 24
- 1.7 Proofs . . . 27
- 1.8 Conclusion . . . 35

2 Option Pricing with Orthogonal Polynomial Expansions 37
- 2.1 Introduction . . . 37
- 2.2 Polynomial Price Series Expansions . . . 38
- 2.3 Polynomial Stochastic Volatility Models . . . 44
- 2.4 Numerical Analysis . . . 49
- 2.5 Basis Construction with Moments . . . 56
- 2.6 Proofs . . . 57
- 2.7 Conclusion . . . 61

3 Linear Credit Risk Models 63
- 3.1 Introduction . . . 63
- 3.2 The Linear Framework . . . 66
- 3.3 The Linear Hypercube Model . . . 74
- 3.4 Extensions . . . 79
- 3.5 Case Studies . . . 83

3.6 Chebyshev Interpolation . . . 97
3.7 Market Price of Risk Specifications . . . 98
3.8 Proofs . . . 99
3.9 Conclusion . . . 106

4 Dependent Defaults and Losses with Factor Copula Models 107
4.1 Introduction . . . 107
4.2 The Factor Copula Framework . . . 110
4.3 Discrete Loss Distributions . . . 116
4.4 Numerical Analysis . . . 123
4.5 Empirical Analysis . . . 128
4.6 Standard Copula Models . . . 133
4.7 Pricing Formulas . . . 135
4.8 Proofs . . . 137
4.9 Conclusion . . . 143

Bibliography 145

Curriculum Vitae 153

List of Figures

1.1 Variance and correlation 9
1.2 European call option. 21
1.3 Implied volatility smile: from Heston to Black–Scholes. 23
1.4 Computational performance. 23
1.5 Forward start and Asian options. 24

2.1 Auxiliary densities and implied volatility convergence. 50
2.2 Fitted implied volatility surfaces for the Heston and Jacobi models. 53

3.1 State space of the LHC model with a single factor. 75
3.2 Comparison of the one-factor LHC and CIR models. 77
3.3 CDS spreads data. 85
3.4 CDS spreads fits and errors. 90
3.5 Factors fitted from CDS spreads. 91
3.6 Payoffs and prices approximations of CDS options. 94
3.7 CDS option price sensitivities. 95
3.8 CDS index option payoff function and Chebyshev error bound. 97

4.1 Computation performance. 124
4.2 Defaults dependence and copula mixture. 125
4.3 Total number of defaults with copula mixture. 125
4.4 Multi-name credit derivatives losses. 126
4.5 Loss distribution and loss amounts dependence. 127
4.6 Number of defaults and loss dependence. 128
4.7 Upfronts on CDX.NA.IG.21 tranches. 130
4.8 Models calibration to tranches on the CDX.NA.IG.21. 131
4.9 Parameters calibrated on CDX.NA.IG.21 tranches. 132
4.10 Diagnostic of models calibrated on CDX.NA.IG.21 tranches. 133

List of Tables

1.1 Implied volatility values and errors. 22

2.1 Implied volatility errors for the Jacobi model. 51
2.2 Fitted parameters for the Heston and Jacobi models. 52
2.3 Implied volatility errors for the Stein-Stein model. 55

3.1 CDS spreads summary statistics. 84
3.2 Fitted parameters for the LHC models. 88
3.3 Comparison of CDS spreads fits for the LHC models. 89

4.1 Tranches structure on the CDX.NA.IG.21. 129
4.2 Summary statistics for the tranches on the CDX.NA.IG.21. 129
4.3 Archimedean copulas 136

Introduction

The development of adequate models for the dynamical evolution of market prices and interest rates is of strategic importance for the financial and insurance industries. Mathematical models are nowadays routinely applied to price securities, hedge derivatives, optimize portfolios, invest pension funds, and evaluate capital requirements. This thesis in Mathematical finance presents new classes of flexible and tractable stochastic dynamic models for equity volatility risk and fixed income credit risk, as well as fast and accurate methods to approximate the prices of derivatives whose values are exposed to these risks.

Mathematical finance, also known as Quantitative finance, is the branch of applied Mathematics concerned with financial markets and the relative valuation of financial securities. A financial mathematician will typically take the share price of a company as given and use it to value derivatives of the stock. Derivatives prices are given by the discounted expected value of their payoffs taken under a risk-neutral measure, an alternative measure equivalent to the real-world measure. In practice, market imperfections and investor preferences are reflected in the risk premiums that connect the real-world measure to the risk-neutral measures as implied from market prices. Therefore, the quality of the computed derivatives prices and hedging strategies crucially depends the accuracy of the models that describe the dynamical behavior of the underlying assets, as well as the risk premiums. A particular interest therefore rises for tractable models capable of parsimoniously reproducing the risk factors dynamics, and for which efficient numerical solutions exist to price derivatives.

Affine models have been the most extensively studied and applied class of continuous stochastic dynamical models in the past twenty years. An affine model is built upon an affine jump-diffusion process which is a particular type of semimartingale. The key property of affine processes is that their characteristic functions are exponential-linear in the initial process value where the coefficients are given by the solution of Riccati equations that can numerically be solved efficiently. This enables the use of Fourier transform techniques to value, for examples, European call and put options with stochastic stock price volatility, and bonds with stochastic interest rates and credit spreads. However, the tractability that affine processes exhibit comes at the cost of constrained dynamics. They may therefore face difficulties in parsimoniously reproducing realistic asset price dynamics.

Polynomial models strictly extend affine models, and are still at the beginning of their development. A polynomial model is built upon a polynomial jump-diffusion process. Formally, the extended generator of a polynomial process maps any polynomial to a polynomial of the same or lower degree. As a consequence, polynomial processes admit closed form conditional moments which are given by the explicit solutions of linear systems of differential equations. These solutions are polynomials evaluated at the initial process value where the coefficients are given by the action of a matrix exponential on a vector that can efficiently be calculated using modern computing capabilities and algorithms. Polynomial processes offer new modeling possibilities because their dynamics is strictly more flexible than those of affine processes. However, their characteristic functions cannot be retrieved explicitly and different methods must be developed to price financial assets and derivatives.

This thesis presents new stochastic volatility and credit risk polynomial models, and new option pricing techniques using polynomial approximations and moments.

Chapter 1 presents a model for the stock price in which the volatility of stock returns is allowed to fluctuate over time between a lower and an upper bound. These fluctuations significantly affect the distribution of returns and, as a consequence, impact derivatives prices. In this model the joint density of any finite sequence of log returns admits a series representation in terms of the Hermite polynomials and the Gaussian distribution. Closed-form series representations are in turn derived for option prices whose discounted payoffs functions depend on the asset price trajectory at finitely many dates. This includes European call, put, and digital options, forward start options, and can be applied to discretely monitored Asian options. A numerical analysis shows that the price approximations obtained by truncating the option price series at a finite order become accurate within short CPU time. This chapter is based on (Ackerer, Filipović, and Pulido 2016).

Chapter 2 refines the option price series representations developed in the first chapter. This method builds on the series expansion of the ratio between the log price density and some auxiliary density, with respect to an orthonormal basis of polynomials in a weighted Lebesgue space. The particular choice of auxiliary density is essential to obtain efficient and accurate option price approximations. The set of tractable auxiliary density is enlarged from a Gaussian density to mixture distribution whose components can belong to various parametric density families. This leads to option price series representations with significantly faster convergence rates. A natural candidate of auxiliary density for univariate diffusive stochastic volatility models is the finite Gaussian mixture. Theoretical and numerical arguments show that fast and precise option price approximations can be obtained for many stochastic volatility models. This chapter contains material in preparation for submission.

Chapter 3 introduces a class of credit risk models in which the firm default probability is linear in the factors. The prices of defaultable bonds and credit default swaps are given by linear-rational expressions in the factors with explicit coefficients. A polynomial model with compact support is developed for which default intensities are bounded. The price of a CDS option can in turn be uniformly approximated by polynomials in the factors. Multi-name models with simultaneous defaults, positively and negatively correlated default intensities, and stochastic interest rates can be constructed. An empirical study illustrates the versatility of these models by fitting CDS spread time series. A numerical analysis validates the efficiency of the option price approximation method. This chapter is based on (Ackerer and Filipović 2016).

Chapter 4 describes a static framework to model dependent defaults for a large number of firms using factor copulas. This is a different approach from the third chapter in which the default intensities are correlated but the default times are virtually not. The framework is generic, as it nests many standard models, and enables parsimonious constructions of high dimensional models. The prices of complex credit derivatives such as collateral debt obligation (CDO) and CDO squared can be easily computed because portfolio losses mass functions can be exactly recovered. This allows the design of models consistently pricing various complex instruments. Dynamic stochastic credit risk models with substantial default dependence can in turn be constructed by combining polynomial factor copulas and linear credit risk models. This chapter is based on (Ackerer and Vatter 2016).

Statement of Originality

I certify that the content of this thesis is my own work, where some parts are the result of collaborations with my thesis supervisor Prof. Damir Filipović, as well as my co-authors Prof. Sergio Pulido and Dr. Thibault Vatter. No other person's work has been used without due acknowledgement.

1 The Jacobi Stochastic Volatility Model

In this chapter we introduce a novel stochastic volatility model where the squared volatility of the asset return follows a Jacobi process. It contains the Heston model as a limit case. We show that the joint density of any finite sequence of log returns admits a Gram–Charlier A expansion with closed-form coefficients. We derive closed-form series representations for option prices whose discounted payoffs are functions of the asset price trajectory at finitely many time points. This includes European call, put, and digital options, forward start options, and can be applied to discretely monitored Asian options. In a numerical analysis we find that the price approximations become accurate within short CPU time.

1.1 Introduction

Stochastic volatility models for asset returns are popular among practitioners and academics because they can generate implied volatility surfaces that match option price data to a great extent. They resolve the shortcomings of the Black–Scholes model (Black and Scholes 1973), where the return has constant volatility. Among the the most widely used stochastic volatility models is the Heston model (Heston 1993), where the squared volatility of the return follows an affine square-root diffusion. European call and put option prices in the Heston model can be computed using Fourier transform techniques, which have their numerical strengths and limitations; see for instance (Carr and Madan 1999), (Bakshi and Madan 2000), (Duffie, Filipović, and Schachermayer 2003), (Fang and Oosterlee 2009), and (Chen and Joslin 2012).

In this chapter we introduce a novel stochastic volatility model, henceforth the Jacobi model, where the squared volatility V_t of the log price X_t follows a Jacobi process with values in some compact interval $[v_{min}, v_{max}]$. As a consequence, Black–Scholes implied volatilities are bounded from below and above by $\sqrt{v_{min}}$ and $\sqrt{v_{max}}$. The Jacobi model (V_t, X_t) belongs to the class of polynomial diffusions studied in (Eriksson and Pistorius 2011), (Cuchiero, Keller-Ressel, and Teichmann 2012), and (Filipović and Larsson 2016). It includes the Black–Scholes model as a special case and converges weakly in the path space to the Heston model for $v_{max} \to \infty$ and $v_{min} = 0$.

We show that the log price X_T has a density g that admits a Gram–Charlier A series expansion with respect to any Gaussian density w with sufficiently large variance. More specifically, the likelihood ratio function $\ell = g/w$ lies in the weighted space L^2_w of square-integrable functions with respect to w. Hence it can be expanded as a generalized Fourier series with respect to the corresponding orthonormal basis of Hermite polynomials $H_0, H_1, \ldots$. Boundedness of V_t is essential, as the Gram–Charlier A series of g does not converge for the Heston model.

The Fourier coefficients ℓ_n of ℓ are given by the Hermite moments of X_T, $\ell_n = \mathbb{E}[H_n(X_T)]$. Due to the polynomial property of (V_t, X_t) the Hermite moments admit easy to compute closed-form expressions. This renders the Jacobi model extremely useful for option pricing. Indeed, the price π_f of a European option with discounted payoff $f(X_T)$ for some function f in L^2_w is given by the L^2_w-scalar product $\pi_f = \langle f, \ell \rangle_w = \sum_{n\geq 0} f_n \ell_n$. The Fourier coefficients f_n of f are given in closed-form for many important examples, including European call, put, and digital options. We approximate π_f by truncating the price series at some finite order N and derive truncation error bounds.

We extend our approach to price exotic options whose discounted payoff $f(Y)$ depends on a finite sequence of log returns $Y = (X_{t_1} - X_0, X_{t_2} - X_{t_1}, \ldots, X_{t_d} - X_{t_{d-1}})$. As in the univariate case we derive the Gram–Charlier A series expansion of the density g of Y with respect to a properly chosen multivariate Gaussian density w. Assuming that f lies in L^2_w the option price π_f is obtained as a series representation of the L^2_w-scalar product in terms of the Fourier coefficients of f and of the likelihood ratio function $\ell = g/w$ given by the corresponding Hermite moments of Y. Due to the polynomial property of (V_t, X_t) the Hermite moments admit closed-form expressions, which can be efficiently computed. The Fourier coefficients of f are given in closed-form for various examples, including forward start options and forward start options on the underlying return.

Consequently, the pricing of these options is extremely efficient and does not require any numerical integration. Even when the Fourier coefficients of the discounted payoff function f are not available in closed-form, e.g. for Asian options, prices can be approximated by integrating f with respect to the Gram–Charlier A density approximation of g. This boils down to a numerically feasible integration with respect to the underlying Gaussian density w. In a numerical analysis we find that the price approximations become accurate within short CPU time. This is in contrast to the Heston model, for which the pricing of exotic options using Fourier transform techniques is cumbersome and creates numerical difficulties as reported in (Kruse and Nögel 2005), (Kahl and Jäckel 2005), and (Albrecher, Mayer, Schoutens, and Tistaert 2006). In view of this, the Jacobi model also provides a viable alternative to approximate option prices in the Heston model.

The Jacobi process, also known as Wright–Fisher diffusion, was originally used to model gene frequencies; see for instance (Karlin and Taylor 1981) and (Ethier and Kurtz 1986). More recently, the Jacobi process has also been used to model financial factors. For example, (Delbaen and Shirakawa 2002) model interest rates by the Jacobi process and study moment-based techniques for pricing bonds. In their framework, bond prices admit a series representation in terms of Jacobi polynomials. These polynomials constitute an orthonormal basis of eigenfunctions of the infinitesimal generator and the stationary beta distribution of the Jacobi process; additional properties of the Jacobi process can be found in (Mazet 1997) and (Demni and Zani 2009). The multivariate Jacobi process has been studied in (Gourieroux and Jasiak 2006) where the authors suggest it to model smooth regime shifts and give an example of stochastic volatility model without leverage effect. The Jacobi process has been also applied recently to model stochastic correlation matrices in (Ahdida and Alfonsi 2013) and credit default swap indexes in (Bernis and Scotti 2017).

Density series expansion approaches to option pricing were pioneered by (Jarrow and Rudd 1982). They propose expansions of option prices that can be interpreted as corrections to the pricing biases of the Black–Scholes formula. They study density expansions for the law of underlying prices, not the log returns, and express them in terms of cumulants. Evidently, since convergence cannot be guaranteed in general, their study is based on strong assumptions that imply convergence. In subsequent work, (Corrado and Su 1996) and (Corrado and Su 1997) study Gram–Charlier A expansions of 4^{th} order for options on the S&P 500 index. These expansions contain skewness and kurtosis adjustments to option prices and implied volatility with respect to the Black–Scholes formula. The skewness and kurtosis correction terms, which depend on the cumulants of 3^{rd} and 4^{th} order, are estimated from data. Due to the instability of the estimation procedure, higher order expansions are not studied. Similar studies on the biases of the Black–Scholes formula using Gram–Charlier A expansions include (Backus, Foresi, and Wu 2004) and (Li and Melnikov 2012). More recently, (Drimus, Necula, and Farkas 2013) and (Necula, Drimus, and Farkas 2015) study related expansions with physicist Hermite polynomials instead of probabilist Hermite polynomials. In order to guarantee the convergence of the Gram–Charlier A expansion for a general class of diffusions, (Ait-Sahalia 2002) develop a technique based on a suitable change of measure. As pointed out in (Filipović, Mayerhofer, and Schneider 2013), in the affine and polynomial settings this change of measure usually destroys the polynomial property and the ability to calculate moments efficiently. More recently a similar study has been carried out by (Xiu 2014). Gram–Charlier A expansions, under a change of measure, are also mentioned in the work of (Madan and Milne 1994), and the subsequent studies of (Longstaff 1995), (Abken, Madan, and Ramamurtie 1996) and (Brenner and Eom 1997), where they use these moment expansions to test the martingale property with financial data and hence the validity of a given model.

The research in this chapter is similar to (Filipović, Mayerhofer, and Schneider 2013) that provides a generic framework to perform density expansions using orthonormal polynomial basis in weighted L^2 spaces for affine models. They show that a bilateral Gamma density weight works for the Heston model. However, that expansion is numerically more cumbersome than the Gram–Charlier A expansion because the orthonormal basis of polynomials has to be constructed using Gram–Schmidt orthogonalization. In a related paper (Heston and Rossi 2016) study polynomial expansions of prices in the Heston, Hull-White and Variance Gamma models using logistic weight functions.

The remainder of the chapter is as follows. In Section 1.2 we introduce the Jacobi stochastic volatility model. In Section 1.3 we derive European option prices based on the Gram–Charlier A series expansion. In Section 1.4 we extend this to the multivariate case, which forms the basis for exotic option pricing and contains the European options as special case. In Section 1.5 we give some numerical examples. In Section 1.6 we explain how to efficiently compute the Hermite moments. All proofs are collected in Section 1.7. In Section 1.8 we conclude.

1.2 Model Specification

We study a stochastic volatility model where the squared volatility follows a Jacobi process. Fix some real parameters $0 \le v_{min} < v_{max}$, and define the quadratic function

$$Q(v) = \frac{(v - v_{min})(v_{max} - v)}{(\sqrt{v_{max}} - \sqrt{v_{min}})^2}.$$

Inspection shows that $v \ge Q(v)$, with equality if and only if $v = \sqrt{v_{min} v_{max}}$, and $Q(v) \ge 0$ for all $v \in [v_{min}, v_{max}]$, see Figure 1.1 for an illustration.

We consider the diffusion process (V_t, X_t) given by

$$\begin{aligned} dV_t &= \kappa(\theta - V_t)\,dt + \sigma\sqrt{Q(V_t)}\,dW_{1t} \\ dX_t &= (r - \delta - V_t/2)\,dt + \rho\,\sqrt{Q(V_t)}\,dW_{1t} + \sqrt{V_t - \rho^2\,Q(V_t)}\,dW_{2t} \end{aligned} \tag{1.1}$$

for real parameters $\kappa > 0$, $\theta \in (v_{min}, v_{max}]$, $\sigma > 0$, interest rate r, dividend yield δ, and $\rho \in [-1, 1]$, and where W_{1t} and W_{2t} are independent standard Brownian motions on some filtered probability space $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{Q})$. The following theorem shows that (V_t, X_t) is well defined.

Theorem 1.2.1. *For any deterministic initial state $(V_0, X_0) \in [v_{min}, v_{max}] \times \mathbb{R}$ there exists a unique solution (V_t, X_t) of (1.1) taking values in $[v_{min}, v_{max}] \times \mathbb{R}$ and satisfying*

$$\int_0^\infty \mathbf{1}_{\{V_t = v\}}\,dt = 0 \quad \text{for all } v \in [v_{min}, v_{max}). \tag{1.2}$$

Moreover, V_t takes values in (v_{min}, v_{max}) if and only if $V_0 \in (v_{min}, v_{max})$ and

$$\frac{\sigma^2(v_{max} - v_{min})}{(\sqrt{v_{max}} - \sqrt{v_{min}})^2} \le 2\kappa \min\{v_{max} - \theta, \theta - v_{min}\}. \tag{1.3}$$

Remark 1.2.2. *Property (1.2) implies that no state $v \in [v_{min}, v_{max})$ is absorbing. It also implies that conditional on $\{V_t,\ t \in [0, T]\}$, the increments $X_{t_i} - X_{t_{i-1}}$ are non-degenerate Gaussian for any $t_{i-1} < t_i \le T$ as will be shown in the proof of Theorem 1.4.1. Taking $v_{min} = 0$ and the limit as $v_{max} \to \infty$, condition (1.3) coincides with the known condition that precludes the zero lower bound for the CIR process, $\sigma^2 \le 2\kappa\theta$.*

We specify the price of a traded asset by $S_t = \mathrm{e}^{X_t}$. Then $\sqrt{V_t}$ is the stochastic volatility of the

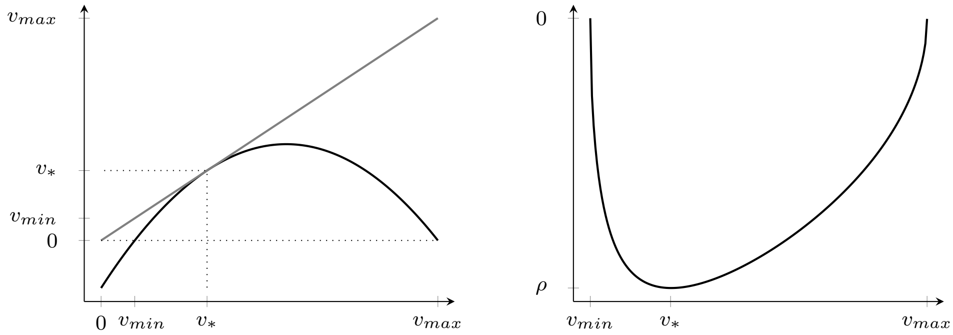


Figure 1.1 – Variance and correlation

The quadratic variation of the Jacobi model (black line) and of the Heston model (gray line) are displayed in the left panel as a function of the instantaneous variance. The right panel displays the instantaneous correlation between the processes X_t and V_t as a function of the instantaneous variance. We denote $v_* = \sqrt{v_{min} v_{max}}$ and assumed that $\rho < 0$.

asset return, $d\langle X, X\rangle_t = V_t\, dt$. The cumulative dividend discounted price process $e^{-(r-\delta)t} S_t$ is a martingale. In other words, $\mathbb{Q}$ is a risk-neutral measure. The parameter ρ tunes the instantaneous correlation between the asset return and the squared volatility,

$$\frac{d\langle V, X\rangle_t}{\sqrt{d\langle V, V\rangle_t}\sqrt{d\langle X, X\rangle_t}} = \rho\sqrt{Q(V_t)/V_t}.$$

This correlation is equal to ρ if $V_t = \sqrt{v_{min} v_{max}}$, see Figure 1.1. In general, we have $\sqrt{Q(V_t)/V_t} \leq 1$. Empirical evidences suggest that ρ is negative when S_t is a stock price or index. This is commonly referred as the leverage effect, that is, an increase in volatility often goes along with a decrease in asset value.

Since the instantaneous squared volatility V_t follows a bounded Jacobi process on the interval $[v_{min}, v_{max}]$, we refer to (1.1) as the *Jacobi model*. For $V_0 = \theta = v_{max}$ we have constant volatility $V_t = V_0$ for all $t \geq 0$ and we obtain the Black–Scholes model

$$dX_t = (r - \delta - V_0/2)\, dt + \sqrt{V_0}\, dW_{2t}. \tag{1.4}$$

For $v_{min} = 0$ and the limit $v_{max} \to \infty$ we have $Q(v) \to v$, and we formally obtain the Heston model as limit case of (1.1),

$$\begin{aligned} dV_t &= \kappa(\theta - V_t)\, dt + \sigma\sqrt{V_t}\, dW_{1t} \\ dX_t &= (r - \delta - V_t/2)\, dt + \sqrt{V_t}\left(\rho\, dW_{1t} + \sqrt{(1-\rho^2)}\, dW_{2t}\right). \end{aligned} \tag{1.5}$$

In fact, the Jacobi model (1.1) is robust with respect to perturbations, or mis-specifications, of the model parameters v_{min}, v_{max} and initial state (V_0, X_0). Specifically, the following theorem shows that the diffusion (1.1) is weakly continuous in the space of continuous paths with respect to v_{min}, v_{max} and (V_0, X_0). In particular, the Heston model (1.5) is indeed a limit case of our model (1.1).

Consider a sequence of parameters $0 \leq v_{min}^{(n)} < v_{max}^{(n)}$ and deterministic initial states $(V_0^{(n)}, X_0^{(n)}) \in [v_{min}^{(n)}, v_{max}^{(n)}] \times \mathbb{R}$ converging to $0 \leq v_{min} < v_{max} \leq \infty$ and $(V_0, X_0) \in [0, \infty) \times \mathbb{R}$ as $n \to \infty$, respectively. We denote by $(V_t^{(n)}, X_t^{(n)})$ and (V_t, X_t) the respective solutions of (1.1), or (1.5) if $v_{max} = \infty$. Here is our main convergence result.

Theorem 1.2.3. *The sequence of diffusions* $(V_t^{(n)}, X_t^{(n)})$ *converges weakly in the path space to* (V_t, X_t) *as* $n \to \infty$.

As the discounted put option payoff function $f_{put}(x) = \mathrm{e}^{-rT}\left(\mathrm{e}^k - \mathrm{e}^x\right)^+$ is bounded and continuous on $\mathbb{R}$, it follows from the weak continuity stated in Theorem 1.2.3 that the put option prices based on $(V_t^{(n)}, X_t^{(n)})$ converge to the put option price based on the limiting model (V_t, X_t) as $n \to \infty$. The put-call parity, $\pi_{call} - \pi_{put} = \mathrm{e}^{-\delta T} S_0 - \mathrm{e}^{-rT+k}$, then implies that also call option prices converge as $n \to \infty$. This carries over to more complex path-dependent options with bounded continuous payoff functional.

Polynomial Property

Moments in the Jacobi model (1.1) are given in closed-form. Indeed, let

$$\mathscr{G} f(v,x) = b(v)^\top \nabla f(v,x) + \frac{1}{2}\operatorname{Tr}(a(v)\nabla^2 f(v,x))$$

denote the generator of (V_t, X_t) with drift vector $b(v)$ and the diffusion matrix $a(v)$ given by

$$b(v) = \begin{bmatrix} \kappa(\theta - v) \\ r - \delta - v/2 \end{bmatrix}, \quad a(v) = \begin{bmatrix} \sigma^2 Q(v) & \rho\sigma Q(v) \\ \rho\sigma Q(v) & v \end{bmatrix}. \tag{1.6}$$

Observe that $a(v)$ is continuous in the parameters v_{min}, v_{max}, so that for $v_{min} = 0$ and $v_{max} \to \infty$ we obtain

$$a(v) \to \begin{bmatrix} \sigma^2 v & \rho\sigma v \\ \rho\sigma v & v \end{bmatrix},$$

which corresponds to the generator of the Heston model (1.5). Let Pol_n be the vector space of polynomials in (v,x) of degree less than or equal to n. It then follows by inspection that the components of $b(v)$ and $a(v)$ lie in Pol_1 and Pol_2, respectively. As a consequence, $\mathscr{G}$ maps any polynomial of degree n onto a polynomial of degree n or less, $\mathscr{G}\,\mathrm{Pol}_n \subset \mathrm{Pol}_n$, so that (V_t, X_t) is a polynomial diffusion, see (Filipović and Larsson 2016, Lemma 2.2). From this we can easily calculate the conditional moments of (V_T, X_T) as follows. For $N \in \mathbb{N}$, let $M = (N+2)(N+1)/2$ denote the dimension of Pol_N. Let $h_1(v,x), \ldots, h_M(v,x)$ be a basis of polynomials of Pol_N and denote by G the matrix representation of the linear map $\mathscr{G}$ restricted to Pol_N with respect to this basis.

Theorem 1.2.4. *For any polynomial* $p \in \mathrm{Pol}_N$ *and* $0 \le t \le T$ *we have*

$$\mathbb{E}[p(V_T, X_T)|\mathscr{F}_t] = [h_1(V_t, X_t), \dots, h_M(V_t, X_t)]\mathrm{e}^{(T-t)G}\vec{p}$$

where $\vec{p} \in \mathbb{R}^M$ *is the coordinate representation of the polynomial* $p(v, x)$ *with respect to the basis* $h_1(v, x), \dots, h_M(v, x)$.

Proof. See (Filipović and Larsson 2016, Theorem 3.1). □

The moment formula in Theorem 1.2.4 is crucial in order to efficiently implement the numerical schemes described below.

1.3 European Option Pricing

Henceforth we assume that $(V_0, X_0) \in [v_{min}, v_{max}] \times \mathbb{R}$ is a deterministic initial state and fix a finite time horizon $T > 0$. We first establish some key properties of the distribution of X_T. Denote the quadratic variation of the second martingale component of X_t in (1.1) by

$$C_t = \int_0^t \left(V_s - \rho^2 Q(V_s)\right) ds. \tag{1.7}$$

The following theorem is a special case of Theorem 1.4.1 below.

Theorem 1.3.1. *Let* $\epsilon < 1/(2v_{max}T)$. *The distribution of* X_T *admits a density* $g_T(x)$ *on* $\mathbb{R}$ *that satisfies*

$$\int_{\mathbb{R}} \mathrm{e}^{\epsilon x^2} g_T(x)\, dx < \infty. \tag{1.8}$$

If

$$\mathbb{E}\left[C_T{}^{-1/2-k}\right] < \infty \tag{1.9}$$

for some $k \in \mathbb{N}_0$ *then* $g_T(x)$ *and* $\mathrm{e}^{\epsilon x^2} g_T(x)$ *are uniformly bounded and* $g_T(x)$ *is* k*-times continuously differentiable on* $\mathbb{R}$. *A sufficient condition*[1] *for* (1.9) *to hold for any* $k \ge 0$ *is*

$$v_{min} > 0 \text{ and } \rho^2 < 1. \tag{1.10}$$

The condition that $\epsilon < 1/(2v_{max}T)$ is sharp for (1.8) to hold. Indeed, consider the Black–Scholes model (1.4) where $V_t = \theta = v_{max}$ for all $t \ge 0$. Then X_T is Gaussian with variance $C_T = v_{max}T$. Hence the integral in (1.8) is infinite for any $\epsilon \ge 1/(2v_{max}T)$.

---

[1] We conjecture that (1.9) holds for any $k \ge 0$ also when $v_{min} = 0$ (and $\kappa\theta > 0$) or $\rho^2 = 1$. For the Heston model (1.5) with $Q(v) = v$ and $\rho^2 < 1$ the conjecture follows from (Dufresne 2001, Theorem 4.1).

Since any uniformly bounded and integrable function on $\mathbb{R}$ is square integrable on $\mathbb{R}$, as an immediate consequence of Theorem 1.3.1 we have the following corollary.

Corollary 1.3.2. *Assume* (1.9) *holds for* $k = 0$. *Then*

$$\int_{\mathbb{R}} \frac{g_T(x)^2}{w(x)}\, dx < \infty \tag{1.11}$$

for any Gaussian density $w(x)$ *with variance* σ_w^2 *satisfying*

$$\sigma_w^2 > \frac{v_{max} T}{2}. \tag{1.12}$$

Remark 1.3.3. *It follows from the proof that the statements of Theorem 1.3.1 also hold for the Heston model* (1.5) *with* $Q(v) = v$ *and* $\epsilon = 0$. *However, the Heston model does not satisfy* (1.8) *for any* $\epsilon > 0$. *Indeed, otherwise its moment generating function*

$$\widehat{g_T}(z) = \int_{\mathbb{R}} \mathrm{e}^{zx} g_T(x)\, dx \tag{1.13}$$

would extend to an entire function in $z \in \mathbb{C}$. *But it is well known that* $\widehat{g_T}(z)$ *becomes infinite for large enough* $z \in \mathbb{R}$, *see (Andersen and Piterbarg 2007). As a consequence, the Heston model does not satisfy* (1.11) *for any finite* σ_w. *Indeed, by the Cauchy-Schwarz inequality,* (1.11) *implies* (1.8) *for any* $\epsilon < 1/(4\sigma_w^2)$.

We now compute the price at time $t = 0$ of a European claim with discounted payoff $f(X_T)$ at expiry date $T > 0$. We henceforth assume that (1.9) holds with $k = 0$, and we let $w(x)$ be a Gaussian density with mean μ_w and variance σ_w^2 satisfying (1.12). We define the weighted Lebesgue space

$$L_w^2 = \left\{ f(x) : \|f\|_w^2 = \int_{\mathbb{R}} f(x)^2\, w(x) dx < \infty \right\},$$

which is a Hilbert space with scalar product

$$\langle f, g\rangle_w = \int_{\mathbb{R}} f(x) g(x)\, w(x) dx.$$

The space L_w^2 admits the orthonormal basis of generalized Hermite polynomials $H_n(x)$, $n \geq 0$, given by

$$H_n(x) = \frac{1}{\sqrt{n!}} \mathscr{H}_n\left(\frac{x - \mu_w}{\sigma_w}\right) \tag{1.14}$$

where $\mathscr{H}_n(x)$ are the standard probabilist Hermite polynomials defined by

$$\mathscr{H}_n(x) = (-1)^n \mathrm{e}^{\frac{x^2}{2}} \frac{d^n}{dx^n} \mathrm{e}^{-\frac{x^2}{2}}, \tag{1.15}$$

see (Feller 1960, Section XVI.1). In particular, $\deg H_n(x) = n$, and $\langle H_m, H_n\rangle_w = 1$ if $m = n$ and zero otherwise.

Corollary 1.3.2 implies that the likelihood ratio function $\ell(x) = g_T(x)/w(x)$ of the density $g_T(x)$ of the log price X_T with respect to $w(x)$ belongs to L^2_w. We henceforth assume that also the discounted payoff function $f(x)$ is in L^2_w. This hypothesis is satisfied for instance in the case of European call and put options. It implies that the price, denoted by π_f, is well defined and equals

$$\pi_f = \int_{\mathbb{R}} f(x) g_T(x)\, dx = \langle f, \ell \rangle_w = \sum_{n\geq 0} f_n \ell_n, \tag{1.16}$$

for the *Fourier coefficients* of $f(x)$

$$f_n = \langle f, H_n \rangle_w, \tag{1.17}$$

and the Fourier coefficients of $\ell(x)$ that we refer to as *Hermite moments*

$$\ell_n = \langle \ell, H_n \rangle_w = \int_{\mathbb{R}} H_n(x) g_T(x)\, dx. \tag{1.18}$$

We approximate the price π_f by truncating the series in (1.16) at some order $N \geq 1$ and write

$$\pi_f^{(N)} = \sum_{n=0}^{N} f_n \ell_n, \tag{1.19}$$

so that $\pi_f^{(N)} \to \pi_f$ as $N \to \infty$. Due to the polynomial property of the Jacobi model, (1.19) induces an efficient price approximation scheme because the Hermite moments ℓ_n are linear combinations of moments of X_T and thus given in closed-form, see Theorem 1.2.4. In particular, since $H_0(x) = 1$, we have $\ell_0 = 1$. More details on the computation of ℓ_n are given in Section 1.6.

With the Hermite moments ℓ_n available, the computation of the approximation (1.19) boils down to a numerical integration,

$$\pi_f^{(N)} = \sum_{n=0}^{N} \left\langle f, \ell_n H_n \right\rangle_w = \int_{\mathbb{R}} f(x) \ell^{(N)}(x)\, w(x) dx, \tag{1.20}$$

of $f(x)\ell^{(N)}(x)$ with respect to the Gaussian distribution $w(x)dx$, where $\ell^{(N)}(x) = \sum_{n=0}^{N} \ell_n H_n(x)$ is in closed-form. The integral (1.20) can be computed by quadrature or Monte-Carlo simulation. In specific cases, we find closed-form formulas for the Fourier coefficients f_n and no numerical integration is needed. This includes European call, put, and digital options, as shown below.

Remark 1.3.4. *Formula* (1.20) *shows that $g_T^{(N)}(x) = \ell^{(N)}(x) w(x)$ serves as an approximation for the density $g_T(x)$. In fact, we readily see that $g_T^{(N)}(x)$ integrates to one and converges to $g_T(x)$ in $L^2_{1/w}$ as $N \to \infty$. Hence, we have convergence of the Gram–Charlier A series expansion of the* *density of the log price X_T in $L^2_{1/w}$.*[2] *In view of Remark 1.3.3, this does not hold for the Heston model.*

Matching the first moment or the first two moments of $w(x)$ and $g_T(x)$, we further obtain

$$\ell_1 = \int_{\mathbb{R}} H_1(x) g_T(x)\, dx = \langle H_0, H_1 \rangle_w = 0 \quad \text{if } \mu_w = \mathbb{E}[X_T],$$

and similarly,

$$\ell_1 = \ell_2 = 0 \quad \text{if } \mu_w = \mathbb{E}[X_T] \text{ and } \sigma_w^2 = \mathrm{var}[X_T]. \tag{1.21}$$

Matching the first moment or the first two moments of $w(x)$ and $g_T(x)$ can improve the convergence of the approximation (1.19). Note however that (1.12) and (1.21) imply $\mathrm{var}[X_T] > v_{max} T/2$, so that second moment matching is not always feasible in empirical applications.

Remark 1.3.5. *If $\mu_w = X_0 + (r-\delta)T - \sigma_w^2/2$, then $f_0 = \int_{\mathbb{R}} f(x) w(x) dx$ is the Black–Scholes option price with volatility $\sigma_{BS} = \sigma_w/\sqrt{T}$. Because $\mathbb{E}[X_T] = X_0 + (r-\delta)T - \mathrm{var}[X_T]/2$, this holds in particular if the first two moments of $w(x)$ and $g_T(x)$ match, see* (1.21). *In this case, the higher order terms in $\pi_f^{(N)} = f_0 + \sum_{n=3}^N f_n \ell_n$ can be thought of as corrections to the corresponding Black–Scholes price f_0 due to stochastic volatility.*

The following result, which is a special case of Theorem 1.4.4 below, provides universal upper and lower bounds on the implied volatility of a European option with discounted payoff $f(X_T)$ at T and price π_f. The implied volatility σ_{IV} is defined as the volatility parameter that renders the corresponding Black–Scholes option price equal to π_f.

Theorem 1.3.6. *Assume that the discounted payoff function $f(\log(s))$ is convex in $s > 0$. Then the implied volatility satisfies $\sqrt{v_{min}} \le \sigma_{\mathrm{IV}} \le \sqrt{v_{max}}$.*

Examples

We now present examples of discounted payoff functions $f(x)$ for which closed-form formulas for the Fourier coefficients f_n exist. The first example is a call option.[3]

Theorem 1.3.7. *Consider the discounted payoff function for a call option with log strike k,*

$$f(x) = \mathrm{e}^{-rT}\left(\mathrm{e}^x - \mathrm{e}^k\right)^+. \tag{1.22}$$

---

[2] A Gram–Charlier A series expansion of a density function $g(x)$ is formally defined as $g(x) = \sum_{n\ge 0} c_n H_n(x) w(x)$ for some real numbers c_n, $n \ge 0$.

[3] Similar recursive relations of the Fourier coefficients for the Physicist Hermite polynomial basis can be found in (Drimus, Necula, and Farkas 2013).

Its Fourier coefficients f_n in (1.17) *are given by*

$$\begin{aligned} f_0 &= \mathrm{e}^{-rT+\mu_w} I_0\left(\frac{k-\mu_w}{\sigma_w};\sigma_w\right) - \mathrm{e}^{-rT+k}\Phi\left(\frac{\mu_w-k}{\sigma_w}\right); \\ f_n &= \mathrm{e}^{-rT+\mu_w}\frac{1}{\sqrt{n!}}\sigma_w I_{n-1}\left(\frac{k-\mu_w}{\sigma_w};\sigma_w\right), \quad n\geq 1. \end{aligned} \tag{1.23}$$

The functions $I_n(\mu;\nu)$ are defined recursively by

$$\begin{aligned} I_0(\mu;\nu) &= \mathrm{e}^{\frac{\nu^2}{2}}\Phi(\nu-\mu); \\ I_n(\mu;\nu) &= \mathscr{H}_{n-1}(\mu)\mathrm{e}^{\nu\mu}\phi(\mu) + \nu I_{n-1}(\mu;\nu), \quad n\geq 1, \end{aligned} \tag{1.24}$$

where $\Phi(x)$ denotes the standard Gaussian distribution function and $\phi(x)$ its density.

The Fourier coefficients of a put option can be obtained from the put-call parity. For digital options, the Fourier coefficients f_n are as follows.

Theorem 1.3.8. *Consider the discounted payoff function for a digital option of the form*

$$f(x) = \mathrm{e}^{-rT}\mathbf{1}_{[k,\infty)}(x).$$

Its Fourier coefficients f_n are given by

$$\begin{aligned} f_0 &= \mathrm{e}^{-rT}\Phi\left(\frac{\mu_w-k}{\sigma_w}\right); \\ f_n &= \frac{\mathrm{e}^{-rT}}{\sqrt{n!}}\mathscr{H}_{n-1}\left(\frac{k-\mu_w}{\sigma_w}\right)\phi\left(\frac{k-\mu_w}{\sigma_w}\right), \quad n\geq 1, \end{aligned} \tag{1.25}$$

where $\Phi(x)$ denotes the standard Gaussian distribution function and $\phi(x)$ its density.

For a digital option with generic payoff $\mathbf{1}_{[k_1,k_2)}(x)$ the Fourier coefficients can be derived using Theorem 1.3.8 and $\mathbf{1}_{[k_1,k_2)}(x) = \mathbf{1}_{[k_1,\infty)}(x) - \mathbf{1}_{[k_2,\infty)}(x)$.

Error Bounds and Asymptotics

We first discuss an error bound of the price approximation scheme (1.19). For a fixed order $N\geq 1$, the error of the approximation is $\epsilon^{(N)} = \pi_f - \pi_f^{(N)} = \sum_{n=N+1}^{\infty} f_n\ell_n$. The Cauchy–Schwarz inequality implies the following error bound

$$|\epsilon^{(N)}| \leq \left(\|f\|_w^2 - \sum_{n=0}^{N} f_n^2\right)^{\frac{1}{2}}\left(\|\ell\|_w^2 - \sum_{n=0}^{N}\ell_n^2\right)^{\frac{1}{2}}. \tag{1.26}$$

The L_w^2-norm of $f(x)$ has an explicit expression, $\|f\|_w^2 = \int_{\mathbb{R}} f(x)^2\, w(x)dx$, that can be computed by quadrature or Monte–Carlo simulation. The Fourier coefficients f_n can be computed similarly. The Hermite moments ℓ_n are given in closed-form. It remains to compute the

L^2_w-norm of $\ell(x)$. For further use we define

$$M_t = X_0 + \int_0^t (r - \delta - V_s/2)\, ds + \frac{\rho}{\sigma}\left(V_t - V_0 - \int_0^t \kappa\,(\theta - V_s)\, ds\right), \tag{1.27}$$

so that, in view of (1.1), the log price $X_t = M_t + \int_0^t \sqrt{V_s - \rho^2 Q(V_s)}\, dW_{2s}$. Recall also C_t given in (1.7).

Lemma 1.3.9. *The L^2_w-norm of $\ell(x)$ is given by*

$$\|\ell\|^2_w = \int_{\mathbb{R}} \frac{g_T(x)^2}{w(x)} dx = \mathbb{E}\left[\frac{g_T(X_T)}{w(X_T)}\right] = \mathbb{E}\left[\frac{\phi\left(X_T, \widetilde{M}_T, \widetilde{C}_T\right)}{\phi\left(X_T, \mu_w, \sigma^2_w\right)}\right] \tag{1.28}$$

where $\phi(x, \mu, \sigma^2)$ is the normal density function in x with mean μ and variance σ^2, and the pair of random variables $(\widetilde{M}_T, \widetilde{C}_T)$ is independent from X_T and has the same distribution as (M_T, C_T).

In applications, we compute the right hand side of (1.28) by Monte–Carlo simulation of $(X_T, \widetilde{M}_T, \widetilde{C}_T)$ and thus obtain the error bound (1.26).

We next show that the Hermite moments ℓ_n decay at an exponential rate under some technical assumptions.

Lemma 1.3.10. *Suppose that* (1.10) *holds and $\sigma^2_w > v_{max} T$. Then there exist finite constants $C > 0$ and $0 < q < 1$ such that $\ell^2_n \le C q^n$ for all $n \ge 0$.*

Comparison to Fourier Transform

An alternative dual expression of the price π_f in (1.16) is given by the Fourier integral

$$\pi_f = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(-\mu - \mathrm{i}\lambda)\hat{g}_T(\mu + \mathrm{i}\lambda)\, d\lambda, \tag{1.29}$$

where $\widehat{f}(z)$ and $\widehat{g_T}(z)$ denote the moment generating functions given by (1.13), respectively. Here $\mu \in \mathbb{R}$ is some appropriate dampening parameter such that $\mathrm{e}^{-\mu x} f(x)$ and $\mathrm{e}^{\mu x} g_T(x)$ are Lebesgue integrable and square integrable on $\mathbb{R}$. Indeed, Lebesgue integrability implies that $\widehat{f}(z)$ and $\widehat{g_T}(z)$ are well defined for $z \in \mu + \mathrm{i}\mathbb{R}$ through (1.13). Square integrability and the Plancherel Theorem then yield the representation (1.29). For example, for the European call option (1.22) we have $\widehat{f}(z) = \mathrm{e}^{-rT + k(1+z)}/(z(z+1))$ for $\mathrm{Re}(z) < -1$

Option pricing via (1.29) is the approach taken in the Heston model (1.5), for which there exists a closed-form expression for $\widehat{g_T}(z)$. It is given in terms of the solution of a Riccati equation. The computation of π_f boils down to the numerical integration of (1.29) along with the numerical solution of a Riccati equation for every argument $z \in \mu + \mathrm{i}\mathbb{R}$ that is needed for the integration. The Heston model (which entails $v_{max} \to \infty$) does not adhere to the series representation (1.16) that is based on condition (1.11), see Remark 1.3.3.

The Jacobi model, on the other hand, does not admit a closed-form expression for $\widehat{g_T}(z)$. But the Hermite moments ℓ_n are readily available in closed-form. In conjunction with Theorem 1.3.7, the (truncated) series representation (1.16) thus provides a valuable alternative to the (numerical) Fourier integral approach (1.29) for option pricing. Moreover, the approximation (1.20) can be applied to any discounted payoff function $f(x) \in L^2_w$. This includes functions $f(x)$ that do not necessarily admit closed-form moment generating function $\widehat{f}(z)$ as is required in the Heston model approach. In Section 1.4, we further develop our approach to price path dependent options, which could be a cumbersome task using Fourier transform techniques in the Heston model.

1.4 Exotic Option Pricing

Pricing exotic options with stochastic volatility models is a challenging task. We show that the price of an exotic option whose payoff is a function of a finite sequence of log returns admits a polynomial series representation in the Jacobi model.

Henceforth we assume that $(V_0, X_0) \in [v_{min}, v_{max}] \times \mathbb{R}$ is a deterministic initial state. Consider time points $0 = t_0 < t_1 < t_2 < \cdots < t_d$ and denote the log returns $Y_{t_i} = X_{t_i} - X_{t_{i-1}}$ for $i = 1, \ldots, d$. The following theorem contains Theorem 1.3.1 as special case where $d = 1$.

Theorem 1.4.1. *Let* $\epsilon_1, \ldots, \epsilon_d \in \mathbb{R}$ *be such that* $\epsilon_i < 1/(2v_{max}(t_i - t_{i-1}))$ *for* $i = 1, \ldots, d$. *The random vector* $(Y_{t_1}, \ldots, Y_{t_d})$ *admits a density* $g_{t_1,\ldots,t_d}(y) = g_{t_1,\ldots,t_d}(y_1, \ldots, y_d)$ *on* $\mathbb{R}^d$ *satisfying*

$$\int_{\mathbb{R}^d} \mathrm{e}^{\sum_{i=1}^d \epsilon_i y_i^2} g_{t_1,\ldots,t_d}(y)\, dy < \infty.$$

If

$$\mathbb{E}\left[\prod_{i=1}^d (C_{t_i} - C_{t_{i-1}})^{-1/2-n_i}\right] < \infty \tag{1.30}$$

for all $(n_1 \ldots, n_d) \in \mathbb{N}_0^d$ *with* $\sum_{i=1}^d n_i \le k \in \mathbb{N}_0$, *for some* $k \in \mathbb{N}_0$, *then* $g_{t_1,\ldots,t_d}(y)$ *and* $\mathrm{e}^{\sum_{i=1}^d \epsilon_i y_i^2} g_{t_1,\ldots,t_d}(y)$ *are uniformly bounded and* $g_{t_1,\ldots,t_d}(y)$ *is* k*-times continuously differentiable on* $\mathbb{R}^d$. *Property* (1.10) *implies* (1.30) *for any* $k \ge 0$.

Since any uniformly bounded and integrable function on $\mathbb{R}^d$ is square integrable on $\mathbb{R}^d$, as an immediate consequence of Theorem 1.4.1 we have the following corollary.

Corollary 1.4.2. *Assume* (1.30) *holds for* $k = 0$. *Then*

$$\int_{\mathbb{R}^d} \frac{g_{t_1,\ldots,t_d}(y)^2}{\prod_{i=1}^d w_i(y_i)}\, dy < \infty$$

for all Gaussian densities $w_i(y_i)$ *with variances* $\sigma^2_{w_i}$ *satisfying*

$$\sigma^2_{w_i} > \frac{v_{max}(t_i - t_{i-1})}{2}, \quad i = 1, \ldots, d. \tag{1.31}$$

Remark 1.4.3. *There is a one-to-one correspondence between the vector of log returns* $(Y_{t_1}, \ldots, Y_{t_d})$ *and the vector of log prices* $(X_{t_1}, \ldots, X_{t_d})$*. Indeed,*

$$X_{t_i} = X_0 + \sum_{j=1}^{i} Y_{t_j}.$$

Hence, a crucial consequence of Theorem 1.4.1 is that the finite-dimensional distributions of the process X_t *admit densities with nice decay properties. More precisely, the density of* $(X_{t_1}, \ldots, X_{t_d})$ *is* $g_{t_1,\ldots,t_d}(x_1 - X_0, \ldots, x_d - x_{d-1})$*.*

Suppose that the discounted payoff of an exotic option is of the form $f(X_{t_1}, ..., X_{t_d})$. Assume that (1.30) holds with $k = 0$. Set $w(y) = w(y_1, \ldots, y_d) = \prod_{i=1}^d w_i(y_i)$, where $w_i(y)$ is a Gaussian density with mean μ_{w_i} and variance $\sigma^2_{w_i}$ satisfying (1.31). Define

$$\widetilde{f}(y) = \widetilde{f}(y_1, \ldots, y_d) = f(X_0 + y_1, X_0 + y_1 + y_2, \ldots, X_0 + y_1 + \cdots + y_d).$$

Then by similar arguments as in Section 1.3 the price of the option is

$$\pi_f = \mathbb{E}[f(X_{t_1}, ..., X_{t_d})] = \sum_{n_1,\ldots,n_d \geq 0} \widetilde{f}_{n_1,\ldots,n_d} \ell_{n_1,\ldots,n_d}$$

where the Fourier coefficients $\widetilde{f}_{n_1,\ldots,n_d}$ and the Hermite moments $\ell_{n_1,\ldots,n_d}$ are given by

$$\widetilde{f}_{n_1,\ldots,n_d} = \langle \widetilde{f}, H_{n_1,\ldots,n_d} \rangle_w = \int_{\mathbb{R}^d} \widetilde{f}(y) H_{n_1,\ldots,n_d}(y) w(y)\, dy$$

and

$$\ell_{n_1,\ldots,n_d} = \mathbb{E}\left[H_{n_1,\ldots,n_d}(Y_{t_1}, \ldots, Y_{t_d})\right] \tag{1.32}$$

with $H_{n_1,\ldots,n_d}(y_1, \ldots, y_d) = \prod_{i=1}^d H^{(i)}_{n_i}(y_i)$, where $H^{(i)}_{n_i}(y_i)$ is the generalized Hermite polynomial of degree n_i associated to parameters μ_{w_i} and σ_{w_i}, see (1.14). The price approximation at truncation order $N \geq 1$ is given, in analogy to (1.19), by

$$\pi^{(N)}_f = \sum_{n_1+\cdots+n_d=0}^{N} \widetilde{f}_{n_1,\ldots,n_d} \ell_{n_1,\ldots,n_d}, \tag{1.33}$$

so that $\pi^{(N)}_f \to \pi_f$ as $N \to \infty$.

We now derive universal upper and lower bounds on the implied volatility for the exotic option with discounted payoff function $f(X_{t_1},\dots,X_{t_d})$ and price π_f. We denote by

$$dS_t^{\mathrm{BS}} = S_t^{\mathrm{BS}}(r-\delta)\,dt + S_t^{\mathrm{BS}}\sigma_{\mathrm{BS}}\,dB_t \tag{1.34}$$

the Black–Scholes price process with volatility $\sigma_{\mathrm{BS}} > 0$ where B_t is some Brownian motion. The Black–Scholes price is defined by

$$\pi_f^{\sigma_{\mathrm{IV}}} = \mathbb{E}\left[f(\log S_{t_1}^{\mathrm{BS}},\dots,\log S_{t_d}^{\mathrm{BS}})\right].$$

The implied volatility σ_{IV} is the volatility parameter σ_{BS} that renders the Black–Scholes option price $\pi_f^{\sigma_{\mathrm{IV}}} = \pi_f$. The following theorem provides bounds on the values that σ_{IV} may take.

Theorem 1.4.4. *Assume that* $f(\log(s_1),\dots,\log(s_d))$ *is convex in* $(s_1,\dots,s_d) \in (0,\infty)^d$. *Then the implied volatility satisfies* $\sqrt{v_{min}} \le \sigma_{\mathrm{IV}} \le \sqrt{v_{max}}$.

Examples

We provide some examples of exotic options on the asset with price $S_t = \mathrm{e}^{X_t}$ for which our method applies.

The payoff of a *forward start call option on the underlying return* between dates t and T, and with strike K is $(S_T/S_t - K)^+$ and its discounted payoff function is given by

$$\tilde{f}(y) = \mathrm{e}^{-rT}\left(\mathrm{e}^{y_2} - K\right)^+$$

with the times $t_1 = t$ and $t_2 = T$. Note that $\tilde{f}(y) = \tilde{f}(y_2)$ only depends on y_2, so that this example reduces to the univariate case. In particular, the Fourier coefficients $\tilde{f}_n$ coincide with those of a call option and, as we shall see in Theroem 1.6.3, the *forward* Hermite moments $\ell_n^* = \mathbb{E}[H_n(X_{t_2} - X_{t_1})]$ can be computed efficiently. Theorem 1.4.4 applies in particular to the implied volatility of a forward start call option on the underlying return. This is in contrast to the Heston model for which the implied volatility explodes (except at the money) when the time to maturity of the underlying call option decreases to zero, $T \to t$, see (Jacquier and Roome 2015) for more details.

The payoff of a *forward start call option* with maturity T, strike fixing date t and proportional strike K is $(S_T - KS_t)^+$ and its discounted payoff function is given by

$$\tilde{f}(y) = \mathrm{e}^{-rT}\left(\mathrm{e}^{X_0+y_1+y_2} - K\mathrm{e}^{X_0+y_1}\right)^+$$

with the times $t_1 = t$ and $t_2 = T$. In this case the Fourier coefficients have the form

$$\begin{aligned}\tilde{f}_{n_1,n_2} &= \mathrm{e}^{X_0-rT}\int_{\mathbb{R}^2} \mathrm{e}^{y_1} H_{n_1}(y_1)\,w_1(y_1)(\mathrm{e}^{y_2}-K)^+ H_{n_2}(y_2)\,w_2(y_2)\,dy_1\,dy_2\\ &= \mathrm{e}^{X_0-rT} f_{n_1}^{(0,-\infty)} f_{n_2}^{(0,\log K)} = f_{n_2}^{(0,\log K)} \frac{\sigma_w^{n_1}}{\sqrt{n_1!}} \mathrm{e}^{X_0-rT+\mu_{w_1}+\sigma_{w_1}^2/2},\end{aligned}$$

where $f_n^{(r,k)}$ denotes the Fourier coefficient of a call option for interest rate r and log strike k as in (1.23). Here we have used (1.23)–(1.24) to deduce that $f_{n_1}^{(0,-\infty)} = \frac{\sigma_w^{n_1}}{\sqrt{n_1!}} \mathrm{e}^{\mu_{w_1}+\sigma_{w_1}^2/2}$. In particular no numerical integration is needed. Additionally, the Hermite moments

$$\ell_{n_1,n_2} = \mathbb{E}[H_{n_1}(Y_{t_1})H_{n_2}(Y_{t_2})]$$

can be calculated efficiently as explained in Theorem 1.6.3. The pricing of forward start call options (on the underlying return) in the Black–Scholes model is straightforward. Analytical expressions for forward start call options (on the underlying return) have been provided in the Heston model by (Kruse and Nögel 2005). However, these integral expressions involve the Bessel function of first kind and are therefore rather difficult to implement numerically.

The payoff of an *Asian call option* with maturity T, discrete monitoring dates $t_1 < \cdots < t_d = T$, and fixed strike K is $(\sum_{i=1}^d S_{t_i}/d - K)^+$ and its discounted payoff function is given by

$$\tilde{f}(y) = \mathrm{e}^{-rT}\left(\frac{1}{d}\sum_{i=1}^{d} \mathrm{e}^{X_0+\sum_{j=1}^{i} y_i} - K\right)^+.$$

Similarly, the payoff of an *Asian call option with floating strike* is $(S_T - K\sum_{i=1}^d S_{t_i}/d)^+$ and its discounted payoff function is given by

$$\tilde{f}(y) = \mathrm{e}^{-rT}\left(\mathrm{e}^{X_0+\sum_{j=1}^{d} y_j} - \frac{K}{d}\sum_{i=1}^{d} \mathrm{e}^{X_0+\sum_{j=1}^{i} y_j}\right)^+.$$

The valuation of Asian options with continuously monitoring in the Black–Scholes model has been studied in (Rogers and Shi 1995) and (Yor 2001) among others.

Remark 1.4.5. *The Fourier coefficients may not be available in closed-form for some exotic options, such as the Asian options. In this case, we compute the multi-dimensional version of the approximation* (1.19) *via numerical integration of* (1.20) *with respect to a Gaussian density* $w(x)$ *in* $\mathbb{R}^d$. *This can be efficiently implemented using Gauss-Hermite quadrature, see for example (Jäckel 2005). Specifically, denote* $z_m \in \mathbb{R}^d$ *and* $w_m \in (0,1)$ *the* m*-th point and weight of an* d*-dimensional standard Gaussian cubature rule with* M *points. The price approximation can then be computed as follows*

$$\begin{aligned}\pi_f^{(N)} &= \int_{\mathbb{R}^d} \tilde{f}\big(\mu+\Sigma z\big)\,\ell^{(N)}\big(\mu+\Sigma z\big)\,\frac{1}{(2\pi)^{\frac{d}{2}}}\mathrm{e}^{-\frac{\|z\|^2}{2}}dz\\ &\approx \sum_{m=1}^{M} w_m \tilde{f}_m \sum_{n_1+\cdots+n_d\le N} \ell_{n_1,\ldots,n_d}\prod_{i=1}^{d}\frac{1}{\sqrt{n_i!}}\mathscr{H}_{n_i}(z_{m,i})\end{aligned}$$

where $\mu = (\mu_{w_1},\ldots,\mu_{w_d})^\top$, $\Sigma = \mathrm{diag}(\sigma_{w_1},\ldots,\sigma_{w_d})$, $\tilde{f}_m = \tilde{f}(\mu+\Sigma z_m)$, *and* $\mathscr{H}_n$ *denotes the standard probabilist Hermite polynomial* (1.15). *We emphasize that many elements in the above expression can be precomputed. A numerical example is given for the Asian option in Section 1.5.2 below.*

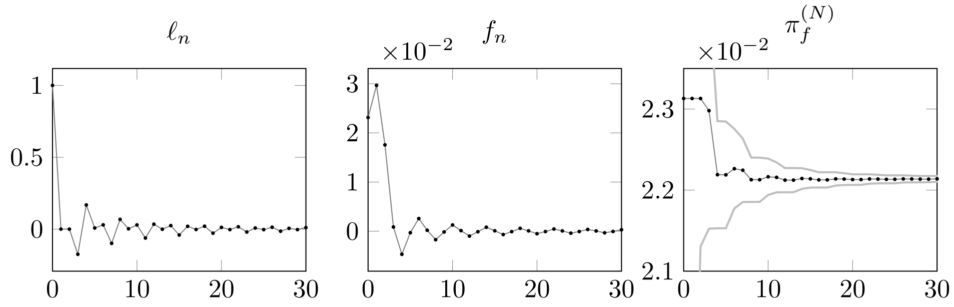


Figure 1.2 – European call option.

Hermite moments ℓ_n, Fourier coefficients f_n, and approximation prices $\pi_f^{(N)}$ with error bounds as functions of the order n (truncation order N).

1.5 Numerical Analysis

We analyse the performance of the price approximation (1.19) with closed-form Fourier coefficients and numerical integration of (1.20) for European call options, forward start and Asian options. This includes price approximation error, model implied volatility, and computational time. The model parameters are fixed as: $r = \delta = X_0 = 0$, $\kappa = 0.5$, $\theta = V_0 = 0.04$, $v_{min} = 10^{-4}$, $v_{max} = 0.08$, $\rho = -0.5$, and $\sigma = 1$.

1.5.1 European Call Option

Figure 1.2 displays Hermite moments ℓ_n, Fourier coefficients f_n, and approximation option prices $\pi_f^{(N)}$ for a European call option with maturity $T = 1/12$ and log strike $k = 0$ (ATM) as functions of the truncation order N. The first two moments of the Gaussian density $w(x)$ match the first two moments of X_T, see (1.21).[4] We observe that the ℓ_n and f_n sequences oscillate and converge toward zero. The amplitudes of these oscillations negatively impact the speed at which the approximation price sequence converges. The gray lines surrounding the price sequence are the upper and lower price error bounds computed as in (1.26) and Lemma 1.3.9, using 10^5 Monte-Carlo samples. The price approximation converges rapidly.

Table 1.1 reports the implied volatility values and absolute errors in percentage points for the log strikes $k = \{-0.1, 0, 0.1\}$ and for various truncation orders. The reference option prices have been computed at truncation order $N = 50$. For all strikes the truncation order $N = 10$ is sufficient to be within 10 basis points of the reference implied volatility.

Figure 1.3 displays the implied volatility smile for various v_{min} and v_{max} such that $\sqrt{v_{min} v_{max}} = \theta$, and for the Heston model (1.5). We observe that the smile of the Jacobi model approaches the Heston smile when v_{min} is small and v_{max} is large. Somewhat surprisingly, a relatively small value for v_{max} seems to be sufficient for the two smiles to coincide for options around

---

[4] In practice, depending on the model parameters, this may not always be feasible, in which case the truncation order N should be increased.

	$k = -0.1$		$k = 0$		$k = 0.1$	
N	IV	error	IV	error	IV	error
0–2	20.13	2.62	20.09	0.86	20.08	0.83
3	22.12	0.63	19.96	0.73	16.60	2.65
4	23.02	0.27	19.27	0.04	18.88	0.37
5	23.03	0.28	19.27	0.04	18.88	0.37
6	22.93	0.18	19.33	0.10	18.72	0.53
7	22.76	0.01	19.32	0.09	19.11	0.14
8	22.83	0.08	19.22	0.01	19.18	0.07
9	22.82	0.07	19.22	0.01	19.19	0.06
10	22.83	0.08	19.25	0.02	19.22	0.03
15	22.74	0.01	19.23	0.00	19.32	0.07
20	22.75	0.00	19.23	0.00	19.28	0.03
30	22.75	0.00	19.23	0.00	19.25	0.00

Table 1.1 – Implied volatility values and errors.
The values and absolute errors are reported in percentage points, for multiple truncation orders N and log strikes k. The reference option price is the 100-th truncation order approximation.

the money. Indeed, although the variance process has an unbounded support in the Heston model, the probability that it will visit values beyond some large threshold can be extremely small. Figure 1.3 also illustrates how the implied volatility smile flattens when the variance support shrinks, $v_{max} \downarrow \theta$. In the limit $v_{max} = \theta$, we obtain the flat implied volatility smile of the Black–Scholes model. This shows that the Jacobi model lies between the Black–Scholes model and the Heston model and that the parameters v_{min} and v_{max} offer additional degrees of flexibility to model the volatility surface.

As reported in Figure 1.4, the Fourier coefficients can be computed in less than a millisecond thanks to the recursive scheme (1.23)-(1.24). Computing the Hermite moments is more costly, however they can be used to price all options with the same maturity. The most expensive task appears to be the construction of the matrix G_n, which however is a one-off. The Hermite moment ℓ_n in turn derives from the vector $v_{n,T} = \mathrm{e}^{G_n T} \boldsymbol{e}_{\pi(0,n)}$ which can be used for any initial state (V_0, X_0). Note that specific numerical methods have been developed to compute the action of the matrix exponential $\mathrm{e}^{G_n T}$ on the basis vector $\boldsymbol{e}_{\pi(0,n)}$, see for example (Al-Mohy and Higham 2011; Hochbruck and Lubich 1997) and references therein. The running times were realized with a standard desktop computer using a single 3.5 Ghz 64 bits CPU and the R programming language.

1.5.2 Forward Start and Asian Options

The left panels of Figure 1.5 display the approximation prices of a forward start call option with strike fixing time $t_1 = 1/52$ and maturity $t_2 = 5/52$, so that $d = 2$, and of an Asian call option with weekly discrete monitoring and maturity four weeks, $t_i = i/52$ for $i \leq d = 4$. Both options have

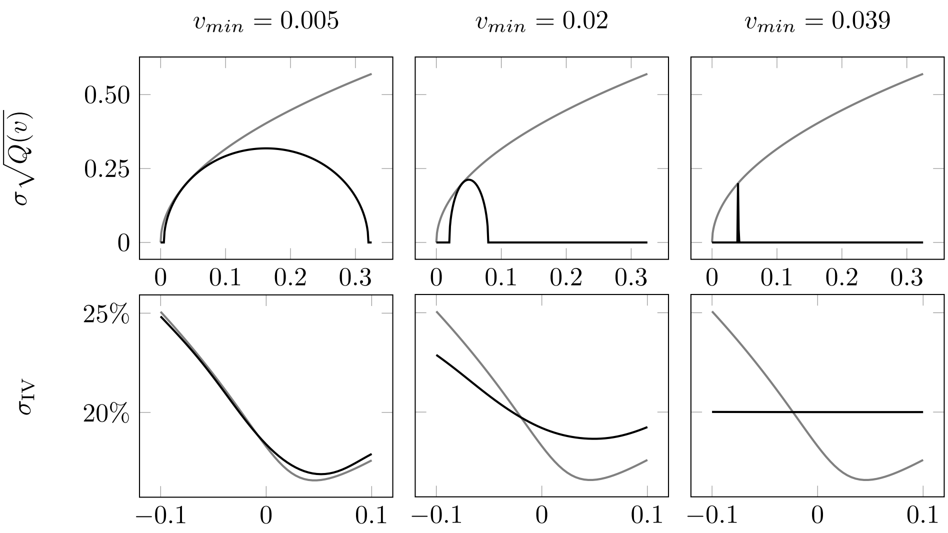


Figure 1.3 – Implied volatility smile: from Heston to Black–Scholes.
The first row displays the variance process' diffusion function in the Jacobi model (black line) and in the Heston model (gray line). The second row displays the implied volatility as a function of the log strike k in the Jacobi model (black line) and in the Heston model (gray line).

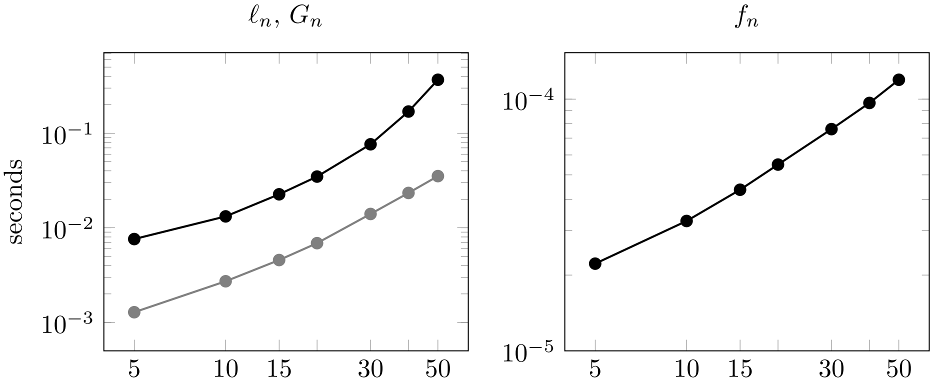


Figure 1.4 – Computational performance.
The left panel displays the computing time to derive the Hermite moments ℓ_n (black line) and the matrix G_n (gray line) as functions of the order n. The right panel displays the same relation for the Fourier coefficients f_n (black line).

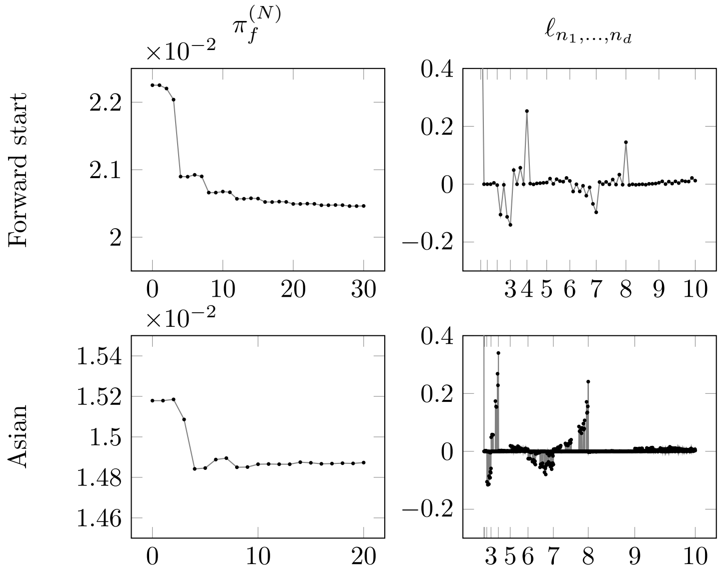


Figure 1.5 – Forward start and Asian options.
The left panels display the approximation prices as functions of the truncation order N. The right panels display the corresponding Hermite moments for multi-orders $n_1 + \cdots + n_d = 1, \dots, 10$.

log strike $k = 0$. The price approximations at order N have been computed using (1.33). For the forward start call option, we match the first two moments of $w_i(y_i)$ and Y_{t_i}. For the Asian call option, we chose $\sigma_{w_i} = \sqrt{v_{max}/104} + 10^{-4}$ and $\mu_{w_i} = E[X_{1/52}]$, which is in line with (1.31) but does not match the first two moments of Y_{t_i}. The Fourier coefficients are not available in closed-form for the Asian call option, therefore we integrated its payoff function with respect to the density approximation using Gaussian cubature as described in Remark 1.4.5. We observe that with exotic payoffs the price approximation sequence may require a larger order before stabilizing. For example, for the forward start price approximation it seems necessary to truncate beyond $N = 15$ in order to obtain a accurate price approximation.

The right panels of Figure 1.5 display the multi-index Hermite moments $\ell_{n_1,\dots,n_d}$ with multi-orders $n_1 + \cdots + n_d = 1, \dots, 10$. Note that there are $\binom{N+d}{N}$ Hermite moments $\ell_{n_1,\dots,n_d}$ of total order $n_1 + \cdots + n_d \le N$. The dimensionality to be handled may therefore become a computational challenge. Yet, we observe that a significant proportion of the Hermite moments is negligible and may simply be set to zero in practice.

1.6 Hermite Moments

We apply Theorem 1.2.4 to describe more explicitly how the Hermite moments $\ell_0, \dots, \ell_N$ in (1.18) can be efficiently computed for any fixed truncation order $N \ge 1$. We let $M = \dim \mathrm{Pol}_N$

and $\pi : \mathscr{E} \to \{1, \dots, M\}$ be an enumeration of the set of exponents

$$\mathscr{E} = \{(m,n) : m,n \geq 0;\ m+n \leq N\}.$$

The polynomials

$$h_{\pi(m,n)}(v,x) = v^m H_n(x), \quad (m,n) \in \mathscr{E} \tag{1.35}$$

then form a basis of Pol_N. In view of the elementary property

$$H_n'(x) = \frac{\sqrt{n}}{\sigma_w} H_{n-1}(x), \quad n \geq 1,$$

we obtain that the $M \times M$–matrix G representing $\mathscr{G}$ on Pol_N has at most 7 nonzero elements in column $\pi(m,n)$ with $(m,n) \in \mathscr{E}$ given by

$$\begin{aligned}
G_{\pi(m-2,n),\pi(m,n)} &= -\frac{\sigma^2 m(m-1) v_{max} v_{min}}{2(\sqrt{v_{max}} - \sqrt{v_{min}})^2}, \quad m \geq 2;\\
G_{\pi(m-1,n-1),\pi(m,n)} &= -\frac{\sigma \rho m \sqrt{n} v_{max} v_{min}}{\sigma_w (\sqrt{v_{max}} - \sqrt{v_{min}})^2}, \quad m,n \geq 1;\\
G_{\pi(m-1,n),\pi(m,n)} &= \kappa\theta m + \frac{\sigma^2 m(m-1)(v_{max} + v_{min})}{2(\sqrt{v_{max}} - \sqrt{v_{min}})^2}, \quad m \geq 1;\\
G_{\pi(m,n-1),\pi(m,n)} &= \frac{(r-\delta)\sqrt{n}}{\sigma_w} + \frac{\sigma \rho m \sqrt{n}(v_{max} + v_{min})}{\sigma_w (\sqrt{v_{max}} - \sqrt{v_{min}})^2}, \quad n \geq 1;\\
G_{\pi(m+1,n-2),\pi(m,n)} &= \frac{\sqrt{n(n-1)}}{2\sigma_w^2}, \quad n \geq 2;\\
G_{\pi(m,n),\pi(m,n)} &= -\kappa m - \frac{\sigma^2 m(m-1)}{2(\sqrt{v_{max}} - \sqrt{v_{min}})^2}\\
G_{\pi(m+1,n-1),\pi(m,n)} &= -\frac{\sqrt{n}}{2\sigma_w} - \frac{\sigma \rho m \sqrt{n}}{\sigma_w (\sqrt{v_{max}} - \sqrt{v_{min}})^2}, \quad n \geq 1.
\end{aligned}$$

Theorem 1.2.4 now implies the following result.

Theorem 1.6.1. *The coefficients ℓ_n are given by*

$$\ell_n = [h_1(V_0, X_0), \dots, h_M(V_0, X_0)]\, \mathrm{e}^{TG}\, \mathbf{e}_{\pi(0,n)}, \quad 0 \leq n \leq N, \tag{1.36}$$

where $\mathbf{e}_i$ is the i–th standard basis vector in $\mathbb{R}^M$.

Remark 1.6.2. *The choice of the basis polynomials $h_{\pi(m,n)}$ in* (1.35) *is convenient for our purposes because: 1) each column of the $M \times M$-matrix G has at most seven nonzero entries. 2) The coefficients ℓ_n in the expansion of prices* (1.16)*, can be obtained directly from the action of $\mathrm{e}^{G_n T}$ on $\mathbf{e}_{\pi_{(0,n)}}$ as specified in* (1.36)*. In practice, it is more efficient to compute directly this action, rather than computing the matrix exponential $\mathrm{e}^{G_n T}$ and then selecting the $\pi_{(0,n)}$-column.*

We now extend Theorem 1.6.1 to a multi-dimensional setting. The following theorem provides an efficient way to calculate multi-dimensional Hermite moments as defined in (1.32). Before stating the theorem we fix some notation. Set $N = \sum_{i=1}^{d} n_i$ and $M = \dim \mathrm{Pol}_N$. Let $G_N^{(i)}$ be the matrix representation of the linear map $\mathscr{G}$ restricted to Pol_N with respect to the basis, in row vector form,

$$h^{(i)}(v,x) = [h_1^{(i)}(v,x), \dots, h_M^{(i)}(v,x)],$$

with $h^{(i)}_{\pi(m,n)}(v,x) = v^m H_n^{(i)}(x)$ as in (1.35) where $H_n^{(i)}$ is the generalized Hermite polynomial of degree n associated to the parameters μ_{w_i} and σ_{w_i}, see (1.14). Define the $M \times M$-matrix $A^{(k,l)}$ by

$$A^{(k,l)}_{i,j} = \begin{cases} H_n^{(l)}(0) & \text{if } i = \pi(m,k) \text{ and } j = \pi(m,n) \text{ for some } m, n \in \mathbb{N} \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 1.6.3. *For any $n_1, \dots, n_d \in \mathbb{N}_0$, the multi-dimensional Hermite moment in* (1.32) *can be computed through*

$$\ell_{n_1,\dots,n_d} = h^{(1)}(V_0, 0) \left(\prod_{i=1}^{d-1} \mathrm{e}^{G_N^{(i)} \Delta t_i} A^{(n_i, i+1)} \right) \mathrm{e}^{G_N^{(d)} \Delta t_d} \boldsymbol{e}_{\pi(0,n_d)},$$

where $\Delta t_i = t_i - t_{i-1}$.

Proof. By an inductive argument it is sufficient to illustrate the case $n = 2$. Applying the law of iterated expectation we obtain

$$\ell_{n_1,n_2} = \mathbb{E}[H_{n_1}^{(1)}(Y_{t_1}) H_{n_2}^{(2)}(Y_{t_2})] = \mathbb{E}[H_{n_1}^{(1)}(X_{t_1} - X_0) \mathbb{E}_{t_1}[H_{n_2}^{(2)}(X_{t_2} - X_{t_1})]].$$

Since the increment $X_{t_2} - X_{t_1}$ does not depend on X_{t_1} we can rewrite, using Theorem 1.2.4,

$$\mathbb{E}_{t_1}[H_{n_2}^{(2)}(X_{t_2} - X_{t_1})] = \mathbb{E}[H_{n_2}^{(2)}(X_{\Delta t_2}) \mid X_0 = 0, V_0 = V_{t_1}] = h^{(2)}(V_{t_1}, 0) v^{(n_2,2)}$$

where $v^{(n_2,2)} = e^{G_N^{(2)} \Delta t_2} \boldsymbol{e}_{\pi(0,n_2)}$. Note that this last expression is a polynomial solely in V_{t_1}

$$h^{(2)}(V_{t_1}, 0) v^{(n_2,2)} = \sum_{n=0}^{n_2} a_n V_{t_1}^n, \quad \text{with } a_n = \sum_{n+j \le n_2} H_j^{(2)}(0)\, v^{(n_2,2)}_{\pi(n,j)}.$$

Theorem 1.2.4 now implies that the Hermite coefficient is given by

$$\ell_{n_1,n_2} = \mathbb{E}[p(V_{t_1}, X_{t_1}) \mid X_0 = 0] = h^{(1)}(V_0, 0) \mathrm{e}^{G_N^{(1)} \Delta t_1} \vec{p}$$

where $\vec{p}$ is the vector representation in the basis $h^{(1)}(v,x)$ of the polynomial

$$p(v,x) = \sum_{n=0}^{n_2} a_n v^n H_{n_1}(x) = h^{(1)}(v,x) \vec{p}.$$

We conclude by observing that the coordinates of the vector $\vec{p}$ are given by $\mathrm{e}_i^\top \vec{p} = a_n$ if $i = \pi(n, n_1)$ for some integer $n \le n_2$ and equal to zero otherwise, which in turn shows that $\vec{p} = A^{(n_1,2)}\, v^{(n_2,2)}$. □

1.7 Proofs

This Section contains the proofs of all theorems and propositions in the main text.

Proof of Theorem 1.2.1

For strong existence and uniqueness of (1.1), it is enough to show strong existence and uniqueness for the SDE for V_t,

$$dV_t = \kappa(\theta - V_t)\, dt + \sigma\sqrt{Q(V_t)}\, dW_{1t}. \tag{1.37}$$

Since the interval $[0,1]$ is an affine transformation of the unit ball in $\mathbb{R}$, weak existence of a $[v_{min}, v_{max}]$-valued solution can be deduced from (Larsson and Pulido 2017, Theorem 2.1). Path-wise uniqueness of solutions follows from (Yamada and Watanabe 1971, Theorem 1). Strong existence of solutions for the SDE (1.37) is a consequence of path-wise uniqueness and weak existence of solutions, see for instance (Yamada and Watanabe 1971, Corollary 1).

Now let $v \in [v_{min}, v_{max})$. The occupation times formula (Revuz and Yor 1999, Corollary VI.1.6) implies

$$\int_0^\infty \mathbf{1}_{\{V_t = v\}} \sigma^2 Q(v)\, dt = 0, \quad \int_0^\infty \mathbf{1}_{\{V_t = v\}} \kappa(\theta - v)\, dt = 0.$$

For $v > v_{min}$ we have $\sigma^2 Q(v) > 0$ and for $v = v_{min}$ we have $\kappa(\theta - v) > 0$, which proves (1.2).

To conclude, Proposition 2.2 in (Larsson and Pulido 2017) shows that $V_t \in (v_{min}, v_{max})$ if and only if $V_0 \in (v_{min}, v_{max})$ and condition (1.3) holds.

Proof of Theorem 1.2.3

The proof of Theorem 1.2.3 builds on the following four lemmas.

Lemma 1.7.1. *Suppose that Y and $Y^{(n)}$, $n \ge 1$, are random variables in $\mathbb{R}^d$ for which all moments exist. Assume further that*

$$\lim_n \mathbb{E}[p(Y^{(n)})] = \mathbb{E}[p(Y)], \tag{1.38}$$

for any polynomial $p(y)$ and that the distribution of Y is determined by its moments. Then the sequence $Y^{(n)}$ converges weakly to Y as $n \to \infty$.

Proof. Theorem 30.2 in (Billingsley 1995) proves this result for the case $d = 1$. Inspection shows that the proof is still valid for the general case. □

Lemma 1.7.2. *The moments of the finite-dimensional distributions of the diffusions* $(V_t^{(n)}, X_t^{(n)})$ *converge to the respective moments of the finite-dimensional distributions of* (V_t, X_t). *That is, for any* $0 \le t_1 < \cdots < t_d < \infty$ *and for any polynomials* $p_1(v,x), \ldots, p_d(v,x)$ *we have*

$$\lim_n \mathbb{E}\left[\prod_{i=1}^{d} p_i(V_{t_i}^{(n)}, X_{t_i}^{(n)})\right] = \mathbb{E}\left[\prod_{i=1}^{d} p_i(V_{t_i}, X_{t_i})\right]. \tag{1.39}$$

Proof. Let $N = \sum_{i=1}^{d} \deg p_i$. Throughout the proof we fix a basis $h_1(v,x), \ldots, h_M(v,x)$ of Pol_N, where $M = \dim \mathrm{Pol}_N$, and for any polynomial $p(v,x)$ we denote by $\vec{p}$ its coordinates with respect to this basis. We denote by G and $G^{(n)}$ the respective $M \times M$-matrix representations of the generators restricted to Pol_N of (V_t, X_t) and $(V_t^{(n)}, X_t^{(n)})$, respectively. We then define recursively the polynomials $q_i(v,x)$ and $q_i^{(n)}(v,x)$ for $1 \le i \le d$ by

$$\begin{aligned} q_d(v,x) &= q_d^{(n)}(v,x) = p_d(v,x), \\ q_i(v,x) &= p_i(v,x)[h_1(v,x), \ldots, h_M(v,x)] \mathrm{e}^{(t_{i+1}-t_i)G} \overrightarrow{q_{i+1}}, \quad 1 \le i < d, \\ q_i^{(n)}(v,x) &= p_i(v,x)[h_1(v,x), \ldots, h_M(v,x)] \mathrm{e}^{(t_{i+1}-t_i)G^{(n)}} \overrightarrow{q_{i+1}^{(n)}}, \quad 1 \le i < d. \end{aligned}$$

As in the proof of Theorem 1.6.3, a successive application of Theorem 1.2.4 and the law of iterated expectation implies that

$$\begin{aligned} \mathbb{E}\left[\prod_{i=1}^{d} p_i(V_{t_i}, X_{t_i})\right] &= \mathbb{E}\left[\prod_{i=1}^{d-1} p_i(V_{t_i}, X_{t_i}) \mathbb{E}\left[p_d(V_{t_d}, X_{t_d}) \mid \mathscr{F}_{t_{d-1}}\right]\right] \\ &= \cdots = [h_1(V_0, X_0), \ldots, h_M(V_0, X_0)]\, \mathrm{e}^{t_1 G} \overrightarrow{q_1}. \end{aligned}$$

and similarly,

$$\mathbb{E}\left[\prod_{i=1}^{d} p_i(V_{t_i}^{(n)}, X_{t_i}^{(n)})\right] = \left[h_1(V_0^{(n)}, X_0^{(n)}), \ldots, h_M(V_0^{(n)}, X_0^{(n)})\right] \mathrm{e}^{t_1 G^{(n)}} \overrightarrow{q_1^{(n)}}.$$

We deduce from (1.6) that

$$\lim_n G^{(n)} = G. \tag{1.40}$$

Note that this is valid also for the limit case $v_{max} = \infty$, that is, $Q(v) = v - v_{min}$. This fact together with an inductive argument shows that $\lim_n \overrightarrow{q_1^{(n)}} = \overrightarrow{q_1}$. This combined with (1.40) proves (1.39). □

Lemma 1.7.3. *The finite-dimensional distributions of* (V_t, X_t) *are determined by their moments.*

Proof. The proof of this result is contained in the proof of (Filipović and Larsson 2016, Lemma

4.1). □

Lemma 1.7.4. *The family of diffusions* $(V_t^{(n)}, X_t^{(n)})$ *is tight.*

Proof. Fix a time horizon $N \in \mathbb{N}$. We first observe that by (Karatzas and Shreve 1991, Problem V.3.15) there is a constant K independent of n such that

$$\mathbb{E}[\|(V_t^{(n)}, X_t^{(n)}) - (V_s^{(n)}, X_s^{(n)})\|^4] \leq K|t-s|^2, \quad 0 \leq s < t \leq N. \tag{1.41}$$

Now fix any positive $\alpha < 1/4$. Kolmogorov's continuity theorem (see (Revuz and Yor 1999, Theorem I.2.1)) implies that

$$\mathbb{E}\left[\left(\sup_{0\leq s<t\leq N} \frac{\|(V_t^{(n)}, X_t^{(n)}) - (V_s^{(n)}, X_s^{(n)})\|}{|t-s|^\alpha}\right)^4\right] \leq J$$

for a finite constant J that is independent of n. The modulus of continuity

$$\Delta(\delta, n) = \sup\left\{\|(V_t^{(n)}, X_t^{(n)}) - (V_s^{(n)}, X_s^{(n)})\| \mid 0 \leq s < t \leq N, |t-s| < \delta\right\}$$

thus satisfies

$$\mathbb{E}[\Delta(\delta, n)^4] \leq \delta^\alpha J.$$

Using Chebyshev's inequality we conclude that, for every $\epsilon > 0$,

$$\mathbb{Q}[\Delta(\delta, n) > \epsilon] \leq \frac{\mathbb{E}[\Delta(\delta, n)^4]}{\epsilon^4} \leq \frac{\delta^\alpha J}{\epsilon^4},$$

and thus $\sup_n \mathbb{Q}[\Delta(\delta, n) > \epsilon] \to 0$ as $\delta \to 0$. This together with the property that the initial states $(V_0^{(n)}, X_0^{(n)})$ converge to (V_0, X_0) as $n \to \infty$ proves the lemma, see (Rogers and Williams 2000, Theorem II.85.3).[5] □

Remark 1.7.5. *Kolmogorov's continuity theorem (see (Revuz and Yor 1999, Theorem I.2.1)) and* (1.41) *imply that the paths of* (V_t, X_t) *are* α*-Hölder continuous for any* $\alpha < 1/4$.

Lemmas 1.7.1–1.7.3 imply that the finite-dimensional distributions of the diffusions $(V_t^{(n)}, X_t^{(n)})$ converge weakly to those of (V_t, X_t) as $n \to \infty$. Theorem 1.2.3 thus follows from Lemma 1.7.4 and (Rogers and Williams 2000, Lemma II.87.3).

---

[5] The derivation of the tightness of $(V_t^{(n)}, X_t^{(n)})$ from (1.41) is also stated without proof in (Rogers and Williams 2000, Theorem II.85.5). For the sake of completeness we give a short self-contained argument here.

Proof of Theorem 1.3.7

We claim that the solution of the recursion (1.24) is given by

$$I_n(\mu;\nu) = \int_\mu^\infty \mathscr{H}_n(x)\mathrm{e}^{\nu x}\phi(x)\,dx, \quad n \geq 0. \tag{1.42}$$

Indeed, for $n = 0$ the right hand side of (1.42) equals

$$\int_\mu^\infty \mathscr{H}_0(x)\mathrm{e}^{\nu x}\phi(x)\,dx = \mathrm{e}^{\frac{\nu^2}{2}}\int_{\mu-\nu}^\infty \phi(x)\,dx,$$

which is $I_0(\mu;\nu)$. For $n \geq 1$, we recall that the standard Hermite polynomials $\mathscr{H}_n(x)$ satisfy

$$\mathscr{H}_n(x) = x\mathscr{H}_{n-1}(x) - \mathscr{H}'_{n-1}(x). \tag{1.43}$$

Integration by parts and (1.43) then show that

$$\begin{aligned}\int_\mu^\infty \mathscr{H}_n(x)\mathrm{e}^{\nu x}\phi(x)\,dx &= \int_\mu^\infty \mathscr{H}_{n-1}(x)\mathrm{e}^{\nu x}x\phi(x)\,dx - \int_\mu^\infty \mathscr{H}'_{n-1}(x)\mathrm{e}^{\nu x}\phi(x)\,dx\\ &= -\mathscr{H}_{n-1}(x)\mathrm{e}^{\nu x}\phi(x)\big|_\mu^\infty + \int_\mu^\infty \mathscr{H}_{n-1}(x)\nu\mathrm{e}^{\nu x}\phi(x)\,dx.\\ &= \mathscr{H}_{n-1}(\mu)\mathrm{e}^{\nu\mu}\phi(\mu) + \nu\int_\mu^\infty \mathscr{H}_{n-1}(x)\mathrm{e}^{\nu x}\phi(x)\,dx,\end{aligned}$$

which proves (1.42).

A change of variables, using (1.14) and (1.42), shows

$$\begin{aligned}f_n &= \mathrm{e}^{-rT}\int_k^\infty \left(\mathrm{e}^x - \mathrm{e}^k\right) H_n(x)\,w(x)\,dx\\ &= \mathrm{e}^{-rT}\int_{\frac{k-\mu_w}{\sigma_w}}^\infty \left(\mathrm{e}^{\mu_w+\sigma_w z} - \mathrm{e}^k\right) H_n(\mu_w+\sigma_w z)\,w(\mu_w+\sigma_w z)\sigma_w\,dz\\ &= \mathrm{e}^{-rT}\frac{1}{\sqrt{n!}}\int_{\frac{k-\mu_w}{\sigma_w}}^\infty \left(\mathrm{e}^{\mu_w+\sigma_w z} - \mathrm{e}^k\right)\mathscr{H}_n(z)\phi(z)\,dz\\ &= \mathrm{e}^{-rT+\mu_w}\frac{1}{\sqrt{n!}}I_n\left(\frac{k-\mu_w}{\sigma_w};\sigma_w\right) - \mathrm{e}^{-rT+k}\frac{1}{\sqrt{n!}}I_n\left(\frac{k-\mu_w}{\sigma_w};0\right).\end{aligned}$$

Formulas (1.23) follow from the recursion formula (1.24).

Proof of Theorem 1.3.8

As before, a change of variables, using (1.14) and (1.42), shows

$$f_n = \mathrm{e}^{-rT}\int_k^\infty H_n(x)\,w(x)\,dx = \frac{\mathrm{e}^{-rT}}{\sqrt{n!}}\int_{\frac{k-\mu_w}{\sigma_w}}^\infty \mathscr{H}_n(z)\phi(z)\,dz = \frac{\mathrm{e}^{-rT}}{\sqrt{n!}}I_n\left(\frac{k-\mu_w}{\sigma_w};0\right).$$

Formulas (1.25) follow directly from (1.24).

Proof of Lemma 1.3.9

We use similar notation as in the proof of Theorem 1.4.1. In particular, with C_T as in (1.7) and M_T as in (1.27), we denote by

$$G_T(x) = (2\pi C_T)^{-\frac{1}{2}} \exp\left(-\frac{(x-M_T)^2}{2C_T}\right) \tag{1.44}$$

the conditional density of X_T given $\{V_t : t \in [0,T]\}$, so that $g_T(x) = \mathbb{E}[G_T(x)]$ is the unconditional density of X_T. Lemma 1.3.9 now follows from observing that $G_T(x) = \phi(x, M_T, C_T)$ and $w(x) = \phi(x, \mu_w, \sigma_w^2)$.

Proof of Lemma 1.3.10

We first recall that by Cramér's inequality (see for instance (Erdélyi, Magnus, Oberhettinger, and Tricomi 1953, Section 10.18)) there exists a constant $K > 0$ such that for all $n \geq 0$

$$\mathrm{e}^{-(x-\mu_w)^2/4\sigma_w^2} |H_n(x)| = (n!)^{-1/2} \mathrm{e}^{-(x-\mu_w)^2/4\sigma_w^2} \left|\mathscr{H}_n\left(\frac{x-\mu_w}{\sigma_w}\right)\right| \leq K. \tag{1.45}$$

Additionally, as in the proof Theorem 1.4.1, since $1/4\sigma_w^2 < 1/(2v_{max}T)$,

$$\mathbb{E}\left[\int_{\mathbb{R}} \mathrm{e}^{(x-\mu_w)^2/4\sigma_w^2} G_T(x)\, dx\right] < \infty,$$

where $G_T(x)$ is given in (1.44). This implies

$$\begin{aligned}\mathbb{E}\left[\int_R |H_n(x)| G_T(x)\, dx\right] &= \mathbb{E}\left[\int_R |H_n(x)| \mathrm{e}^{-(x-\mu_w)^2/4\sigma_w^2} \mathrm{e}^{(x-\mu_w)^2/4\sigma_w^2} G_T(x)\, dx\right] \\ &\leq K\mathbb{E}\left[\int_R \mathrm{e}^{(x-\mu_w)^2/4\sigma_w^2} G_T(x)\, dx\right] < \infty.\end{aligned}$$

We can therefore use Fubini's theorem to deduce

$$\ell_n = \int_{\mathbb{R}} H_n(x) g_T(x)\, dx = \mathbb{E}\left[\int_{\mathbb{R}} H_n(x) G_T(x)\, dx\right] = \mathbb{E}[Y_n]. \tag{1.46}$$

We now analyze the term inside the expectation in (1.46). A change of variables shows

$$Y_n = \int_R H_n(x) G_T(x)\, dx = (2\pi n!)^{-1/2} \int_{\mathbb{R}} \mathscr{H}_n(\alpha y + \beta) \mathrm{e}^{-y^2/2}\, dy,$$

where we define $\alpha = \frac{\sqrt{C_T}}{\sigma_w}$ and $\beta = \frac{M_T - \mu_w}{\sigma_w}$. We recall that

$$0 < (1-\rho^2) v_{min} T \leq C_T \leq v_{max} T < \sigma_w. \tag{1.47}$$

The inequalities in (1.47) together with the fact that V_t is a bounded process yield the following uniform bounds for α, β,

$$1-q=\frac{(1-\rho^2)v_{min}T}{\sigma_w^2}\le \alpha^2\le v_{max}T/\sigma_w^2<1, \quad |\beta|\le R, \tag{1.48}$$

with constants $0<q<1$ and $R>0$. Define

$$x_n=(2\pi)^{-1/2}\int_{\mathbb{R}}\mathscr{H}_n(\alpha y+\beta)\mathrm{e}^{-y^2/2}\,dy,$$

so that

$$Y_n=\int_R H_n(x)G_T(x)\,dx=(n!)^{-1/2}x_n.$$

An integration by parts argument using (1.43) and the identity

$$\mathscr{H}_n'(x)=n\mathscr{H}_{n-1}(x)$$

shows the following recursion formula

$$x_n=\beta x_{n-1}-(n-1)(1-\alpha^2)x_{n-2},$$

with $x_0=1$ and $x_1=\beta$. This recursion formula is closely related to the recursion formula of the Hermite polynomials which helps us deduce the following explicit expression

$$x_n=n!\sum_{m=0}^{\lfloor n/2\rfloor}\frac{(\alpha^2-1)^m}{m!(n-2m)!}\frac{\beta^{n-2m}}{2^m}. \tag{1.49}$$

Recall that

$$\mathscr{H}_n(x)=n!\sum_{m=0}^{\lfloor n/2\rfloor}\frac{(-1)^m}{m!(n-2m)!}\frac{x^{n-2m}}{2^m}. \tag{1.50}$$

By (1.49) and (1.50) we have

$$x_n=n!(1-\alpha^2)^{\frac{n}{2}}\sum_{m=0}^{\lfloor n/2\rfloor}\frac{(-1)^m}{m!(n-2m)!}\frac{((1-\alpha^2)^{-\frac{1}{2}}\beta)^{n-2m}}{2^m}=(1-\alpha^2)^{\frac{n}{2}}\mathscr{H}_n\left((1-\alpha^2)^{-\frac{1}{2}}\beta\right)$$

and

$$\ell_n=\mathbb{E}\left[(1-\alpha^2)^{\frac{n}{2}}n!^{-\frac{1}{2}}\mathscr{H}_n\left((1-\alpha^2)^{-\frac{1}{2}}\beta\right)\right].$$

Cauchy-Schwarz inequality and (1.45) yield

$$\ell_n^2\le\mathbb{E}\left[(n!^{-\frac{1}{2}}\mathscr{H}_n\left((1-\alpha^2)^{-\frac{1}{2}}\beta\right))^2\right]\mathbb{E}[(1-\alpha^2)^n]\le K^2\mathbb{E}\left[\exp(\beta^2/(2(1-\alpha^2)))\right]\mathbb{E}[(1-\alpha^2)^n]. \tag{1.51}$$

Inequalities (1.48) and (1.51) imply the existence of constants $C>0$ and $0<q<1$ such that

$\ell_n^2 \le Cq^n$.

Proof of Theorem 1.4.1

In order to shorten the notation we write $\Delta Z_{t_i} = Z_{t_i} - Z_{t_{i-1}}$ for any process Z_t. From (1.1) we infer that the log price $X_t = M_t + \int_0^t \sqrt{V_s - \rho^2 Q(V_s)}\, dW_{2s}$ where M_t is defined in (1.27). In particular the log returns $Y_{t_i} = \Delta X_{t_i}$ have the form

$$Y_{t_i} = \Delta M_{t_i} + \int_{t_{i-1}}^{t_i} \sqrt{V_s - \rho^2 Q(V_s)}\, dW_{2s}.$$

In view of property (1.2) we infer that $\Delta C_{t_i} > 0$ for $i = 1,\dots,d$. Motivated by (Broadie and Kaya 2006), we notice that, conditional on $\{V_t,\, t \in [0,T]\}$, the random variable $(Y_{t_1},\dots,Y_{t_d})$ is Gaussian with mean vector $(\Delta M_{t_1},\dots,\Delta M_{t_d})$ and covariance matrix $\operatorname{diag}(\Delta C_{t_1},\dots,\Delta C_{t_d})$. Its density $G_{t_1,\dots,t_d}(y) = G_{t_1,\dots,t_d}(y_1,\dots,y_d)$ has the form

$$G_{t_1,\dots,t_d}(y) = (2\pi)^{-d/2} \prod_{i=1}^d (\Delta C_{t_i})^{-1/2} \exp\left[-\sum_{i=1}^d \frac{(y_i - \Delta M_{t_i})^2}{2\Delta C_{t_i}} \right].$$

Fubini's theorem implies that $g_{t_1,\dots,t_d}(y) = \mathbb{E}\left[G_{t_1,\dots,t_d}(y)\right]$ is measurable and satisfies, for any bounded measurable function $f(y)$,

$$\mathbb{E}\left[f(Y_{t_1},\dots,Y_{t_d})\right] = \mathbb{E}\left[\int_{\mathbb{R}^d} f(y) G_{t_1,\dots,t_d}(y)\, dy\right] = \int_{\mathbb{R}^d} f(y) g_{t_1,\dots,t_d}(y)\, dy.$$

Hence the distribution of $(Y_{t_1},\dots,Y_{t_d})$ admits the density $g_{t_1,\dots,t_d}(y)$ on $\mathbb{R}^d$. Dominated convergence implies that $g_{t_1,\dots,t_d}(y)$ is uniformly bounded and k–times continuously differentiable on $\mathbb{R}^d$ if (1.30) holds. The arguments so far do not depended on ϵ_i and also apply to the Heston model, which proves Remark 1.3.3.

For the rest of the proof we assume, without loss of generality, that $\epsilon_i > 0$ for $i = 1,\dots,d$. Observe that the mean vector and covariance matrix of $G_{t_1,\dots,t_d}(y)$ admit the uniform bounds

$$|\Delta M_{t_i}| \le K, \quad |\Delta C_{t_i}| \le v_{max}(t_i - t_{i-1}),$$

for some finite constant K. Define $\Delta_i = 1 - 2\epsilon_i \Delta C_{t_i}$ and $\delta_i = 1 - 2\epsilon_i v_{max}(t_i - t_{i-1})$. Then $\delta_i \in (0,1)$ and $\Delta_i \ge \delta_i$. Completing the square implies

$$\begin{aligned}
e^{\sum_{i=1}^d \epsilon_i y_i^2} G_{t_1,\dots,t_d}(y) &= \prod_{i=1}^d (2\pi\Delta C_{t_i})^{-\frac{1}{2}} \exp\left[\epsilon_i y_i^2 - \frac{(y_i - \Delta M_{t_i})^2}{2\Delta C_{t_i}}\right] \\
&= \prod_{i=1}^d (2\pi\Delta C_{t_i})^{-\frac{1}{2}} \exp\left[-\frac{\Delta_i}{2\Delta C_{t_i}}\left(y_i - \frac{\Delta M_{t_i}}{\Delta_i}\right)^2 + \frac{\Delta M_{t_i}^2}{2\Delta C_{t_i}}\left(\frac{1}{\Delta_i} - 1\right)\right] \qquad (1.52) \\
&= \prod_{i=1}^d (2\pi\Delta C_{t_i})^{-\frac{1}{2}} \exp\left[-\frac{\Delta_i}{2\Delta C_{t_i}}\left(y_i - \frac{\Delta M_{t_i}}{\Delta_i}\right)^2 + \frac{\epsilon_i \Delta M_{t_i}^2}{\Delta_i}\right].
\end{aligned}$$

Integration of (1.52) then gives

$$\int_{\mathbb{R}^d} \mathrm{e}^{\sum_{i=1}^d \epsilon_i y_i^2} G_{t_1,\dots,t_d}(y)\,dy = \prod_{i=1}^d \frac{1}{\sqrt{\Delta_i}} \exp\left[\frac{\epsilon_i \Delta M_{t_i}^2}{\Delta_i}\right] \le \prod_{i=1}^d \frac{1}{\sqrt{\delta_i}} \exp\left[\frac{\epsilon_i K^2}{\delta_i}\right].$$

Hence (1.8) follows by Fubini's theorem after taking expectation on both sides. We also derive from (1.52) that

$$\mathrm{e}^{\sum_{i=1}^d \epsilon_i y_i^2} g_{t_1,\dots,t_d}(y) = \mathbb{E}\left[\mathrm{e}^{\sum_{i=1}^d \epsilon_i y_i^2} G_{t_1,\dots,t_d}(y)\right] \le \mathbb{E}\left[\prod_{i=1}^d (2\pi \Delta C_{t_i})^{-\frac{1}{2}}\right] \prod_{i=1}^d \exp\left[\frac{\epsilon_i K^2}{\delta_i}\right].$$

Hence $\mathrm{e}^{\sum_{i=1}^d \epsilon_i y_i^2} g_{t_1,\dots,t_d}(y)$ is uniformly bounded and continuous on $\mathbb{R}^d$ if (1.30) holds. In fact, for this to hold it is enough suppose that (1.30) holds with $k = 0$. Moreover, (1.10) implies that $\Delta C_{t_i} \ge (t_i - t_{i-1})(1-\rho^2)v_{min} > 0$ and (1.30) follows.

Proof of Theorem 1.4.4

We assume the Brownian motions B_t and (W_{1t}, W_{2t}) in (1.34) and (1.1) are independent. We denote by $\pi_{f,t}$ the time-t price of the exotic option in the Jacobi model.

For any $t_{i-1} \le t < t_i$ and given a realization $X_{t_1}, \dots, X_{t_{i-1}}$, the time-$t$ Black–Scholes price of the option is a function $\pi_f^{\sigma_{\mathrm{BS}}}(t, S_t)$ of t and the spot price S_t defined by

$$\begin{aligned} \mathrm{e}^{-rt}\pi_f^{\sigma_{\mathrm{BS}}}(t,s) &= \mathbb{E}\left[f(X_{t_1},\dots,X_{t_{i-1}}, \log S_{t_i}^{\mathrm{BS}}, \dots, \log S_{t_d}^{\mathrm{BS}}) \mid \mathscr{F}_t,\, S_t^{\mathrm{BS}} = s\right] \\ &= \mathbb{E}\left[f\left(X_{t_1},\dots,X_{t_{i-1}}, \log\left(sR_{t,t_i}^{\mathrm{BS}}\right), \dots, \log\left(sR_{t,t_d}^{\mathrm{BS}}\right)\right) \mid \mathscr{F}_t\right] \end{aligned}$$

where we write

$$R_{t,t_i}^{\mathrm{BS}} = \mathrm{e}^{\left(r-\delta-\frac{1}{2}\sigma_{\mathrm{BS}}^2\right)(t_i-t)+\sigma_{\mathrm{BS}}\left(B_{t_i}-B_t\right)}.$$

By assumption, we infer that $\pi_f^{\sigma_{\mathrm{BS}}}(t,s)$ is convex in $s > 0$. Moreover, $\pi_f^{\sigma_{\mathrm{BS}}}(t,s)$ satisfies the following PDE

$$r\pi_f^{\sigma_{\mathrm{BS}}}(t,s) = \frac{\partial \pi_f^{\sigma_{\mathrm{BS}}}(t,s)}{\partial t} + (r-\delta)s\frac{\partial \pi_f^{\sigma_{\mathrm{BS}}}(t,s)}{\partial s} + \frac{1}{2}\sigma_{\mathrm{BS}}^2 s^2 \frac{\partial^2 \pi_f^{\sigma_{\mathrm{BS}}}(t,s)}{\partial s^2} \tag{1.53}$$

and has terminal value satisfying $\pi_f^{\sigma_{\mathrm{BS}}}(T, S_T) = \pi_{f,T}$. Write

$$\pi_{f,t}^{\sigma_{\mathrm{BS}}} = \pi_f^{\sigma_{\mathrm{BS}}}(t,S_t), \quad \Theta_{f,t}^{\sigma_{\mathrm{BS}}} = -\frac{\partial \pi_f^{\sigma_{\mathrm{BS}}}(t,S_t)}{\partial t}, \quad \Delta_{f,t}^{\sigma_{\mathrm{BS}}} = \frac{\partial \pi_f^{\sigma_{\mathrm{BS}}}(t,S_t)}{\partial s}, \quad \Gamma_{f,t}^{\sigma_{\mathrm{BS}}} = \frac{\partial^2 \pi_f^{\sigma_{\mathrm{BS}}}(t,S_t)}{\partial s^2}$$

and $dN_t = \rho\sqrt{Q(V_t)}\,dW_{1t} + \sqrt{V_t - \rho^2 Q(V_t)}\,dW_{2t}$ for the martingale driving the asset return in (1.1) such that, using (1.53),

$$\begin{aligned}
d(e^{-rt}\pi^{\sigma_{\mathrm{BS}}}_{f,t}) &= \mathrm{e}^{-rt}\left(-r\pi^{\sigma_{\mathrm{BS}}}_{f,t} - \Theta^{\sigma_{\mathrm{BS}}}_{f,t} + (r-\delta)S_t\Delta^{\sigma_{\mathrm{BS}}}_{f,t} + \frac{1}{2}V_tS_t^2\Gamma^{\sigma_{\mathrm{BS}}}_{f,t}\right)dt + \mathrm{e}^{-rt}\Delta^{\sigma_{\mathrm{BS}}}_{f,t}S_t\,dN_t \\
&= \frac{1}{2}\mathrm{e}^{-rt}(V_t - \sigma^2_{\mathrm{BS}})S_t^2\Gamma^{\sigma_{\mathrm{BS}}}_{f,t}\,dt + \mathrm{e}^{-rt}\Delta^{\sigma_{\mathrm{BS}}}_{f,t}S_t\,dN_t.
\end{aligned}$$

Consider the self-financing portfolio with zero initial value, long one unit of the exotic option, and short $\Delta^{\sigma_{\mathrm{BS}}}_{f,t}$ units of the underlying asset. Let Π_t denote the time-t value of this portfolio. Its discounted price dynamics then satisfies

$$\begin{aligned}
d(e^{-rt}\Pi_t) &= d(e^{-rt}\pi_{f,t}) - \Delta^{\sigma_{\mathrm{BS}}}_{f,t}\left(d(e^{-rt}S_t) + \mathrm{e}^{-rt}S_t\delta\,dt\right) \\
&= d(e^{-rt}\pi_{f,t}) - \Delta^{\sigma_{\mathrm{BS}}}_{f,t}e^{-rt}S_t\,dN_t \\
&= d(e^{-rt}\pi_{f,t}) - d(e^{-rt}\pi^{\sigma_{\mathrm{BS}}}_{f,t}) + \frac{1}{2}e^{-rt}(V_t - \sigma^2_{\mathrm{BS}})S_t^2\Gamma^{\sigma_{\mathrm{BS}}}_{f,t}\,dt.
\end{aligned}$$

Integrating in t gives

$$\mathrm{e}^{-rT}\Pi_T = -\pi_{f,0} + \pi^{\sigma_{\mathrm{BS}}}_{f,0} + \frac{1}{2}\int_0^T e^{-rt}(V_t - \sigma^2_{\mathrm{BS}})S_t^2\Gamma^{\sigma_{\mathrm{BS}}}_{f,t}\,dt \tag{1.54}$$

as $\pi_{f,T} - \pi^{\sigma_{\mathrm{BS}}}_{f,T} = 0$.

We now claim that the time-0 option price $\pi_{f,0} = \pi_f$ lies between the Black–Scholes option prices for $\sigma_{\mathrm{BS}} = \sqrt{v_{min}}$ and $\sigma_{\mathrm{BS}} = \sqrt{v_{max}}$,

$$\pi^{\sqrt{v_{min}}}_{f,0} \le \pi_f \le \pi^{\sqrt{v_{max}}}_{f,0}. \tag{1.55}$$

Indeed, let $\sigma_{\mathrm{BS}} = \sqrt{v_{min}}$. Because $\Gamma^{\mathrm{BS}}_{f,t} \ge 0$ by assumption, it follows from (1.54) that $\mathrm{e}^{-rT}\Pi_T \ge -\pi_{f,0} + \pi^{\sqrt{v_{min}}}_{f,0}$. Absence of arbitrage implies that Π_T must not be bounded away from zero, hence $-\pi_{f,0} + \pi^{\sqrt{v_{min}}}_{f,0} \le 0$. This proves the left inequality in (1.55). The right inequality follows similarly, whence the claim (1.55) is proved.

A similar argument shows that the Black–Scholes price $\pi^{\sigma_{\mathrm{BS}}}_{f,0}$ is non-decreasing in σ_{BS}, whence $\sqrt{v_{min}} \le \sigma_{\mathrm{IV}} \le \sqrt{v_{max}}$, and the theorem is proved.

1.8 Conclusion

In this chapter we have introduced the Jacobi model and shown that it is a highly tractable and versatile stochastic volatility model. It contains the Heston stochastic volatility model as a limit case. The moments of the finite dimensional distributions of the log prices can be calculated explicitly thanks to the polynomial property of the model. As a result, the series approximation techniques based on the Gram–Charlier A expansions of the joint distributions of finite sequences of log returns allow us to efficiently compute prices of options whose payoff depends on the underlying asset price at finitely many time points. Compared to the Heston model, the Jacobi model offers additional flexibility to fit a large range of Black–Scholes implied volatility surfaces. Our numerical analysis shows that the series approximations of European call, put and digital option prices in the Jacobi model are computationally comparable to the widely used Fourier transform techniques for option pricing in the Heston model. The truncated series of prices, whose computations do not require any numerical integration, can be implemented efficiently and reliably up to orders that guarantee accurate approximations as shown by our numerical analysis. The pricing of forward start options, which does not involve any numerical integration, is significantly simpler and faster than the iterative numerical integration method used in the Heston model. The minimal and maximal volatility parameters are universal bounds for Black–Scholes implied volatilities and provide additional stability to the model. In particular, Black–Scholes implied volatilities of forward start options in the Jacobi model do not experience the explosions observed in the Heston model. Furthermore, our density approximation technique in the Jacobi model circumvents some limitations of the Fourier transform techniques in affine models and allows us to price discretely monitored Asian options.

2 Option Pricing with Orthogonal Polynomial Expansions

We refine the series representations for option prices presented in Chapter 1 by letting the auxiliary density be a mixture distribution. We show how to efficiently compute the polynomial basis and Fourier coefficients for various choices of mixture components. The log price density of univariate diffusive stochastic volatility models is given by a Gaussian mixture density with a continuum of components which can be discretized and then used as auxiliary density to approximate option prices. We indicate that the computed option price approximations can be accurate even when the log price density does not belong to the relevant weighted space. Numerical examples with the Jacobi and Stein-Stein models illustrate the improvement in approximation accuracy and the applicability of this method to other stochastic volatility models.

2.1 Introduction

In this chapter we improve the option pricing technique described in Chapter 1 where option prices have series representations when the likelihood ratio $\ell = g/w$ of the log price density g with respect to the auxiliary density w belongs to the weighted Lebesgue space L^2_w. In this case, and when the discounted payoff function f also belongs to L^2_w, the corresponding option price has a series representation of the form $\pi_f = \sum_{n\geq 0} f_n \ell_n$ where the coefficients f_n and ℓ_n are explicit. An option price approximation is then obtained by truncating this series at a finite order N. This option price approximation is accurate for a small truncation order N when the true density g is statistically close to the auxiliary density w. Indeed, the density w is the initial approximation of g. We thus let the auxiliary density w be a mixture distribution, that is a convex combination of K component densities v_k, in order to gain flexibility in the construction of w. We show how the orthonormal polynomial basis H_n with respect to the mixture w can be recursively computed when the orthonormal polynomial basis of its components are known. The likelihood coefficients $\ell_n = \langle \ell, H_n \rangle_w$ are then given explicitly for polynomial models. Similarly, we show that the payoff coefficients $f_n = \langle f, H_n \rangle_w$ can be efficiently computed when the payoff coefficients $f^k_m = \langle f, H_m \rangle_{v_k}$ are available for each component k and order $m \leq n$. The payoff coefficients are, in particular, given by closed-form recursive systems of equations for the European call options with the Gaussian and Gamma auxiliary densities. As a consequence, we can construct mixture distributions that better approximate the log price density and, therefore, we can efficiently computed fast converging option price approximations.

We then study a particular class of stochastic volatility models in which the squared volatility is given by a polynomial in a univariate polynomial diffusion. This class nests multiple models of interest having bounded and unbounded volatility support. For examples it includes the Jacobi, Heston, Stein-Stein, and Hull-White models. We show that the log price distribution g for these models is given by a Gaussian mixture with an infinite number of components. Letting w be a finite Gaussian mixture, we show that $\ell \in L^2_w$ under some conditions on the dynamics of the log price X_t and the volatility V_t, and on the parameters of w. Even when ℓ does not belong to L^2_w so that the option price approximation $\sum_{n=0}^N f_n \ell_n$ does not converge as $N \to \infty$, we demonstrate that this series truncated at some finite order N can provide an accurate approximation of the true price π_f. We also provide simple algorithms to construct the Gaussian mixture to be used as auxiliary density.

We validate our approach on several use cases. In the Jacobi model with a Gaussian auxiliary density, the accuracy of the option price approximation for a fixed N decreases rapidly as the upper bound of the volatility support increases. We therefore let the auxiliary density w be a mixture of two Gaussian distributions whose first two moments are matching the log price density. We show that the option price series converges significantly faster using this Gaussian mixture density in comparison to a single Gaussian density. This allows to approximate option prices rapidly and accurately for any parameters choice. We hence calibrate the Jacobi model on a volatility surface sample and report a twice smaller implied volatility root mean squared error than with the Heston model. Next, we approximate the option prices in the Stein-Stein model using a Gaussian mixture with different number of components and parameters choices, and compare them to the true option price that can be computed using Fourier techniques. Despite the fact that the likelihood ratio in the Stein-Stein model does not belong to L^2_w when w is a Gaussian mixture density, we find that our approach still produces accurate option price approximations.

Section 2.2 discusses the density expansion and option price series representation with an auxiliary mixture distribution. Section 2.3 presents the polynomial stochastic volatility models and describe some Gaussian mixture constructions for the auxiliary density. Section 2.4 contains the numerical applications. Alternative moments based methods to construct orthonormal polynomial basis can be found in Section 2.5. The proofs are collected in Section 2.6. Section 2.7 concludes.

2.2 Polynomial Price Series Expansions

We first recap the density expansion approach described in (Filipović, Mayerhofer, and Schneider 2013) along with the option price series representation further developed in Chapter 1. We then show how the elements of these option price series can be efficiently computed when the auxiliary density is a mixture distribution. We conclude the section by giving some examples of component densities which are convenient to work with.

2.2.1 European Option Price Series Representation

Fix a maturity $T > 0$ and assume that the log price X_T admits a density $g(x)$ at the initial time. We are interested in computing the option price

$$\pi_f = \mathbb{E}\left[f(X_T)\right] = \int_{\mathbb{R}} f(x)g(x)dx \tag{2.1}$$

for some discounted payoff function $f(x)$. The expectation is always taken with respect to the risk-neutral measure. Let $w(x)$ be an auxiliary density measure that dominates $g(x)$

$$w(x) = 0 \quad \Rightarrow \quad g(x) = 0$$

and denote $\ell(x)$ the likelihood ratio given by

$$g(x) = \ell(x)\,w(x).$$

We define the weighted Lebesgue space,

$$L^2_w = \left\{ f(x) \mid \|f\|^2_w = \int_E f(x)^2 w(x)dx < \infty \right\}$$

which is a Hilbert space with the scalar product

$$\langle f, g\rangle_w = \int_E f(x)g(x)w(x)dx.$$

Assume that the polynomials are dense in L^2_w and that the functions $\ell(x)$ and $f(x)$ are in L^2_w. Then, the option price has the following series representation

$$\pi_f = \sum_{n\geq 0} f_n \ell_n \tag{2.2}$$

with the *likelihood coefficients*

$$\ell_n = \langle \ell, H_n\rangle_w = \int_E H_n(x)g(x)dx, \tag{2.3}$$

and the *payoff coefficients*

$$f_n = \langle f, H_n\rangle_w = \int_E H_n(x)w(x)dx, \tag{2.4}$$

and where $H_n(x)$ denotes the orthonormal polynomials basis of L^2_w such that $\deg H_n(x) = n$ and $\langle H_m, H_n\rangle_w = 1$ if $m = n$ and zero otherwise. In practice we approximate the option price by truncating the option price series in (2.2) as follows

$$\pi_f^{(N)} = \sum_{n=0}^{N} f_n \ell_n \tag{2.5}$$

for some positive integer N. The accuracy of such approximation with a low order N crucially depends on the statistical distance between the true density and the auxiliary density. For example if $w(x) = g(x)$ then $\pi_f = f_0\ell_0 = f_0$, as $\ell_0 = 1$ and $\ell_n = 0$ for all $n \geq 1$. The efficiency of this approach depends on how fast the operations required to compute the coefficients in (2.3) and (2.4) can be performed.

In what follows we always have in mind that the log price X_t, jointly with another process $Z_t \in \mathbb{R}^m$, has the polynomial property. Fix the order N, this in particular implies that the payoff coefficients in (2.3) can be computed explicitly as follows

$$\ell_n = \mathbb{E}[H_n(X_T)] = (1, \mathscr{H}(X_0, Z_0))\mathrm{e}^{GT}\vec{H}_n, \quad n = 0, \ldots, N \tag{2.6}$$

where the polynomials $(1, \mathscr{H}(x, z))$ forms a basis of $\mathrm{Pol}_N(\mathbb{R}^{1+m})$, G is the matrix representation of the infinitesimal generator of (X_t, Z_t) in this basis, and $\vec{H}_n$ is the vector representation of the polynomial $H_n(x)$ in this basis.

Remark 2.2.1. *The options Greeks are computed by differentiating the option price with respect to one or multiple variables. For the sensitivity analysis we fix the auxiliary density $w(x)$, hence the basis $H_n(x)$ and the coefficients f_n, and let only $\ell(x)$ through $g(x)$ depend on the perturbed parameters. The sensitivity of π_f with respect to the variable y is hence given by*

$$\partial_y \pi_f = \sum_{n \geq 0} \ell_n \partial_y f_n + f_n \partial_y \ell_n$$

with the partial derivative $\partial_y = \partial / \partial y$. The sensitivity of ℓ_n with respect to y is given by

$$\partial_y \ell_n = (0, \partial_y \mathscr{H}(X_0, Z_0))\, \mathrm{e}^{GT}\, \vec{H}_n + (1, \mathscr{H}(X_0, Z_0))\, \partial_y \mathrm{e}^{GT}\, \vec{H}_n.$$

The derivative of the exponential operator e^{GT} with respect to y is given by

$$\frac{\partial \mathrm{e}^{GT}}{\partial y} = \int_0^1 \mathrm{e}^{xGT}\, \frac{\partial GT}{\partial y}\, \mathrm{e}^{(1-x)GT}\, dx$$

as proved in (Wilcox 1967). In particular for the Delta, which is the derivative of option price with respect to $y = \exp(X_0)$, we have that $\partial_y \mathrm{e}^{GT} = 0$.

2.2.2 Auxiliary Mixture Distribution

From now on we let the auxiliary density $w(x)$ be a mixture distribution as defined by

$$w(x) := \sum_{k=1}^{K} c_k v_k(x)$$

for some mixture weights $c_k > 0$ satisfying $\sum_{k=1}^{K} c_k = 1$, and some mixture components $v_k(x)$ which are also probability densities. To each density $v_k(x)$ is associated an orthonormal polynomial basis $H_n^k(x)$, $n \geq 0$. Let a_n^k and b_n^k denote the coefficients that define the recurrence relation for $H_n^k(x)$, which always holds for univariate bases,

$$xH_n^k(x) = b_{n+1}^k H_{n+1}^k(x) + a_n^k H_n^k(x) + b_n^k H_{n-1}^k(x) \tag{2.7}$$

for all $n \geq 0$ with $H_{-1}^k = 0$ and $H_0^k = 1$. We define the following matrices

$$J_N^k = \begin{pmatrix} a_0^k & b_1^k & & & \\ b_1^k & a_1^k & b_2^k & & \\ & \ddots & \ddots & \ddots & \\ & & b_{N-1}^k & a_{N-1}^k & b_N^k \\ & & & b_N^k & a_N^k \end{pmatrix}, \quad \text{for } k = 1, \ldots, K.$$

The recurrence relation (2.7) to construct the orthonormal polynomial basis is explicitly known for many densities taking values both on compact and unbounded supports, see (Schoutens 2012, Chapter 1) for an overview of them. For example, when v_k is a Gaussian density with mean μ_k and variance σ_k^2, the coefficients in (2.7) are given by $a_n^k = \mu_k$ and $b_n^k = \sqrt{n}\sigma_k$.

Remark 2.2.2. *The coefficients a_n^k and b_n^k can be inferred from the basis $H_n^k(x)$. Denote $\alpha_{n,i}^k$ the coefficient in front of the monomial x^i of the polynomial $H_n^k(x)$. Then by inspection of Equation* (2.7) *we have*

$$b_{n+1}^k = \frac{\alpha_{n,n}^k}{\alpha_{n+1,n+1}^k} \quad \text{and} \quad a_n^k = \frac{\alpha_{n,n-1}^k - b_{n+1}^k \alpha_{n+1,n}^k}{\alpha_{n,n}^k}.$$

The following Proposition gives an algorithm to compute the coefficients a_n and b_n in the recursion of the orthonormal polynomial basis $H_n(x)$ associated with the mixture distribution $w(x)$,

$$xH_n(x) = b_{n+1}H_{n+1}(x) + a_n H_n(x) + b_n H_{n-1}(x)$$

for all $n \geq 0$ with $H_{-1} = 0$ and $H_0 = 1$.

Proposition 2.2.3. *The recurrence relation coefficients of the mixture distribution are given by*

$b_n = \sqrt{\psi_n/\psi_{n-1}}$ for $n = 1,\dots,N$ and $a_n = \phi_n/\psi_n$ for $n = 0,\dots,N-1$ where

$$\begin{aligned}\psi_n &= \sum_{k=1}^{K} c_k\,(z_n^k)^\top z_n^k\\ \phi_n &= \sum_{k=1}^{K} c_k\,(z_n^k)^\top J_N^k\, z_n^k\\ z_{n+1}^k &= (J_N^k - a_n I) z_n^k - (b_n)^2 z_{n-1}^k \quad \text{for all } k = 1,\dots,K\end{aligned}$$

with $z_{-1}^k = 0$, and $z_0^k = \mathrm{e}_1$ is the vector whose first coordinate is equal to one and zero otherwise.

This algorithm is fast and performs well numerically. For example, with $N, K \sim 10^2$, it takes few milliseconds on a modern CPU to construct the orthonormal basis. There are others moments based approaches to construct the orthonormal basis, for examples the Gram-Schmidt and the Mysovskikh algorithms are described in Section 2.5. However these methods are typically subject to numerical problems and may be slow even for a relatively small order N.

The following proposition shows that the payoff coefficients with respect to the mixture distribution can efficiently be computed when the corresponding coefficients are known for the mixture components.

Proposition 2.2.4. *The payoff coefficients are equal to*

$$f_N = \langle f, H_N\rangle_w = \sum_{k=1}^{K}\sum_{n=0}^{N} c_k\, q_{N,n}^k\, f_n^k \quad \text{with } f_n^k = \langle f, H_n^k\rangle_{v_k} \tag{2.8}$$

and where $q_N^k \in \mathbb{R}^N$ is the vector representation of $H_N(x)$ in the basis $H_n^k(x)$

$$H_N(x) = \sum_{n=0}^{N} q_{N,n}^k\, H_n^k(x), \quad k = 1,\dots,K. \tag{2.9}$$

The usefulness of this Proposition lies on the premise that the coefficients f_n^k can easily be computed. In a situation where they are numerically costly to compute then one may directly integrate the payoff function with respect to the following density approximation

$$w^{(N)}(x) = \sum_{n=0}^{N} \ell_n H_n(x) \sum_{k=1}^{K} c_k v_k(x). \tag{2.10}$$

Remark 2.2.5. *The vectors q_N^k can be computed efficiently. Let $\mathbf{H}_N^k \in \mathbb{R}^{(N+1)\times(N+1)}$ denote the matrix whose (i,j)-th element is given by the coefficient in front of the monomial x^{j-1} in the $(i-1)$-th polynomial of the basis $H_n^k(x)$. Define similarly the matrix $\mathbf{H}_N$ with respect to the polynomial basis $H_n(x)$. The matrices $\mathbf{H}_N^k$ for $k = 1,\dots,K$, and $\mathbf{H}_N$ are upper triangular. We are interested in the upper triangular matrix $\mathbf{Q}_N^k \in \mathbb{R}^{(N+1)\times(N+1)}$ for $k = 1,\dots,K$ whose (i,j)-th element is equal to $q_{j,i}^k$. It is also given by $\mathbf{H}_N = \mathbf{H}_N^k \mathbf{Q}_N^k$ which forms a triangular system of equations and can thus be solved efficiently.*

2.2.3 Component Density Examples

In practice, to ensure efficient option prices approximations we need to select components $v_k(x)$ whose orthonormal polynomial basis $H_n^k(x)$ and payoff coefficients f_n^k can efficiently be computed. The Gaussian and uniform distributions have been successfully used in Chapter 1 and Chapter 3 respectively. The logistic distribution whose tailed decreases at an exponential-linear rate is used in (Heston and Rossi 2016) but for which the payoff coefficients are given by complex expressions involving special functions. The bilateral Gamma distribution of (Küchler and Tappe 2008) was used in (Filipović, Mayerhofer, and Schneider 2013) to accurately approximate option prices by numerical integration of the discounted payoff function with respect to the density approximation. However an explicit recursion formula to construct the orthonormal polynomial basis was not provided.

We study hereinbelow the Gamma distribution whose single tail decays at a polynomial-exponential-linear rate and for which simple recursive expressions can be derived for the payoff coefficients. By mixing two Gamma distributions one can obtain a density distribution on the entire real line. The Gamma distribution on the half-line $(\xi,+\infty)$ for some $\xi\in\mathbb{R}$ is defined by

$$v_k(x) = \mathbb{1}_{\{x>\xi\}}\frac{\beta^\alpha}{\Gamma(\alpha)}(x-\xi)^{\alpha-1}\mathrm{e}^{-\beta(x-\xi)} \tag{2.11}$$

for some shape parameter $\alpha\geq 1$, rate parameter $\beta>0$, and where $\Gamma(\alpha)$ is the upper incomplete Gamma function defined by

$$\Gamma(\alpha)=\Gamma(\alpha,0) \quad \text{with} \quad \Gamma(\alpha,z)=\int_z^\infty x^{\alpha-1}\mathrm{e}^{-x}dx$$

such that $\Gamma(n)=(n-1)!$ for any positive integer n. The Gamma distribution $v_k(x)$ admits an orthonormal polynomial basis $H_n^k(x)$ given by

$$H_n^k(x)=\sqrt{\frac{n!}{\Gamma(\alpha+n)}}\,\mathscr{L}_n^{\alpha-1}(\beta(x-\xi))$$

where $\mathscr{L}_n^{\alpha-1}(x)$ denotes the n-th order generalized Laguerre polynomial with parameter $\alpha-1$ and defined by

$$\mathscr{L}_n^{\alpha-1}(x)=\frac{x^{-\alpha+1}\mathrm{e}^x}{n!}\frac{\partial^n}{\partial x^n}(\mathrm{e}^{-x}x^{\alpha-1+n}).$$

The generalized Laguerre polynomials are recursively given by

$$\begin{aligned}
\mathscr{L}_0^{\alpha-1}(x)&=1\\
\mathscr{L}_1^{\alpha-1}(x)&=\alpha+x\\
\mathscr{L}_{n+1}^{\alpha-1}(x)&=\frac{2n+\alpha-x}{n+1}\mathscr{L}_n^{\alpha-1}(x)-\frac{n+\alpha-1}{n+1}\mathscr{L}_{n-1}^{\alpha-1}(x)\quad\text{for all } n\geq 1.
\end{aligned}$$

The following theorem shows that the payoff coefficients f_n^k can be computed recursively for the Gamma distribution $v_k(x)$.

Theorem 2.2.6. *Consider the discounted payoff function of a call option with log strike k,*

$$f(x) = \mathrm{e}^{-rT}\left(\mathrm{e}^{x} - \mathrm{e}^{k}\right)^{+}.$$

Its payoff coefficients f_n are given by

$$f_n = \sqrt{\frac{n!}{\Gamma(\alpha+n)}}\frac{1}{\Gamma(\alpha)}\left(\mathrm{e}^{\xi} I_n^{\alpha-1}\left(\mu;\beta^{-1}\right) + \mathrm{e}^{k} I_n^{\alpha-1}\left(0;\beta^{-1}\right)\right), \quad n \geq 0.$$

with $\mu = \max(0, \beta(k-\xi))$ and where the functions $I_n^{\alpha-1}(\mu;\nu)$ are recursively defined by

$$\begin{aligned}
I_0^{\alpha-1}(\mu;\nu) &= (1-\nu)^{-\alpha}\,\Gamma(\alpha)\,\Gamma(\alpha,\mu(1-\nu)) \\
I_1^{\alpha-1}(\mu;\nu) &= \alpha I_0^{\alpha-1}(\mu;\nu) + I_0^{\alpha}(\mu;\nu) \\
I_n^{\alpha-1}(\mu;\nu) &= \left(2 + \frac{\alpha-2}{n}\right) I_{n-1}^{\alpha-1}(\mu;\nu) - \left(1 + \frac{\alpha-2}{n}\right) I_{n-2}^{\alpha-1}(\mu;\nu)\cdot \\
&\quad - \frac{1}{n}\left(I_{n-1}^{\alpha}(\mu;\nu) - I_{n-2}^{\alpha}(\mu;\nu)\right), \quad n \geq 2.
\end{aligned} \tag{2.12}$$

Note that the calculation of the term $I_n^{\alpha-1}(\mu;\nu)$ requires calculations of the terms $I_{n-1}^{\alpha}(\mu;\nu)$ and therefore the dimension of the recursive system grows at the rate n^2.

The Gamma distribution on the half-line $(-\infty,\xi)$ for some $\xi \in \mathbb{R}$ together with its polynomial basis and Fourier coefficients can be similarly derived. A Gamma mixture on $\mathbb{R}$ can thus be constructed when $K \geq 2$ by letting, for example, the first component support be $(-\infty,\xi]$ and the second $[\xi,\infty)$.

2.3 Polynomial Stochastic Volatility Models

We present a class of stochastic volatility models for which the log price distribution is given by a Gaussian mixture with an infinite number of components. We then describe a simple methodology to approximate the log price density by a Gaussian mixture with a finite number of components. The section terminates by studying the option price approximation error when the likelihood ratio $\ell = g/w$ does not belong to the weighted space L^2_w.

2.3.1 Definition and Basic Properties

We fix a stochastic basis $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{Q})$ where $\mathscr{F}_t$ is the filtration generated by two independent Brownian motions W_{1t} and W_{2t}. Let $Y_t \in E_Y \subset \mathbb{R}$ be an autonomous polynomial diffusion whose dynamics is given by

$$dY_t = \kappa(\theta - Y_t)\,dt + \sigma(Y_t)\,dW_{1t} \quad \text{with} \quad \sigma(y)^2 = \alpha + ay + Ay^2 \tag{2.13}$$

for some real parameters $\kappa, \theta, \alpha, a, A$ such that for any $Y_0 \in E_Y$ a solution of (2.13) taking values in E_Y exists. We then define the dynamics of the log price X_t as follows

$$dX_t = (r - \delta)\,dt - \frac{1}{2}d\langle X\rangle_t + \Sigma_1(Y_t)\,dW_{1t} + \Sigma_2(Y_t)\,dW_{2t}$$

for the interest rate r, the dividend yield $\delta \geq 0$, and such that $\Sigma_1(y)^2 + \Sigma_2(y)^2 \in \mathrm{Pol}_m$, $\Sigma_2(y)^2 \in \mathrm{Pol}_{m+1}$, and $\Sigma_1(y)\sigma(y) \in \mathrm{Pol}_{m+1}$ for some $m \in \mathbb{N}$. The process (X_t, Z_t) is hence a polynomial diffusion where $Z_t = (Y_t, Y_t^2, \ldots, Y_t^m)$ since Y_t is an autonomous polynomial diffusion and so must be Z_t, see (Filipović and Larsson 2017). Some classical stochastic volatility models are nested in this setup, for examples, the Heston, Jacobi, Stein-Stein, and Hull-White models.

The volatility of the asset log price is $V_t = \sqrt{d\langle X\rangle_t / dt} = \sqrt{\Sigma_1(Y_t)^2 + \Sigma_2(Y_t)^2}$. The leverage effect refers to the generally negative correlation between dV_t^2 and dX_t and is given by

$$\begin{aligned} \mathrm{lev}(X_t) &= \frac{d\langle V^2, X\rangle_t}{\sqrt{d\langle V^2\rangle_t}\,\sqrt{d\langle X\rangle_t}} \\ &= \frac{\Sigma_1(Y_t)}{\sqrt{\Sigma_1(Y_t)^2 + \Sigma_2(Y_t)^2}}\,\mathrm{sign}\left[\left(\Sigma_1'(Y_t)\Sigma_1(Y_t) + \Sigma_2'(Y_t)\Sigma_2(Y_t)\right)\sigma(Y_t)\right]. \end{aligned} \tag{2.14}$$

The volatility of the volatility of the asset log price is

$$\mathrm{volvol}(X_t) = \sqrt{\frac{d\langle V\rangle_t}{dt}} = \frac{1}{2V_t}\sqrt{\frac{d\langle V^2\rangle_t}{dt}} = \sqrt{\frac{\left(\Sigma_1'(Y_t)\Sigma_1(Y_t) + \Sigma_2'(Y_t)\Sigma_2(Y_t)\right)^2 \sigma(Y_t)^2}{\Sigma_1(Y_t)^2 + \Sigma_2(Y_t)^2}} \tag{2.15}$$

The proofs of (2.14)–(2.15) are given in Section 2.6.

We fix a finite time horizon $T > 0$. The following proposition shows that the distribution of the log price is given by a Gaussian mixture density with an infinite number of components, see (McNeil, Frey, and Embrechts 2015, Chapter 6.2).

Proposition 2.3.1. *The distribution of X_T conditional on the trajectory of W_{1t} on $[0, T]$ is normally distributed with mean*

$$M_T = X_0 + (r - \delta)T - \frac{1}{2}\int_0^T \left(\Sigma_1(Y_t)^2 + \Sigma_2(Y_t)^2\right)dt + \int_0^T \Sigma_1(Y_t)dW_{1t} \tag{2.16}$$

and variance

$$C_T = \int_0^T \Sigma_2(Y_t)^2 dt. \tag{2.17}$$

Therefore, when it is well defined, the log price density is of the form

$$g(x) = \mathbb{E}\left[(2\pi C_T)^{-\frac{1}{2}} \exp\left(-\frac{(x-M_T)^2}{2C_T}\right)\right]. \tag{2.18}$$

Similar expressions for the log price density have been previously derived in (Lipton and Sepp 2008) and in (Glasserman and Kim 2011). In practice, we want to approximate the log price density $g(x)$ by a Gaussian mixture density $w_K(x)$ with K components that will in turn be used as auxiliary density to derive option price approximations,

$$w_K(x) = \sum_{k=1}^{K} c_k \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{(x-\mu_k)^2}{2\sigma_k^2}\right) \tag{2.19}$$

for some mixture weights $c_k > 0$ such that $\sum_{k=1}^{K} c_k = 1$, and some constants $\mu_k \in \mathbb{R}$ and $\sigma_k > 0$ for $k = 1, \dots, K$. In Section 2.3.2 we suggest a computationally efficient approach based on weighted simulations of the first Brownian motion W_{1t}.

From Chapter 1 we known that under certain conditions the likelihood ratio $\ell(x)$ belongs to a weighted Lebesgue space $L^2_{w_K}$ where w_K is a Gaussian mixture.

Corollary 2.3.2. *Assume that there exist two constants $C_1, C_2 > 0$ such that*

$$|M_T| < C_1 \text{ and } C_T < C_2 \tag{2.20}$$

then $\ell \in L^2_{w_K}$ with w_K as defined in (2.19) *if $\sigma_k > C_2/2$ for some k in $1, \dots, K$. Sufficient conditions for* (2.20) *to hold are that $y_{min} < Y_t < y_{max}$ for all $t \geq 0$ and some constants y_{min}, y_{max}, and that $\Sigma_1(y)/\sigma(y) \in \mathrm{Pol}_m$ for some m.*

2.3.2 Gaussian Mixture Specification

We present an efficient approach to specify Gaussian mixtures for the auxiliary density. It is based on the discretization of the single source of randomness affecting M_T and C_T in (2.16)-(2.17): the trajectory of W_1 on $[0, T]$. Fix a time grid $0 = t_0 < t_1 < t_2 < \cdots < t_n = T$ with constant step size $t_{i+1} - t_i = \Delta t$. Then, apply a Monte-Carlo approach to obtain K (weighted) vectors $Z^{(k)} \in \mathbb{R}^n$ of normally distributed Brownian increments

$$Z^{(k)} \sim \left(\Delta W_{1,t_1}, \dots, \Delta W_{1,t_n}\right) \sim \mathcal{N}\left(\mathbf{0}_n, \mathrm{diag}(\Delta t \mathbf{1}_n)\right) \tag{2.21}$$

where $\Delta W_{1,t_i} = W_{1,t_i} - W_{1,t_{i-1}}$. The stochastic differential equations (SDE) of (Y_t, M_t, C_t) can then be numerically integrated to obtain K triplets $(c_k, M_T^{(k)}, C_T^{(k)})$ where c_k is the weight associated to Z^k. We can then let these triplets be the parameters (c_k, μ_k, σ_k^2) of the Gaussian mixture in (2.19).

A standard simulation with standard i.i.d. simulations such that $c_k = 1/K$ would be costly because the number of triplets K required to obtain a good approximation of $g(x)$ may be large. In addition, a large K makes the computation of the coefficients f_n and ℓ_n computationally more demanding. Therefore, a weighted Monte-Carlo method appears to be a more sensible approach to parametrize the Gaussian mixture. For example, one can use an optimal K-quantization of the multivariate Gaussian distribution. This is a discrete probability distribution in $\mathbb{R}^n$ with K mass points that best approximates the multivariate normal (2.21) in the L^2 sense, we refer to (Pagès and Printems 2003) for more details. The K pairs $(c_k, Z^{(k)})$ can always be precomputed and loaded on demand since they do not depend on the stochastic volatility model at hand.

The discretization scheme used to numerically integrate the SDE may also play an important role. We want a scheme that performs well with a large time step Δt since we are only interested in the log price density at time T. Numerical experiments on multiple models showed that good results can notably be obtained with the Interpolated-Kahl-Jäckel (IJK) scheme introduced in (Kahl and Jäckel 2006). The IJK scheme boils down to a scheme with linear interpolation of the drift of the diffusion, consideration for the correlated diffusive terms, and with a higher order Milstein term. It bears little additional computational cost.

2.3.3 Nonconvergent Option Price Approximations

We provide some explanations why, even when $\ell \notin L^2_w$, the price approximation $\pi_f^{(N)}$ in (2.5) can be an accurate approximation of the option price π_f. More precisely, we estimate the option price approximation error.

Assume that instead of considering the true process X_t we consider the modified process X_t^τ whose dynamics is

$$dX_t^\tau = (r-\delta)dt - \frac{1}{2}d\langle X^\tau\rangle_t + 1_{\{\tau>t\}}\left(\Sigma_1(Y_t)\,dW_{1t} + \Sigma_2(Y_t)\,dW_{2t}\right)$$

where the stopping time τ is defined by

$$\tau = \inf\{t \geq 0 : |M_t| \geq C_1\} \wedge \inf\{t \geq 0 : C_t \geq C_2\}$$

for some positive constants C_1 and C_2. The event $\{\tau < T\}$ is thought be very unlikely so that $\mathbb{P}[X_T = X_T^\tau] > 1-\epsilon$ for some small $\epsilon > 0$. Note that the modified discounted cum-dividend stock price $\mathrm{e}^{-(r-\delta)t + X_t^\tau}$ remains a martingale. As a consequence of this construction, Corollary 2.3.2 applies to the process X_t^τ which implies that the option price has a series representation for some Gaussian mixture density w,

$$\pi_f^\tau = \mathbb{E}[f(X_T^\tau)] = \sum_{n=0}^{\infty} f_n \ell_n^\tau, \tag{2.22}$$

where the payoff coefficients f_n are defined as in (2.4) and the likelihood coefficients ℓ_n^τ are defined by

$$\ell_n^\tau = \mathbb{E}[H_n(X_T^\tau)]. \tag{2.23}$$

However, we do not know the moments of X_T^τ and we will use instead the moments of X_T. We hence obtain the option price approximation

$$\pi_f^{(N)} = \sum_{n=0}^N f_n \ell_n,$$

which is similar to (2.5) but with the important difference that it will not converge to the true price π_f. The pricing error with this approximation can be decomposed into three terms,

$$\begin{aligned}\left|\pi_f - \pi_f^{(N)}\right| &\le \left|\mathbb{E}[(f(X_T) - f(X_t^\tau))1_{\{\tau\le T\}}]\right| && (2.24)\\ &+ \left|\pi_f^\tau - \sum_{n=0}^N f_n \ell_n^\tau\right| && (2.25)\\ &+ \left|\sum_{n=0}^N f_n(\ell_n - \ell_n^\tau)\right|. && (2.26)\end{aligned}$$

The first error term (2.24) is the difference between the option price with X_T and with X_T^τ. It can easily be controlled. For example if $|f(x)|$ is bounded by K on $\mathbb{R}$ then we have

$$\left|\mathbb{E}[(f(X_T) - f(X_T^\tau))1_{\{\tau\le T\}}]\right| \le 2K\epsilon.$$

The difference between the call option prices can also be bounded by using the put-call parity given that $\mathbb{E}[\mathrm{e}^{X_T}] = \mathbb{E}[\mathrm{e}^{X_T^\tau}]$. The second term (2.25) is the option price approximation error for the log price X_T^τ that converges to zero as $N \to \infty$. The third error term (2.26) is the difference between the option price approximations of X_T and X_T^τ, that we expect to be small for a low order N but will typically diverge as $N \to \infty$. This last error term can be further divided

$$\begin{aligned}\left|\sum_{n=0}^N f_n(\ell_n - \ell_n^\tau)\right| &= \left|\mathbb{E}\left[\sum_{n=0}^N f_n(H_n(X_T) - H_n(X_T^\tau))1_{\{\tau\le T\}}\right]\right|\\ &\le \left|\mathbb{E}\left[\sum_{n=0}^N f_n H_n(X_T)1_{\{\tau\le T\}}\right]\right| + \left|\mathbb{E}\left[\sum_{n=0}^N f_n H_n(X_T^\tau)1_{\{\tau\le T\}}\right]\right| && (2.27)\end{aligned}$$

The first term on the right side of (2.27) is precisely the one that is not expected to converge as $N \to \infty$. Applying the Cauchy-Schwarz inequality we derive the upper bound

$$\left|\mathbb{E}\left[\sum_{n=0}^N f_n H_n(X_T)1_{\{\tau\le T\}}\right]\right| \le \sqrt{\mathbb{E}[p_N(X_T)]\epsilon}, \quad \text{where} \quad p_N(x) = \left(\sum_{n=0}^N f_n H_n(x)\right)^2$$

which can be computed explicitly. In practice, this bound may give an indication to whether the option price approximation $\pi_f^{(N)}$ is reasonable.

Remark 2.3.3. *A simple trick to stabilize the option price approximation when Corollary 2.3.2* *does not apply is to match the N^*-th moment of the log price and the auxiliary density, $E[X_T^{N^*}] = \int_{\mathbb{R}} x^{N^*} w(x)dx$. By doing so, the first N^* moments of $g(x)$ and $w(x)$ will be of the same magnitudes which should result in well-behaved values for the ℓ_n coefficients. This can be achieved, for example, by including a component with small mixture weight and which is used to tune the N^*-th moment.*

2.4 Numerical Analysis

In this section we show that a simple Gaussian mixture with two components can be used to improve the convergence rate of the option price approximations in the Jacobi model. We use this technique to swiftly calibrate the model on a sample of option prices. We then derive accurate option price approximations in the Stein-Stein model using Gaussian mixtures with many components for the auxiliary density.

2.4.1 Jacobi Model

The dynamics of (X_t, Y_t) in the Jacobi model is of the form

$$\begin{aligned} dY_t &= \kappa(\theta - Y_t)\,dt + \sigma\sqrt{Q(Y_t)}\,dW_{1t} \\ dX_t &= \left(r - \delta - \frac{1}{2}Y_t\right)dt + \rho\sqrt{Q(Y_t)}\,dW_{1t} + \sqrt{Y_t - \rho^2 Q(Y_t)}\,dW_{2t} \end{aligned}$$

with the function $Q(y) = \frac{(y-y_{min})(y_{max}-y)}{(\sqrt{y_{max}}-\sqrt{y_{min}})^2}$. We refer to Chapter 1 for more details on the model. We illustrate the advantages of using a Gaussian mixture with the Jacobi model when $\mathrm{var}[X_T] < y_{max}T/2$ such that the first two moments cannot be matched with a single Gaussian distribution $w_1(x) = v_1(x)$ as auxiliary density. We consider here a Gaussian mixture with two components $w_2(x)$ as auxiliary density given by

$$w_2(x) = c_1\,\frac{1}{\sqrt{2\pi\sigma_1^2}}\exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right) + (1-c_1)\,\frac{1}{\sqrt{2\pi\sigma_2^2}}\exp\left(-\frac{(x-\mu_2)^2}{2\sigma_2^2}\right)$$

for some probability $0 < c_1 < 1$, some mean parameters $\mu_1, \mu_2 \in \mathbb{R}$ and volatility parameters $\sigma_1, \sigma_2 > 0$. We match the the first two moments of X_T which gives the following underdetermined system of equations

$$\mathbb{E}[X_T] = c_1\,\mu_1 + (1-c_1)\,\mu_2 \tag{2.28}$$

$$\mathbb{E}[X_T^2] = c_1\,(\sigma_1^2 + \mu_1^2) + (1-c_1)\,(\sigma_2^2 + \mu_2^2). \tag{2.29}$$

We set $\mathbb{E}[X_T] = \mu_1 = \mu_2$ such that (2.28) is automatically satisfied and (2.29) rewrites

$$c_1 = \frac{\sigma_2^2 - \mathrm{var}[X_T]}{\sigma_2^2 - \sigma_1^2} \tag{2.30}$$

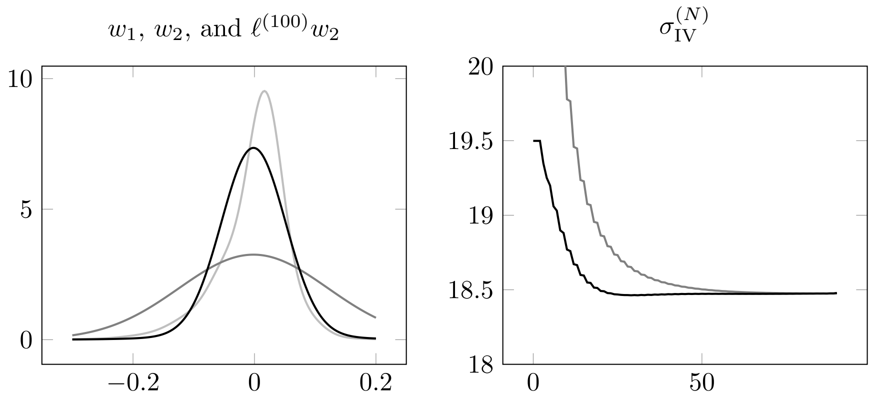


Figure 2.1 – Auxiliary densities and implied volatility convergence.
The left panel displays the Gausian mixture used asauxiliary density with one (grey) and two (black) components as well as the log price density approximation at the order $N = 100$ (light-grey). The right panel displays the implied volatility series for the corresponding two auxiliary densities for a call option with strike $k = 0$.

hence it must be that $\left|\sigma_2^2 - \text{var}[X_T]\right| < \left|\sigma_2^2 - \sigma_1^2\right|$ and that $\left(\sigma_2^2 - \text{var}[X_T]\right)\left(\sigma_2^2 - \sigma_1^2\right) \geq 0$. We set $\sigma_2 = \sqrt{y_{max}T/2} + 10^{-4}$ so that Corollary 2.3.2 applies which ensures that the option price expansion will converge to the true price. Then we arbitrarily fix $c_1 = 95\%$ and solve (2.30) to get σ_1. The reason behind these choices is that, by doing so, the mixture component with large weight c_1 is almost a Gaussian approximation of the log price density since $\mathbb{E}[X_T] = \mu_1$ and $\text{var}[X_T] \approx \sigma_1^2$. In the following numerical example we use the parameters: $r = \delta = x_0 = 0$, $\kappa = 0.5$, $\theta = Y_0 = 0.04$, $\sigma = 1$, $y_{min} = 10^{-4}$, $y_{max} = 0.36$, and $T = 1/12$. The upper bound on the volatility support is therefore 60%.

The left panel of Figure 2.1 displays the Gaussian mixture with one and two components used as auxiliary density along with the log price density approximation (2.10) at the truncation order $N = 100$. We expect that $g(x) \approx \ell^{(100)}(x)w_2(x)$. It is clear from the figure that $w_1(x)$ is a poor approximation of the density $g(x)$ whereas $w_2(x)$ appears more sensible. As a consequence, the implied volatility of a call option with log strike $k = 0$ converges significantly faster using $w_2(x)$ as auxiliary density as can be seen on the right panel of Figure 2.1. The call (put) implied volatility is initially overestimated with $w_1(x)$ since it has significantly more weight in the right (left) tail than $g(x)$. This behavior is confirmed in Table 2.1 which reports the implied volatility error with respect to the approximation at the order $N = 100$ for call options with different moneyness and for different truncation order. More problematic, the option price approximation of the far OTM option is negative between $N = 2$ and $N = 18$ with $w_1(x)$.

Equipped with a Gaussian mixture with two components, we calibrate the Jacobi model on a sample of S&P500 options and compare its suitability to fit the implied volatility surface with the Heston model. We select all the call and put options available on March 30 2017 with maturity in 1, 2, 3, or 4 weeks from the OptionMetrics database. With a linear regression we extract from the put-call parity the risk-free rate $r = 1.66\%$ and the dividend yield $\delta = 1.50\%$.

	$k=-0.1$		$k=0$		$k=0.1$	
N	$K=1$	$K=2$	$K=1$	$K=2$	$K=1$	$K=2$
0–1	23.59	3.67	26.67	1.02	25.97	3.25
2	3.77	3.67	7.47	1.02	2.88	3.25
3	2.63	1.89	7.36	0.87	-	0.17
4	7.38	1.86	4.03	0.77	-	0.03
5	5.55	1.01	4.00	0.72	-	2.47
6	5.18	0.88	2.57	0.58	-	1.80
7	4.22	0.55	2.55	0.55	-	3.00
8	3.23	0.38	1.78	0.42	-	2.10
9	2.73	0.28	1.77	0.40	-	2.32
10	1.91	0.12	1.30	0.29	-	1.66
11	1.64	0.12	1.29	0.28	-	1.57
12	1.02	0.00	0.98	0.19	-	1.15
13	0.88	0.02	0.97	0.19	-	1.02
14	0.42	0.06	0.76	0.12	-	0.78
15	0.37	0.04	0.75	0.12	-	0.67
16	0.04	0.09	0.60	0.07	-	0.53
17	0.03	0.06	0.59	0.07	-	0.44
18	0.21	0.09	0.48	0.04	5.12	0.38
19	0.18	0.07	0.47	0.03	4.43	0.31
20	0.35	0.08	0.39	0.01	3.12	0.28
30	0.39	0.00	0.15	0.01	0.49	0.04
40	0.15	0.04	0.06	0.01	0.16	0.09
50	0.02	0.04	0.02	0.01	0.32	0.10

Table 2.1 – Implied volatility errors for the Jacobi model.

The reported values are absolute percentage errors with respect to implied volatility approximations obtained at the 100-th truncation order for call options with different log strikes k. The auxiliary density is a Gaussian mixture with two components whose two first moments match those of the log price. The ”–” symbol indicates that the implied volatility was not retrievable because the option price approximation was negative.

	$\sqrt{\theta}$	κ	σ	ρ	$\sqrt{Y_0}$	$\sqrt{y_{min}}$	$\sqrt{y_{max}}$	RMSE
Jacobi	0.3035	1.8230	1.4591	-0.8177	0.0728	0.0546	0.3977	0.3102
Heston	0.2519	2.2532	0.7942	-0.6178	0.0732	–	–	0.6976

Table 2.2 – Fitted parameters for the Heston and Jacobi models.
The table reports fitted parameters and the implied volatilities root-mean-squared-error (RMSE) in percentage. The models were calibrated on a subset of S&P500 options with maturity less than one month observed on March 30 2017.

For each maturity we select a sample of 25 call options with a Delta ranging from 5% to 95%. We denote here π_{ij}, σ_{ij}, and ν_{ij} the j-th option price, implied volatility, and Vega from the i-week maturity sample. Similarly $\hat{\pi}_{ij}$ and $\hat{\sigma}_{ij}$ denote the model, Jacobi or Heston, option price and implied volatility. We calibrate the two models to the implied volatility surface by minimizing the weighted root-mean-square-error (RMSE)

$$\sqrt{\frac{1}{100}\sum_{i=1}^{4}\sum_{j=1}^{25}\left(\frac{\pi_{ij}-\hat{\pi}_{ij}}{\nu_{ij}}\right)^2}.$$

This criterion is a computationally efficient approximation for the implied volatility surface RMSE criterion which follows by observing that

$$\sigma_{ij}-\hat{\sigma}_{ij}\approx\frac{\pi_{ij}-\hat{\pi}_{ij}}{\nu_{ij}} \quad \text{when} \quad \pi_{ij}\approx\hat{\pi}_{ij}.$$

Table 2.2 reports the fitted parameters and Figure 2.2 displays the corresponding implied volatility surfaces. We observe that the values of κ, θ, and ρ are relatively similar, but the vol-of-vol parameter σ is almost twice larger for the Jacobi model which suggests that its volatility process may take large values, close to y_{max}, within little time. The fitted volatility support goes, roughly, from 5% to 40% which seems reasonable for this sample. With two additional parameters, the Jacobi model is able to better fit the implied volatility surface than the Heston model. Indeed, the resulting RMSE on the implied volatility is twice smaller for the Jacobi model. In particular, the Jacobi models seems to perform better in capturing the short-term skew and smile curvature.

2.4.2 Extended Stein-Stein Model

The Stein-Stein model was introduced by (Stein and Stein 1991) and generalized by (Schöbel and Zhu 1999) to allow for non-zero leverage. Its dynamics is of the form

$$\begin{aligned} dY_t &= \kappa(\theta - Y_t)dt + \sigma dW_{1t} \\ dX_t &= \left(r-\delta-\frac{1}{2}Y_t^2\right)dt + \rho Y_t dW_{1t} + \sqrt{1-\rho^2}\,Y_t dW_{2t} \end{aligned} \tag{2.31}$$

for some nonnegative constants $\rho \in (-1,1)$, $\kappa, \theta, \delta, r, \sigma$ and some deterministic initial values $(X_0, Y_0) \in \mathbb{R}^2$. The process X_t is the log price, as before, and the process Y_t follows an Ornstein-

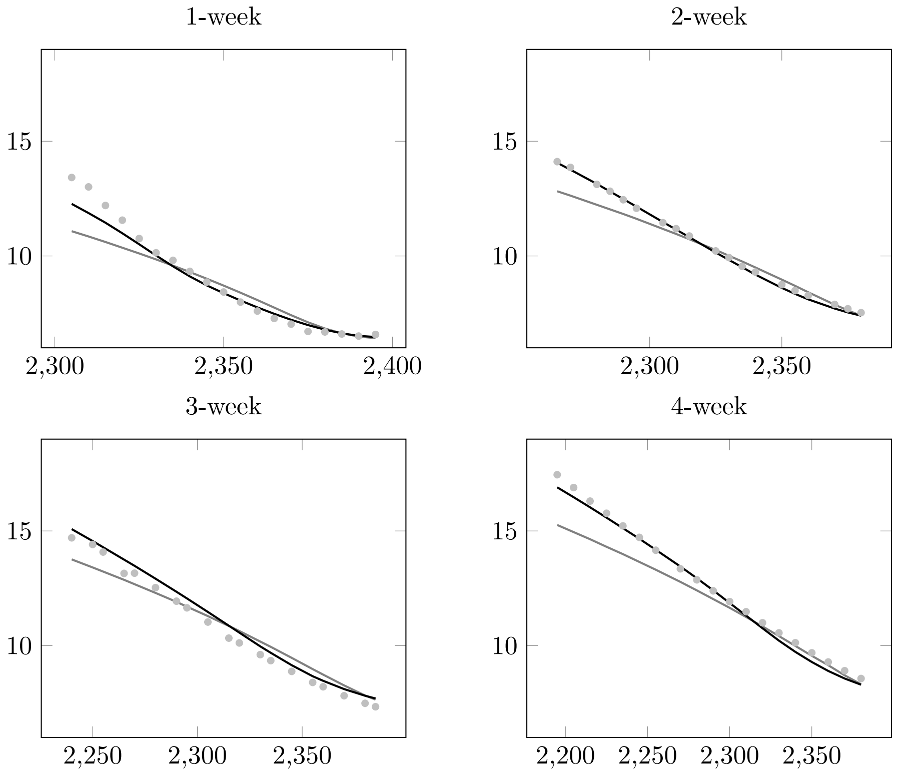


Figure 2.2 – Fitted implied volatility surfaces for the Heston and Jacobi models.

The true (dotted light-gray), the Heston (gray), and the Jacobi (black) implied volatility surfaces are displayed for each maturity as a function of the strike price. The data sample is a subset of S&P500 options with maturity less than one month observed on March 30 2017.

Uhlenbeck process. The volatility process is $V_t = |Y_t|$ since Y_t can become negative.

The Stein-Stein model has the particularity that (X_t, Y_t, Y_t^2) is a polynomial model but also an affine process. This enables the use of standard Fourier techniques to compute option prices that we use as reference option prices, see for details (Carr and Madan 1999), (Duffie, Filipović, and Schachermayer 2003), and (Fang and Oosterlee 2009). We approximate option prices using a Gaussian mixture for the auxiliary density as described in Section 2.3.2. The IJK scheme applied to the linear volatility dynamics V_t is

$$Y_{t_{i+1}} = Y_{t_i} + \kappa(\theta - Y_{t_i})\Delta t + \sigma \Delta W_{1,t_{i+1}} \quad i = 0, \dots, n-1$$

such that the discretized conditional mean and variance are given by

$$M_T = X_0 - \frac{1}{2}\sum_{i=1}^{n} \frac{Y_{t_i}^2 + Y_{t_{i-1}}^2}{2}\Delta t + \rho \sum_{i=1}^{n} Y_{t_i}\Delta W_{1,t_i} + \frac{1}{2}\rho\sigma \sum_{i=1}^{n}(\Delta W_{1,t_i}^2 - \Delta t)$$

and

$$C_T = (1-\rho^2)\sum_{i=1}^{n} \frac{Y_{t_i}^2 + Y_{t_{i-1}}^2}{2}\Delta t.$$

The time to maturity is $T = 1/12$ and we use a single step, $\Delta t = T$. We therefore use the optimal quantizers of the univariate normal distribution to approximate the Brownian increment $\Delta W_{1\Delta t}$. We also approximate option prices using an extended mixture obtained by including one additional component whose purpose is to adjust the moments of the auxiliary density as suggested in Remark 2.3.3. This additional component has a fixed weight equal to 5% while the other component weights are scaled down by 95%, its mean parameter is set equal to zero, and its variance parameter is computed such that the N^*-th moment of the auxiliary density and of the log price density are equal. A different number of components K for each auxiliary density and the first moment of the log price is always matched such that $\ell_1 = 0$. The reference parameters are $N^* = 20$, $r = \delta = X_0 = k = 0$, $\kappa = 0.5$, $\theta = Y_0 = 0.2$, $\sigma = 0.5$, and $\rho = -0.5$.

Table 2.3 reports the implied volatility errors for a call option with log strike $k = 0$ for different K and using the Gaussian mixture and the extended Gaussian mixture as auxiliary density, respectively denoted GM and GM$^+$. We can see that the implied volatility errors rapidly become small as the truncation order N increases for all K with the extended mixture and continue to converge toward π_f well after the matched moment N^*. On the other hand, the implied volatility errors do not appear to converge when K is small with the standard Gaussian mixture. Note that good option price approximations can be achieved by choosing a large number of components K even without moment matching.

	$K=3$		$K=10$		$K=50$	
N	GM	GM$^+$	GM	GM$^+$	GM	GM$^+$
0–1	1.26	0.64	0.01	1.85	0.16	1.95
2	0.75	0.12	0.20	0.24	0.10	0.24
3	0.26	0.02	0.15	0.19	0.16	0.18
4	0.43	0.07	0.09	0.05	0.15	0.05
5	0.08	0.03	0.06	0.03	0.08	0.03
6	0.68	0.06	0.04	0.01	0.03	0.02
7	0.46	0.03	0.01	0.00	0.04	0.01
8	1.27	0.04	0.01	0.01	0.01	0.02
9	1.56	0.03	0.03	0.01	0.00	0.01
10	2.94	0.03	0.04	0.02	0.01	0.03
11	6.06	0.02	0.06	0.01	0.01	0.02
12	8.02	0.02	0.04	0.02	0.01	0.03
13	25.63	0.01	0.05	0.02	0.01	0.02
14	23.63	0.01	0.18	0.02	0.02	0.03
15	–	0.01	0.10	0.02	0.01	0.03
16	–	0.01	0.43	0.02	0.02	0.03
17	–	0.01	0.03	0.02	0.01	0.03
18	–	0.01	1.30	0.02	0.02	0.02
19	–	0.01	0.63	0.02	0.02	0.02
20	–	0.00	3.82	0.02	0.02	0.02
30	–	0.00	–	0.00	0.00	0.00
40	–	0.00	–	0.00	4.65	0.00
50	–	4.34	–	4.60	–	9.96

Table 2.3 – Implied volatility errors for the Stein-Stein model.

The reported values are absolute percentage errors with respect to the implied volatility computed with Fourier technique. The GM$^+$ column refers to option price approximations obtained with a Gaussian mixture auxiliary density whose 20-th moment is matching $\mathbb{E}[X_T^{20}]$. The ”–” symbol indicates that the implied volatility was not retrievable either because the option price approximation was negative or because the implied volatility was larger than 99%.

2.5 Basis Construction with Moments

In this section we present moment-based constructions, alternative to Proposition 2.2.3, for the orthonormal basis (ONB) $H_n(x)$ of the space L^2_w which can also be used when the auxiliary density w is d-valued. Let $\pi : \mathscr{E} \to \{1, \dots, M\}$ be an enumeration of the set of exponents

$$\mathscr{E} = \left\{ n \in \mathbb{N}^d : |n| \le N \right\}$$

for some positive integer N, with $\pi(0) = 1$, and such that $\pi(n) \le \pi(m)$ if $|n| \le |m|$. We denote $\pi_i = \pi^{-1}(i) \in \mathbb{N}^d$ where $\pi^{-1} : \{1, \dots, M\} \to \mathscr{E}$ is the inverse function of π.

A standard approach to construct the ONB is to apply the Gram-Schmidt algorithm outlined below. First ones constructs the orthogonal basis

$$\begin{aligned} u_0(x) &= 1 \\ u_i(x) &= x^{\pi_i} - \sum_{j=0}^{i-1} \frac{\langle x^{\pi_i}, u_j \rangle_w}{\langle u_j, u_j \rangle_w} u_j(x), \quad i \ge 1 \end{aligned}$$

and the ONB is obtained by normalization,

$$H_i(x) = \frac{u_i(x)}{\| u_i \|_w}, \quad i \ge 0.$$

Another interesting approach to construct an ONB of the space L^2_w is as follows. Let $\mathbf{M}$ denote the $(M \times M)$ Gram matrix defined by

$$\mathbf{M}_{i+1, j+1} = \langle x^{\pi_i}, x^{\pi_j} \rangle_w \tag{2.32}$$

which is thus symmetric and positive definite. Let $\mathbf{M} = \mathbf{L}\mathbf{L}^\top$ be the unique Cholesky decomposition of $\mathbf{M}$ where $\mathbf{L}$ is a lower triangular matrix, and defined the lower triangular $\mathbf{S} = \mathbf{L}^{-1}$.

Theorem 2.5.1 ((Mysovskikh 1968))**.** *The polynomials*

$$H_i(x) = \sum_{j=0}^{i} \mathbf{S}_{i+1, j+1} \, x^{\pi_j}$$

form an ONB of L^2_w.

Remark 2.5.2. *The orthonormal basis resulting from the classical Gram-Schmidt described above may not appear orthogonal numerically because of rounding errors, the procedure is said to be numerically unstable. To alleviate this issue the modified Gram-Schmidt implementation is often preferred in practice, the polynomial* $u_i(x)$ *is now computed in multiple steps*

$$u_i^{(j+1)}(x) = u_i^{(j)}(x) - \frac{\langle u_i^{(j)}(x), u_j \rangle_w}{\langle u_j, u_j \rangle_w} u_j(x), \quad j = 0, \dots, i-1$$

with $u_i^{(0)}(x) = x^{\pi_i}$ *and such that* $u_i(x) = u_i^{(i)}(x)$*. Although the two algorithms are equivalent in exact arithmetic, significant difference can be observed in finite-precision arithmetic.*

Remark 2.5.3. *The Gram matrix in Equation* (2.32) *may numerically be singular because of rounding errors either in the computation in its eigenvalues or its moments. One approach to avoid this problem is to consider an approximately orthonormal basis in place of the monomial basis. By doing so, the Gram matrix would be already almost diagonal and thus more likely to be invertible. This may be achieved, for example, by implementing an algorithm that computes the ONB for the first* j *elements by using the ONB of the first* $j-1$ *elements enlarged with the monomial* x^{π_j}*.*

2.6 Proofs

This Section contains the proofs of all theorems and propositions in the main text.

Proof of Proposition 2.2.3

This proof is based on the results in (Fischer and Golub 1992). We first aim to derive a series of orthogonal monic polynomial basis $\{h_n\}_{n\geq 0}$, that is whose leading order coefficient is equal to one, and then normalize it to obtain the desired basis $\{H_n\}_{n\geq 0}$. The recurrence relation for the orthogonal monic basis is given by

$$x h_n(x) = h_{n+1}(x) + \alpha_n h_n(x) + \gamma_n h_{n-1}(x) \tag{2.33}$$

for all $n \geq 0$ with $h_{-1} = 0$ and $h_0 = 1$, and where the coefficients α_n, γ_n are given by

$$\alpha_n = \frac{\langle h_n^*, h_n\rangle_w}{\langle h_n, h_n\rangle_w} \quad \text{and} \quad \gamma_n = \frac{\langle h_n, h_n\rangle_w}{\langle h_{n-1}, h_{n-1}\rangle_w} \tag{2.34}$$

for $n \geq 0$, and where $h_n^*(x) = x h_n(x)$. The orthonormal polynomial basis is then obtained by normalizing the orthogonal monic basis, that is

$$H_n(x) = \frac{h_n(x)}{\sqrt{\langle h_n, h_n\rangle_w}},$$

which in view of Equation (2.33) is equivalent to define the recurrence coefficients as follows

$$a_n = \alpha_n \quad \text{and} \quad b_n = \sqrt{\gamma_n}.$$

The inner products in Equation (2.34) are left to be computed. We show that one can actually compute effectively and accurately the integral $\langle p, 1\rangle_w$ for any polynomial of order less than

$2N$. First, we recall the Gauss quadrature rule associated with the density v_k

$$\langle p, 1\rangle_{v_k} = \int_{\mathbb{R}} p(x)\, v_k(x)\, dx = \sum_{i=0}^{N} (v_{i1}^k)^2\, p(\lambda_i^k)$$

where v_i^k is the eigenvector corresponding to the eigenvalue λ_i^k of the Jacobi matrix J_N^k. Those values however do not need to be computed explicitly. Observe that, the matrix J_N^k being Hermitian, there exists a unitary matrix U_N^k whose columns are the normalized eigenvectors of J_N^k and such that

$$\Sigma_N^k := \operatorname{diag}(\lambda_0^k, \ldots, \lambda_N^k) = (U_N^k)^\top J_N^k\, U_N^k.$$

Combining the above results we obtain

$$\begin{aligned}\langle p, 1\rangle_w &= \sum_{k=1}^{K} c_k \sum_{i=0}^{N} (v_{i1}^k)^2\, p(\lambda_i^k) = \sum_{k=1}^{K} c_k\, \mathrm{e}_1^\top\, U_N^k\, p(\Sigma_N^k)(U_N^k)^\top\, \mathrm{e}_1 \\ &= \sum_{k=1}^{K} c_k\, \mathrm{e}_1^\top\, p(J_N^k)\, \mathrm{e}_1.\end{aligned}$$

Define the vector z_n^k as follows

$$\begin{aligned}z_{n+1}^k &= h_{n+1}(J_N^k)\mathrm{e}_1 = (J_N^k - \alpha_n) h_n(J_N^k)\mathrm{e}_1 - \gamma_n h_{n-1}(J_N^k)\mathrm{e}_1 \\ &= (J_N^k - \alpha_n) z_n^k \mathrm{e}_1 - \gamma_n z_{n-1}^k \mathrm{e}_1\end{aligned}$$

where the second equality follows from Equation (2.33). The inner products then rewrites

$$\langle h_n, h_n\rangle_w = \sum_{k=1}^{K} c_k\, \mathrm{e}_1^\top\, h_n(J_N^k)^\top\, h_n(J_N^k)^\top\, \mathrm{e}_1 = \sum_{k=1}^{K} c_k\, (z_n^k)^\top\, z_k$$

and similarly

$$\langle h_n^*, h_n\rangle_w = \sum_{k=1}^{K} c_k\, \mathrm{e}_1^\top\, h_n(J_N^k)^\top\, (J_N^k)^\top\, h_n(J_N^k)^\top\, \mathrm{e}_1 = \sum_{k=1}^{K} c_k\, (z_n^k)^\top\, J_N^k\, z_k.$$

Proof of Proposition 2.2.4

The proof follows several elementary steps

$$
\begin{aligned}
f_N &= \int_{\mathbb{R}} f(x) H_N(x) w(x) dx = \int_{\mathbb{R}} f(x) H_N(x) \sum_{k=1}^{k} c_k \, v_k(x) dx \\
&= \sum_{k=1}^{K} c_k \int_{\mathbb{R}} f(x) H_N(x) v_k(x) dx = \sum_{k=1}^{K} \sum_{n=0}^{N} c_k \int_{\mathbb{R}} q_{N,n}^k f(x) H_n^k(x) v_k(x) dx \\
&= \sum_{k=1}^{K} \sum_{n=0}^{N} c_k \, q_{N,n}^k \langle f, H_n^k \rangle_{v_k}
\end{aligned}
$$

which proves (2.8) and where the third line results from (2.9) which gives the representation of the polynomial $H_N(x)$ in the polynomials basis $H_n(x)$.

Proof of Theorem 2.2.6

We want to compute

$$
\int_{\mathbb{R}} f(x) L_n(x) v_k(x) dx = \frac{\sqrt{n!}}{\sqrt{\Gamma(\alpha+n)\Gamma(\alpha)}} \int_{\mu}^{\infty} (\mathrm{e}^{\xi + \frac{x}{\beta}} - \mathrm{e}^{k}) \mathscr{L}_n^{\alpha-1}(x) \, x^{\alpha-1} \mathrm{e}^{-x} dx
$$

by a change of variable $y = \beta(x-\xi)$, with $\mu = \max(0, \beta(k-\xi))$ and $v_k(x)$ as in (2.11). We first show that

$$
I_n^{\alpha-1}(\mu;\nu) = \int_{\mu}^{\infty} \mathrm{e}^{\nu x} \mathscr{L}_n^{\alpha-1}(x) \, x^{\alpha-1} \mathrm{e}^{-x} dx
$$

satisfies the recursive system (2.12). This directly follows from the recursive relations

$$
\mathscr{L}_n^{\alpha-1}(x) = \left(2 + \frac{\alpha-2}{n}\right) \mathscr{L}_{n-1}^{\alpha-1}(x) - \frac{1}{n} x \mathscr{L}_{n-1}^{\alpha-1}(x) - \left(1 + \frac{\alpha-2}{n}\right) \mathscr{L}_{n-2}^{\alpha-1}(x)
$$

and the three-point rule

$$
\mathscr{L}_n^{\alpha-1}(x) = \mathscr{L}_n^{\alpha}(x) - \mathscr{L}_{n-1}^{\alpha}(x)
$$

such that we obtain

$$
I_n^{\alpha-1}(\mu;\nu) = \left(2 + \frac{\alpha-2}{n}\right) I_{n-1}^{\alpha-1}(\mu;\nu) - \left(1 + \frac{\alpha-2}{n}\right) I_{n-2}^{\alpha-1}(\mu;\nu) - \frac{1}{n}\left(I_{n-1}^{\alpha}(\mu;\nu) - I_{n-2}^{\alpha}(\mu;\nu)\right).
$$

We conclude by computing

$$
\begin{aligned}
I_0^{\alpha-1}(\mu;\nu) &= \int_{\mu}^{\infty} x^{\alpha-1} \mathrm{e}^{-(1-\nu)x} dx = \frac{\Gamma(\alpha)}{(1-\nu)^{\alpha}} \int_{\mu}^{\infty} \frac{(1-\nu)^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \mathrm{e}^{-(1-\nu)x} dx \\
&= \frac{\Gamma(\alpha)}{(1-\nu)^{\alpha}} \left(1 - \int_{0}^{\mu} \frac{(1-\nu)^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \mathrm{e}^{-(1-\nu)x}) dx\right) = (1-\nu)^{-\alpha} \left(\Gamma(\alpha) - \Gamma(\alpha, \mu(1-\nu))\right)
\end{aligned}
$$

and since $\mathscr{L}_1^{\alpha-1}(x) = (\alpha + x)$ we get that

$$I_1^{\alpha-1}(\mu;\nu) = \int_\mu^\infty (\alpha + x)\mathrm{e}^{-(1-\nu)x} x^{\alpha-1} dx = \alpha I_0^{\alpha-1}(\mu;\nu) + I_0^{\alpha}(\mu;\nu)$$

Proof of Equations (2.14)–(2.15)

The dynamics of the instantaneous variance $V_t^2 = d\langle X\rangle_t / dt = \Sigma_1(Y_t)^2 + \Sigma_2(Y_t)^2$ is of the form

$$dV_t^2 = (\cdots)dt + 2\left(\Sigma_1'(Y_t)\Sigma_1(Y_t) + \Sigma_2'(Y_t)\Sigma_2(Y_t)\right)\sigma(Y_t)\, dW_{1t}.$$

The quadratic covariation between X_t and V_t is therefore given by

$$d\langle X, V^2\rangle_t = 2\left(\Sigma_1'(Y_t)\Sigma_1(Y_t) + \Sigma_2'(Y_t)\Sigma_2(Y_t)\right)\sigma(Y_t)\,\Sigma_1(Y_t)dt$$

and the quadratic variation of V_t by

$$d\langle V^2\rangle_t = 4\left(\Sigma_1'(Y_t)\Sigma_1(Y_t) + \Sigma_2'(Y_t)\Sigma_2(Y_t)\right)^2 \sigma(Y_t)^2 dt.$$

Equation (2.14) directly follows by observing that

$$\frac{2\left(\Sigma_1'(Y_t)\Sigma_1(Y_t) + \Sigma_2'(Y_t)\Sigma_2(Y_t)\right)\sigma(Y_t)}{\sqrt{4\left(\Sigma_1'(Y_t)\Sigma_1(Y_t) + \Sigma_2'(Y_t)\Sigma_2(Y_t)\right)^2 \sigma(Y_t)^2}} = \operatorname{sign}\left[\Sigma_1'(Y_t)\Sigma_1(Y_t) + \Sigma_2'(Y_t)\Sigma_2(Y_t)\sigma(Y_t)\right],$$

and Equation (2.15) follows from the above and

$$d\langle\sqrt{V^2}\rangle_t = \frac{d\langle V^2\rangle_t}{4V_t^2}.$$

Proof of Proposition 2.3.1

Conditional on the trajectory of W_{1t} on $[0,T]$, the trajectory of Y_t is observable and W_{2t} is the only source of randomness in the dynamics of X_T which is thus equivalent to the dynamics of a Gaussian process with time varying parameters. Hence, its conditional distribution is given by a normal distribution with mean M_T and variance C_T as in (2.16) and (2.17). Taking expectation gives (2.18)

Proof of Corollary 2.3.2

The first part of the proof follows from similar arguments as in Theorem 1.3.1. Note that it is sufficient to consider only the component of w with the largest variance parameter.

It is clear that C_T is bounded when Y_t is bounded since $\Sigma_2(y)^2$ is a polynomial. The random variable M_T is equivalently given by the expression

$$M_T = X_0 + (r-\delta)T - \frac{1}{2}\int_0^T \left(\Sigma_1(Y_t)^2 + \Sigma_2(Y_t)^2\right) dt + \int_0^T \frac{\Sigma_1(Y_t)}{\sigma(Y_t)}(dY_t - \kappa(\theta - Y_t)dt).$$

Applying Ito's lemma we get

$$\int_0^T Y_t^n dY_t = \frac{1}{n+1}(Y_T^{n+1} - Y_0^{n+1}) - \frac{1}{2}\int_0^T nY_t^{n-1}\frac{\Sigma_1(Y_t)}{\sigma(Y_t)} dt$$

which is bounded and so is M_T given that that $\Sigma_1(y)/\sigma(y)$ and $\Sigma_1(Y_t)^2 + \Sigma_2(Y_t)^2$ are some polynomials, and that Y_t is itself bounded.

2.7 Conclusion

In this chapter, we showed that option price series representations can also be derived in a tractable way when the auxiliary density is a mixture distribution. We presented methodologies to specify a Gaussian mixture for the auxiliary density for a class of stochastic volatility models, and indicate that accurate option price approximations are possible even when the log price density does not belong to the corresponding weighted space. We then provided numerical examples that illustrate the performance of this approach.

3 Linear Credit Risk Models

We introduce a novel class of credit risk models in which the drift of the survival process of a firm is a linear function of the factors. The prices of defaultable bonds and credit default swaps (CDS) are linear-rational in the factors. The price of a CDS option can be uniformly approximated by polynomials in the factors. Multi-name models with simultaneous defaults, positively and negatively correlated default intensities, and stochastic interest rates can be constructed. An empirical study illustrates the versatility of these models by fitting CDS spread time series. A numerical analysis validates the efficiency of the option price approximation method.

3.1 Introduction

Credit risk is inherent to virtually all financial securities. Breach of contracts caused by the nonpayment of cash flows, as well as variations in asset values caused by changing default risk, are omnipresent in financial markets. The underestimation of credit risk before the financial crisis incited regulators around the globe to force financial institutions to better manage and report credit risk. The complexity of credit risky portfolios and the securities therein renders the valuation of credit risk a challenging task that calls for suitable models.

In this chapter we introduce a novel class of flexible and tractable reduced form models for the term structure of credit risk, the linear credit risk models. We directly specify the survival process of a firm, that is, its conditional survival probability given the economic background information. Specifically, we assume a multivariate factor process with a linear drift and let the drift of the survival process be linear in the factors. Prices of defaultable bonds and credit default swaps (CDS) are given in closed-form by linear-rational functions in the factors. By linearity, the same result holds for the prices of CDSs on indices. The implied default intensity is an explicit linear-rational function of the factors. In contrast, the price of a CDS in an affine default intensity model is a sum of exponential-affine functions in the factors process and whose coefficients are given by the solutions of nonlinear ordinary differential equations which may not be explicit. In addition, the linear credit risk models offer new tractable features such as a multi-name model with negatively correlated default intensity.

Within the linear framework we define the linear hypercube (LHC) model which is a single-firm model. The factor process is diffusive with quadratic diffusion function so that it takes values in a hypercube whose edges' length is given by the survival process. The quadratic diffusion function is concave and bi-monotonic. This feature allows factors to virtually jump between low and high values. This facilitates the persistence and likelihood of term structure shifts. The factors' volatility parameters do not enter the bond and CDS pricing formulas, yet they impact the volatility of CDS spreads and thus affect CDS option prices. This may facilitate the joint calibration of credit spread and option price time series. We discuss in detail the one-factor LHC model and compare it with the one-factor affine default intensity model. We provide an identifiable canonical representation and the market price of risk specifications that preserve the linear drift of the factors.

We present a price approximation methodology for European style options on credit risky underlyings that exploits the compactness of the state space and the closed-form of the conditional moments of the factor process. First, by the Stone–Weierstrass theorem, any continuous payoff function on the compact state space can be approximated by a polynomial to any given level of accuracy. Second, the conditional expectation of any polynomial in the factors is a polynomial in the prevailing factor values. In consequence, the price of a CDS option can be uniformly approximated by polynomials in the factors. This method also applies to the computation of credit valuation adjustments.

We develop multi-name models by letting the survival processes be linear and polynomial combinations of independent LHC models. Bonds and CDSs are still linear but with respect to an extended factor representation. These direct extensions can easily accommodate the inclusion of new factors and new firms. Stochastic short rate models with a similar specification as the survival processes can be introduced while preserving tractability. Simultaneous defaults can be produced either by introducing a common jump process in the survival processes, or by a random time change

We perform an empirical and numerical analysis of the LHC model. Assuming a parsimonious cascading drift structure, we fit two- and three-factor LHC models to the ten-year long time series of weekly CDS spreads on an investment grade and a high yield firm. The three-factor model is able to capture the complex term structure dynamics remarkably well and performs significantly better than the two-factor model. We illustrate the numerical efficiency of the option pricing method by approximating the prices of CDS options with different moneyness. Polynomials of relatively low orders are sufficient to obtain accurate approximations for in-the-money options. Out-of-the money options typically require a higher order. We show that the CDS option on a homogeneous portfolio can be also approximated efficiently.

We now review some of the related literature. Our approach follows a standard doubly stochastic construction of default times as described in (Elliott, Jeanblanc, and Yor 2000) or (Bielecki and Rutkowski 2002). The early contributions by (Lando 1998) and (Duffie and Singleton

1999) already make use of affine factor processes. In contrast, the factor process in the LHC model is a strictly non-affine polynomial diffusion, whose general properties are studied in (Filipović and Larsson 2016). The stochastic volatility models developed in (Hull and White 1987) and Chapter 1 are two other examples of non-affine polynomial models. Factors in the LHC models have a compact support and can exhibit jump-like dynamics similar to the multivariate Jacobi process introduced by (Gourieroux and Jasiak 2006). Our approach has some similarities with the linearity generating process by (Gabaix 2009) and the linear-rational models by (Filipović, Larsson, and Trolle 2017). These models exploit the tractability of the factor processes with linear drift and target non default risky asset pricing. To our knowledge, we are the first to model directly the survival process of a firm with linear drift characteristics.

Options on CDS contracts are complex derivatives and intricate to price. The pricing and hedging of CDS options in a generic hazard process framework is discussed in (Bielecki, Jeanblanc, and Rutkowski 2006) and (Bielecki, Jeanblanc, Rutkowski, et al. 2008), and specialised to the square-root diffusion factor process in (Bielecki, Jeanblanc, and Rutkowski 2011). More recently (Brigo and El-Bachir 2010) developed a semi-analytical expression for CDS option prices in the context of a shifted square-root jump-diffusion default intensity model that was introduced in (Brigo and Alfonsi 2005). Another strand of the literature has focused on developing market models in the spirit of LIBOR market models. We refer the interested reader to (Schönbucher 2000), (Hull and White 2003), (Schönbucher 2004), (Jamshidian 2004), and (Brigo and Morini 2005). Black-Scholes like formulas are then obtained for the prices of CDS options by assuming, for example, that the underlying CDS spread follows a geometric Brownian motion under the survival measure. Although offering more tractability, this approach makes it difficult, if not impossible, to consistently valuate multiple instruments exposed to the same credit risk. (Di Graziano and Rogers 2009) introduced a framework where they obtained explicit expressions which are similar to ours for CDS prices, but under the assumption that the firm default intensity is driven by a continuous-time finite-state irreducible Markov chain. Another tractable approach to price multi-name credit derivatives is to model the dependence between defaults with a copula function, as for examples in (Li 2000) or (Laurent and Gregory 2005). However these models are by construction static and require repeated calibration, and they become intractable when combined with stochastic survival processes as in (Schönbucher and Schubert 2001). The idea of approximating option prices by power series can be traced back to (Jarrow and Rudd 1982). However, most of the previous literature has focussed on approximating the transition density function of the underlying process, see for example (Corrado and Su 1996) and (Filipović, Mayerhofer, and Schneider 2013). In contrast, we approximate directly the payoff function by a polynomial.

The remainder of the chapter is structured as follows. Section 3.2 presents the linear credit risk framework. Section 3.3 describes the LHC model and Section 3.4 some extensions. The numerical and empirical analysis is in Section 3.5. Section 3.6 describes the two-dimensional Chebyshev interpolation. Section 3.7 provides some additional results on market price of risk specifications that preserve the linear drift of the factors. The proofs are collected in Section 3.8. Section 3.9 concludes.

3.2 The Linear Framework

We introduce the linear credit risk model framework and derive closed-form expressions for defaultable bond prices and credit default swap spreads. We also discuss the pricing of credit default swap options, and of credit valuation adjustments.

3.2.1 Survival Process Specification

We fix a stochastic basis $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{Q})$, where $\mathscr{F}_t$ represents the economic background information and $\mathbb{Q}$ is the risk-neutral pricing measure. We consider N firms and let S_t^i be the survival process of firm i. This is a right-continuous $\mathscr{F}_t$-adapted and non-increasing positive process with $S_0^i = 1$. Let $U^1, \dots, U^N$ be standard uniform random variables which are independent from $\mathscr{F}_\infty$ and from each other. For each firm, we define the random time

$$\tau_i := \inf\{t \geq 0 \mid S_t^i \leq U_i\},$$

which is infinity if the set is empty. Let $\mathscr{H}_t^i$ be the filtration generated by the indicator process $H_t^i = \mathbb{1}_{\{\tau_i > t\}}$, which is one as long as firm i has not defaulted by time t, and zero afterwards. The default time τ_i is a stopping time in the enlarged filtration $\mathscr{F}_t \vee \mathscr{H}_t^i$. It is $\mathscr{F}_t$-doubly stochastic in the sense that

$$\mathbb{P}\left[\tau_i > t \mid \mathscr{F}_\infty\right] = \mathbb{P}\left[S_t^i > U_i \mid \mathscr{F}_\infty\right] = S_t^i.$$

The filtration $\mathscr{G}_t = \mathscr{F}_t \vee \mathscr{H}_t^1 \vee \cdots \vee \mathscr{H}_t^N$ contains all the information about the occurrence of firm defaults, as well as the economic background information. Henceforward we omit the index i of the firm and refer to any of the N firms as long as there is no ambiguity.

In a linear credit risk model the survival process of a firm is defined by

$$S_t = a^\top Y_t \tag{3.1}$$

for some firm specific parameter $a \in \mathbb{R}_+^n$, and some common factor process (Y_t, X_t) taking values in $\mathbb{R}_+^n \times \mathbb{R}^m$ with linear drift of the form

$$dY_t = (c\,Y_t + \gamma\,X_t)dt + dM_t^Y \tag{3.2}$$

$$dX_t = (b\,Y_t + \beta\,X_t)dt + dM_t^X \tag{3.3}$$

for some $c \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^{m\times n}$, $\gamma \in \mathbb{R}^{n\times m}$, $\beta \in \mathbb{R}^{m\times m}$, m-dimensional $\mathscr{F}_t$-martingale M_t^X, and n-dimensional $\mathscr{F}_t$-martingale M_t^Y. The process S_t being positive and non-increasing, we necessarily have that its martingale component $M_t^S = a^\top M_t^Y$ is of finite variation and thus purely discontinuous, see (Jacod and Shiryaev 2013, Lemma I.4.14), and that $-S_{t-} < \Delta M_t^S \leq 0$ because $\Delta S_t = \Delta M_t^S$. This observation motivates the decomposition of the factor process into a component X_t and a component Y_t with finite variation. Although we do not specify further the dynamics of the factor process at the moment, it is important to emphasize that additional conditions should be satisfied to ensure that S_t is a valid survival process.

Remark 3.2.1. *In practice we will consider a componentwise non-increasing process Y_t with $Y_0 = \mathbf{1}$. Survival processes can then easily be constructed by choosing any vector $a \in \mathbb{R}^n_+$ such that $a^\top \mathbf{1} = 1$.*

The linear drift of the process (Y_t, X_t) implies that the $\mathscr{F}_t$-conditional expectation of (Y_t, X_t) is linear of the form

$$\mathbb{E}\left[\begin{pmatrix} Y_T \\ X_T \end{pmatrix} \mid \mathscr{F}_t\right] = \mathrm{e}^{A(T-t)} \begin{pmatrix} Y_t \\ X_t \end{pmatrix}, \quad t \leq T, \tag{3.4}$$

where the $(m+n) \times (m+n)$-matrix A is defined by

$$A = \begin{pmatrix} c & \gamma \\ b & \beta \end{pmatrix}. \tag{3.5}$$

Remark 3.2.2. *If S_t is absolutely continuous, so that $a^\top dM^Y_t = 0$ for all $t \geq 0$, the corresponding default intensity λ_t that derives from the relation $S_t = \mathrm{e}^{-\int_0^t \lambda_s ds}$ is linear-rational in (Y_t, X_t) of the form*

$$\lambda_t = -\frac{a^\top (c Y_t + \gamma X_t)}{S_t}.$$

In this framework, the default times are correlated because the survival processes are driven by common factors. Simultaneous defaults are possible and may be caused by the martingale components of Y_t that forces the survival processes to jump downward at the same time. Additionally, and to the contrary of affine default intensities models, the linear credit risk framework allows for negative correlation between default intensities as illustrated by the following stylized example.

Example 3.2.3. *Consider the factor process (Y_t, X_t) taking values in $\mathbb{R}^2_+ \times \mathbb{R}$ defined by*

$$\begin{aligned} dY_t &= \frac{\epsilon}{2}\left(\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} Y_t + \begin{pmatrix} -1 \\ 1 \end{pmatrix} X_t\right) dt \\ dX_t &= -\kappa X_t dt + \sigma \sqrt{(\mathrm{e}^{-\epsilon t} - X_t)(\mathrm{e}^{-\epsilon t} + X_t)}\, dW_t \end{aligned}$$

for some $\kappa > \epsilon > 0$, $\sigma > 0$, $X_0 \in [-1, 1]$, and $\mathscr{F}_t$-adapted univariate Brownian motion W_t. The process X_t takes values in the interval $[-\mathrm{e}^{-\epsilon t}, \mathrm{e}^{-\epsilon t}]$ at time t. Define $N = 2$ survival processes as follows

$$S^1_t = Y_{1t} \quad \textit{and} \quad S^2_t = Y_{2t}$$

so that the implied default intensities of the two firms rewrite

$$\lambda_t^1 = \frac{\epsilon}{2}\left(1 + \frac{X_t}{Y_{1t}}\right) \quad \textit{and} \quad \lambda_t^2 = \frac{\epsilon}{2}\left(1 - \frac{X_t}{Y_{2t}}\right)$$

which results in $d\langle \lambda^1, \lambda^2 \rangle_t \le 0$ *and* $\lambda_t^1, \lambda_t^2 \le \epsilon$ *as shown in Section 3.8.*

3.2.2 Defaultable Bonds

We consider securities with notional value equal to one and exposed to the credit risk of a reference firm. We assume a constant risk-free interest rate equal to r such that the time-t price of the risk-free zero-coupon bond price with maturity T and notional value one is given by $\mathrm{e}^{-r(T-t)}$. The following result gives a closed-form expression for the price of a defaultable bond with constant recovery rate at maturity.

Proposition 3.2.4. *The time-t price of a defaultable zero-coupon bond with maturity T and recovery $\delta \in [0,1]$ at maturity is*

$$B_{\mathrm{M}}(t,T) = \mathbb{E}\left[\mathrm{e}^{-r(T-t)}\left(\mathbb{1}_{\{\tau>T\}} + \delta \mathbb{1}_{\{\tau\le T\}}\right) \mid \mathcal{G}_t\right] = (1-\delta) B_{\mathrm{Z}}(t,T) + \mathbb{1}_{\{\tau>t\}} \delta\, \mathrm{e}^{-r(T-t)}$$

where $B_{\mathrm{Z}}(t,T) = \mathrm{e}^{-r(T-t)}\mathbb{E}\left[\mathbb{1}_{\{\tau>T\}} \mid \mathcal{G}_t\right]$ denotes the time-t price of a defaultable zero-coupon bond with maturity T and zero recovery. It is of the form

$$B_{\mathrm{Z}}(t,T) = \mathbb{1}_{\{\tau>t\}} \frac{1}{S_t} \psi_{\mathrm{Z}}(t,T)^\top \begin{pmatrix} Y_t \\ X_t \end{pmatrix} \tag{3.6}$$

where the vector $\psi_{\mathrm{Z}}(t,T) \in \mathbb{R}^{1+m}$ is given by

$$\psi_{\mathrm{Z}}(t,T)^\top = \mathrm{e}^{-r(T-t)} \begin{pmatrix} a^\top & 0 \end{pmatrix} \mathrm{e}^{A(T-t)}.$$

The next result shows that the price of a defaultable bond paying a constant recovery rate at default can also be retrieved in closed-form.

Proposition 3.2.5. *The time-t price of a defaultable zero-coupon bond with maturity T and recovery $\delta \in [0,1]$ at default is*

$$B_{\mathrm{D}}(t,T) = \mathbb{E}\left[\mathrm{e}^{-r(T-t)}\mathbb{1}_{\{\tau>T\}} + \delta \mathrm{e}^{-r(\tau-t)}\mathbb{1}_{\{t<\tau\le T\}} \mid \mathcal{G}_t\right] = B_{\mathrm{Z}}(t,T) + \delta\, C_{\mathrm{D}}(t,T),$$

where $C_{\mathrm{D}}(t,T) = \mathbb{E}\left[\mathrm{e}^{-r(\tau-t)}\mathbb{1}_{\{t<\tau\le T\}} \mid \mathcal{G}_t\right]$ denotes the time-t price of a contingent claim paying one at default if it occurs between dates t and T. It is of the form

$$C_{\mathrm{D}}(t,T) = \mathbb{1}_{\{\tau>t\}} \frac{1}{S_t} \psi_{\mathrm{D}}(t,T)^\top \begin{pmatrix} Y_t \\ X_t \end{pmatrix} \tag{3.7}$$

where the vector $\psi_{\mathrm{D}}(t,T) \in \mathbb{R}^{1+m}$ *is given by*

$$\psi_{\mathrm{D}}(t,T)^\top = -a^\top \begin{pmatrix} c & \gamma \end{pmatrix} \left(\int_t^T \mathrm{e}^{A_*(s-t)} ds \right) \tag{3.8}$$

where $A_* = A - r\,\mathrm{Id}$.

The price of a security whose only cash flow is proportional to the default time is given in the following corollary. It is of interest to compute the expected accrued interests at default for some contingent securities such as CDS.

Corollary 3.2.6. *The time-t price of a contingent claim paying τ at default if it occurs between date t and T is of the form*

$$C_{\mathrm{D}_*}(t,T) = \mathbb{E}\left[\tau \mathrm{e}^{-r(\tau-t)} \mathbb{1}_{\{\tau \le T\}}) \mid \mathcal{G}_t\right] = \mathbb{1}_{\{\tau > t\}} \frac{1}{S_t} \psi_{\mathrm{D}_*}(t,T)^\top \begin{pmatrix} Y_t \\ X_t \end{pmatrix} \tag{3.9}$$

where the vector $\psi_{\mathrm{D}_*}(t,T)^\top \in \mathbb{R}^{1+m}$ *is given by*

$$\psi_{\mathrm{D}_*}(t,T)^\top = -a^\top \begin{pmatrix} c & \gamma \end{pmatrix} \left(\int_t^T s\mathrm{e}^{A_*(s-t)} ds \right). \tag{3.10}$$

Note the presence of a term s in the integral expression on the right hand side of (3.10) which is absent in (3.8). The following Lemma shows that pricing formulas (3.7)–(3.9) can further simplify with an additional condition.

Lemma 3.2.7. *Assume that the matrix A_* is invertible then the vectors in the contingent cash-flow prices rewrite explicitly as follows*

$$\begin{aligned} \psi_{\mathrm{D}}(t,T)^\top &= -a^\top \begin{pmatrix} c & \gamma \end{pmatrix} A_*^{-1} \left(\mathrm{e}^{A_*(T-t)} - \mathrm{Id}\right) \\ \psi_{\mathrm{D}_*}(t,T)^\top &= -a^\top \begin{pmatrix} c & \gamma \end{pmatrix} \left((T-t) A_*^{-1} \mathrm{e}^{A_*(T-t)} + A_*^{-1}(\mathrm{Id}\, t - A_*^{-1})(\mathrm{e}^{A_*(T-t)} - \mathrm{Id}) \right) \end{aligned}$$

where Id *is the* $(m+n)$*-dimensional identity matrix.*

This is a remarkable result since the prices of contingent cash flows become analytical expressions composed of basic matrix operations and are thus easily computed. Closed-form formulas for defaultables securities renders the linear framework appealing for large scale applications, for example with a large number of firms and contracts, in comparison to standard affine default intensity models that will require further numerical methods in general. For illustration, assume that the survival process S_t is absolutely continuous so that it admits the default intensity λ_t as in Remark 3.2.2. Then $C_{\mathrm{D}}(t,T)$ can be rewritten as

$$C_{\mathrm{D}}(t,T) = \mathbb{1}_{\{\tau > t\}} \int_t^T \mathrm{e}^{-r(u-t)} \mathbb{E}\left[\lambda_u \mathrm{e}^{-\int_t^u \lambda_s ds} \mid \mathcal{F}_t \right] du.$$

With affine default intensity models the expectation to be integrated requires solving Riccati equations, which have an explicit solution only when the default intensity is driven by a sum of independent univariate CIR processes. Numerical methods such as finite difference are usually employed to compute the expectation with time-u cash flow for $u \in [t,T]$. The integral can then only be approximated by means of another numerical method such as quadrature, that necessitates solving the corresponding ODEs at many different points u. For more details on affine default intensity models we refer to (Duffie and Singleton 2012), (Filipović 2009), and (Lando 2009).

3.2.3 Credit Default Swaps

Single-Name CDS

A CDS is an insurance contract that pays at default the realized loss on a reference bond – the protection leg – in exchange of periodic payments that will stop after default – the premium leg. We consider the following discrete tenor structure $t \le T_0 < T_1 < \cdots < T_M$ and a contract offering default protection from date T_0 to date T_M. When $t < T_0$ the contract is usually called a knock out forward CDS and generates cash flows only if the firm has not defaulted by time T_0. We consider a CDS contract with notional value equal to one. The time-t value of the premium leg with spread k is given by $k\,V_{\mathrm{prem}}(t,T_0,T_M)$ where

$$V_{\mathrm{prem}}(t,T_0,T_M) = V_{\mathrm{coup}}(t,T_0,T_M) + V_{\mathrm{ai}}(t,T_0,T_M)$$

is the sum of the value of coupon payments before default

$$V_{\mathrm{coup}}(t,T_0,T_M) = \mathbb{E}\left[\sum_{j=1}^{M} \mathrm{e}^{-r(T_j-t)}(T_j - T_{j-1})\mathbb{1}_{\{T_j<\tau\}} \mid \mathcal{G}_t\right]$$

and the value of the accrued coupon payment at the time of default

$$V_{\mathrm{ai}}(t,T_0,T_M) = \mathbb{E}\left[\sum_{j=1}^{M} \mathrm{e}^{-r(\tau-t)}(\tau - T_{j-1})\mathbb{1}_{\{T_{j-1}<\tau\le T_j\}} \mid \mathcal{G}_t\right].$$

The time-t value of the protection leg is

$$V_{\mathrm{prot}}(t,T_0,T_M) = (1-\delta)\mathbb{E}\left[\mathrm{e}^{-r(\tau-t)}\mathbb{1}_{\{T_0<\tau\le T_M\}} \mid \mathcal{G}_t\right],$$

where $\delta \in [0,1]$ denotes the constant recovery rate at default. The (forward) CDS spread $\mathrm{CDS}(t,T_0,T_M)$ is the spread k that makes the premium leg and the protection leg equal in value at time t. That is,

$$\mathrm{CDS}(t,T_0,T_M) = \frac{V_{\mathrm{prot}}(t,T_0,T_M)}{V_{\mathrm{prem}}(t,T_0,T_M)}.$$

Proposition 3.2.8. *The values of the protection and premium legs are given by*

$$V_{\text{prot}}(t,T_0,T_M)=\mathbb{1}_{\{\tau>t\}}\frac{1}{S_t}\psi_{\text{prot}}(t,T_0,T_M)^\top\begin{pmatrix}Y_t\\X_t\end{pmatrix}\tag{3.11}$$

$$V_{\text{prem}}(t,T_0,T_M)=\mathbb{1}_{\{\tau>t\}}\frac{1}{S_t}\psi_{\text{prem}}(t,T_0,T_M)^\top\begin{pmatrix}Y_t\\X_t\end{pmatrix}\tag{3.12}$$

where the vectors $\psi_{\text{prot}}(t,T_0,T_M),\psi_{\text{prem}}(t,T_0,T_M)\in\mathbb{R}^{1+m}$ *are given by*

$$\begin{aligned}\psi_{\text{prot}}(t,T_0,T_M)&=(1-\delta)\left(\psi_{\text{D}}(t,T_M)-\psi_{\text{D}}(t,T_0)\right),\\ \psi_{\text{prem}}(t,T_0,T_M)&=\sum_{j=1}^{M}(T_j-T_{j-1})\psi_{\text{Z}}(t,T_j)+\psi_{\text{D}_*}(t,T_M)-\psi_{\text{D}_*}(t,T_0)\\ &\quad+T_{M-1}\psi_{\text{D}}(t,T_M)-\sum_{j=1}^{M-1}(T_j-T_{j-1})\psi_{\text{D}}(t,T_j)-T_0\psi_{\text{D}}(t,T_0).\end{aligned}$$

The CDS spread is given by a readily available linear-rational expression,

$$\text{CDS}(t,T_0,T_M)=\mathbb{1}_{\{\tau>t\}}\frac{\psi_{\text{prot}}(t,T_0,T_M)^\top\begin{pmatrix}Y_t\\X_t\end{pmatrix}}{\psi_{\text{prem}}(t,T_0,T_M)^\top\begin{pmatrix}Y_t\\X_t\end{pmatrix}}.\tag{3.13}$$

This is a remarkably simple expression that allows us to see how the factors (Y_t,X_t) affect the CDS spread through the vectors $\psi_{\text{prot}}(t,T_0,T_M)$ and $\psi_{\text{prem}}(t,T_0,T_M)$. For comparison, in an affine default intensity model the two legs $V_{\text{prot}}(t,T_0,T_M)$ and $V_{\text{prem}}(t,T_0,T_M)$ are given as sums of exponential-affine terms that cannot be simplified further.

Multi-Name CDS

A credit default index swap (CDIS) is an insurance on a reference portfolio of N firms with equal weight that we assume to be $1/N$ so that the portfolio total notional is equal to one. The protection buyer pays a regular premium that is proportional to the current nominal value of the CDIS. Let $\delta\in[0,1]$ be the recovery rate determined at inception. Upon default of a firm the protection seller pays $1-\delta$ to the protection buyer and the notional value of the CDIS decreases by $1/N$. These steps repeat until maturity or until all the firms in the reference portfolio have defaulted, whichever comes first.

The CDIS spread simplifies to a double linear-rational expression,

$$\mathrm{CDIS}(t,T_0,T_M)=\frac{\sum_{i=1}^{N}\mathbb{1}_{\{\tau_i>t\}}\,(1/S_t^i)\,\psi_{\mathrm{prot}}^i(t,T_0,T_M)^\top\begin{pmatrix}Y_t\\X_t\end{pmatrix}}{\sum_{i=1}^{N}\mathbb{1}_{\{\tau_i>t\}}\,(1/S_t^i)\,\psi_{\mathrm{prem}}^i(t,T_0,T_M)^\top\begin{pmatrix}Y_t\\X_t\end{pmatrix}}$$

where $\psi_{\mathrm{prot}}^i(t,T_0,T_M)$ and $\psi_{\mathrm{prem}}^i(t,T_0,T_M)$ are defined as in Proposition 3.2.8 for the i-th firm.

Unspanned Factors

The characteristics of the martingales M_t^Y and M_t^X do not appear explicitly in the bond, CDS and CDIS pricing formulas. This leaves the freedom to specify exogenous factors that feed into M_t^Y and M_t^X. Such factors would be unspanned by the term structures of defaultable bonds and CDS and give rise to unspanned stochastic volatility, as described in (Filipović, Larsson, and Trolle 2017). They provide additional flexibility for fitting time series of bond prices and CDS spreads. Furthermore, these unspanned stochastic volatility factors affect the distribution of the survival and factor processes and therefore can be recovered from the prices of credit derivatives such as those discussed hereinafter.

3.2.4 CDS Option and CDIS Option

A CDS option with strike spread k is a European call option on the CDS contract exercisable only if the firm has not defaulted before the option maturity date T_0. Its payoff is

$$\mathbb{1}_{\{\tau>T_0\}}\left(V_{\mathrm{prot}}(T_0,T_0,T_M)-k\,V_{\mathrm{prem}}(T_0,T_0,T_M)\right)^+=\frac{\mathbb{1}_{\{\tau>T_0\}}}{S_{T_0}}\left(\psi_{\mathrm{cds}}(T_0,T_0,T_M,k)^\top\begin{pmatrix}Y_{T_0}\\X_{T_0}\end{pmatrix}\right)^+ \tag{3.14}$$

with

$$\psi_{\mathrm{cds}}(t,T_0,T_M,k)=\psi_{\mathrm{prot}}(t,T_0,T_M)-k\psi_{\mathrm{prem}}(t,T_0,T_M). \tag{3.15}$$

Denote $V_{\mathrm{CDSO}}(t,T_0,T_M,k)$ the price of the CDS option at time t,

$$\begin{aligned}V_{\mathrm{CDSO}}(t,T_0,T_M,k)&=\mathbb{E}\left[\mathrm{e}^{-r(T_0-t)}\frac{\mathbb{1}_{\{\tau>T_0\}}}{S_{T_0}}\left(\psi_{\mathrm{cds}}(T_0,T_0,T_M,k)^\top\begin{pmatrix}Y_{T_0}\\X_{T_0}\end{pmatrix}\right)^+\mid\mathcal{G}_t\right]\\&=\mathbb{1}_{\{\tau>t\}}\frac{\mathrm{e}^{-r(T_0-t)}}{S_t}\mathbb{E}\left[\left(\psi_{\mathrm{cds}}(T_0,T_0,T_M,k)^\top\begin{pmatrix}Y_{T_0}\\X_{T_0}\end{pmatrix}\right)^+\mid\mathcal{F}_t\right]\end{aligned} \tag{3.16}$$

where the second equality follows directly from Lemma 3.8.1.

The price of a CDS option is therefore equal to the expected positive part of a linear function of (Y_{T_0},X_{T_0}), adjusted for time value and realized credit risk. When the characteristic function of the process (Y_{T_0}, X_{T_0}) is available such expression can be computed efficiently using Fourier transform techniques. An alternative method is presented in Section 3.3.2 which is based on the polynomial approximation of the payoff function and the conditional moments of (Y_{T_0}, X_{T_0}).

A CDS index option gives the right at time T_0 to enter a CDS index contract with strike k and maturity T_M on the firms in the reference portfolio which have not defaulted and, simultaneously, to receive the losses realized before the exercise date T_0. Denote by $V_{\mathrm{CDISO}}(t, T_0, T_M, k)$ the price of the CDIS option at time $t \leq T_0$,

$$\begin{aligned} &V_{\mathrm{CDISO}}(t, T_0, T_M, k) = \\ &\quad \mathbb{E}\left[\frac{\mathrm{e}^{-r(T_0-t)}}{N}\left(\sum_{i=1}^{N}\frac{\mathbb{1}_{\{\tau_i>T_0\}}}{S^i_{T_0}}\psi^i_{\mathrm{cds}}(T_0, T_0, T_M, k)^\top\begin{pmatrix}Y_{T_0}\\X_{T_0}\end{pmatrix} + (1-\delta)\mathbb{1}_{\{\tau_i\leq T_0\}}\right)^+ \mid \mathcal{G}_t\right]. \end{aligned} \tag{3.17}$$

Proposition 3.2.9. *The price of a CDIS option is given by*

$$\begin{aligned} V_{\mathrm{CDISO}}(t, T_0, T_M, k) = \sum_{\alpha\in\mathcal{C}}\frac{\mathrm{e}^{-r(T_0-t)}}{N}\mathbb{E}\Bigg[&\left(\sum_{i=1}^{N}\frac{\alpha_i}{S^i_{T_0}}\psi^i_{\mathrm{cds}}(T_0, T_0, T_M, k)^\top\begin{pmatrix}Y_{T_0}\\X_{T_0}\end{pmatrix} + (1-\delta)(1-\alpha_i)\right)^+ \\ &\times\prod_{i=1}^{N}\left(\frac{(S^i_{T_0})^{\alpha_i}(S^i_t - S^i_{T_0})^{1-\alpha_i}}{S^i_t}\mathbb{1}_{\{\tau_i>t\}} + (\mathbb{1}_{\{\tau_i\leq t\}})^{1-\alpha_i}\right) \mid \mathcal{F}_t\Bigg] \end{aligned} \tag{3.18}$$

where $\mathcal{C} = \{0,1\}^N$ *is the set of all possible defaults combinations, with the convention* $0^0 = 0$, *and we write* $\alpha = (\alpha_1, \ldots, \alpha_N)$.

3.2.5 Credit Valuation Adjustment

The unilateral credit valuation adjustment (UCVA) of a position is the present value of losses resulting from its cancellation when a bilateral counterparty defaults.

Proposition 3.2.10. *The time-t price of the UCVA with maturity T and time-u net positive exposure $f(u, S_u, Y_u, X_u)$, for some continuous function $f(u, s, y, x)$, is*

$$\begin{aligned} \mathrm{UCVA}(t, T) &= \mathbb{E}\left[\mathrm{e}^{-r(\tau-t)}\mathbb{1}_{\{t<\tau\leq T\}}f(\tau, S_\tau, Y_\tau, X_\tau) \mid \mathcal{G}_t\right] \\ &= \mathbb{1}_{\{\tau>t\}}\frac{1}{S_t}\int_t^T \mathrm{e}^{-r(u-t)}\mathbb{E}\left[f(u, S_u, Y_u, X_u)\, a^\top(c\, Y_u + \gamma^\top X_u) \mid \mathcal{F}_t\right] du. \end{aligned}$$

where τ is the counterparty default time.

Computing the UCVA therefore boils down to a numerical integration of European style option prices. As is the case for CDS options, these option prices can be uniformly approximated as described in Section 3.3.2. We refer to (Brigo, Capponi, and Pallavicini 2014) for a thorough analysis of bilateral counterparty risk valuation in a doubly stochastic default framework.

3.3 The Linear Hypercube Model

The linear hypercube (LHC) model is a single-name model, that is $n = 1$ so that $S_t = Y_t \in (0,1]$. We first present the single-name model and discuss in greater details the one-factor model. We conclude by developing a methodology to price virtually all credit derivatives within this class of models.

The LHC model assumes that the survival process is absolutely continuous, as in Remark 3.2.2, and that the factor process X_t is diffusive and takes values in a hypercube whose edges' length is given by S_t. More formally the state space of (S_t, X_t) is given by

$$E = \left\{(s,x) \in \mathbb{R}^{1+m} : s \in (0,1] \text{ and } x \in [0,s]^m\right\}.$$

The dynamics of (S_t, X_t) is

$$\begin{aligned} dS_t &= -\gamma^\top X_t\, dt \\ dX_t &= (bS_t + \beta X_t)\, dt + \Sigma(S_t, X_t)\, dW_t \end{aligned} \tag{3.19}$$

for some $\gamma \in \mathbb{R}_+^m$ and some m-dimensional Brownian motion W_t, and where the dispersion matrix $\Sigma(S_t, X_t)$ is given by

$$\Sigma(s,x) = \operatorname{diag}\Big(\sigma_1\sqrt{x_1(s-x_1)}, \ldots, \sigma_m\sqrt{x_m(s-x_m)}\Big) \tag{3.20}$$

with volatility parameters $\sigma_1, \ldots, \sigma_m \geq 0$.

Let (S_t, X_t) be an E-valued solution of (3.19). It is readily verified that S_t is non-increasing and that the parameter γ controls the speed at which it decreases

$$0 \leq \gamma^\top X_t \leq \gamma^\top \mathbf{1}\, S_t$$

which implies

$$0 \leq \lambda_t \leq \gamma^\top \mathbf{1} \text{ and } S_t \geq S_0 \mathrm{e}^{-\gamma^\top \mathbf{1}\, t} > 0 \text{ for any } t \geq 0. \tag{3.21}$$

Note that the default intensity upper bound $\gamma^\top \mathbf{1}$ depends on γ, which is estimated from data. Therefore, a crucial step in the model validation procedure is to verify that the range of possible default intensities is sufficiently wide. The following theorem gives conditions on the parameters such that the LHC model (3.19) is well defined.

Theorem 3.3.1. *Assume that, for all $i = 1, \ldots, m$,*

$$b_i - \sum_{j \neq i} \beta_{ij}^- \geq 0, \tag{3.22}$$

$$\gamma_i + \beta_{ii} + b_i + \sum_{j \neq i} (\gamma_j + \beta_{ij})^+ \leq 0. \tag{3.23}$$

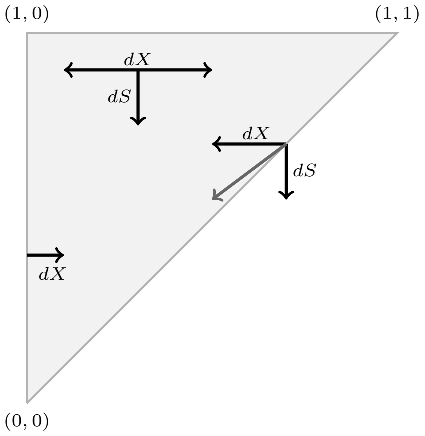


Figure 3.1 – State space of the LHC model with a single factor.
Illustrations of the inward pointing drift conditions at the state space boundaries. The survival process value is given by the y-axis and the factor value by the x-axis.

Then for any initial law of (S_0, X_0) *with support in* E *there exists a unique in law* E*-valued solution* (S_t, X_t) *of* (3.19). *It satisfies the boundary non-attainment, for some* $i = 1, \dots, m$*,*

1. $X_{it} > 0$ *for all* $t \geq 0$ *if* $X_{i0} > 0$ *and*

$$b_i - \sum_{j \neq i} \beta_{ij}^- \geq \frac{\sigma_i^2}{2}, \tag{3.24}$$

2. $X_{it} < S_t$ *for all* $t \geq 0$ *if* $X_{i0} < S_0$ *and*

$$\gamma_i + \beta_{ii} + b_i + \sum_{j \neq i} (\gamma_j + \beta_{ij})^+ \leq -\frac{\sigma_i^2}{2}. \tag{3.25}$$

The state space E is a regular $(m+1)$-dimensional hyperpyramid. Figure 3.1 shows E when $m = 1$ and illustrates the drift inward pointing conditions (3.22)–(3.23) at the boundaries of E.

Remark 3.3.2. *One may consider a more generic specification of the process with a dispersion matrix given by*

$$\Sigma(s, x) = \operatorname{diag}\left(\sigma_1 \sqrt{x_1 (L_1 s - x_1)}, \dots, \sigma_m \sqrt{x_m (L_m s - x_m)}\right)$$

for some positive constants $L_1, \dots, L_m$*. Lemma 3.8.3 shows that such model is observationally equivalent to the above specification.*

In Section 3.7 we describe all possible market price of risk specifications under which the drift function of (S_t, X_t) remains linear.

Remark 3.3.3. *The instantaneous volatility of X_{it} is maximal at the center of its support and decreases to zero at its boundaries. As a consequence, some factors may alternate visits to the lower part and upper part of their supports, and therefore may mimic regime-shifting behaviors.*

3.3.1 One-Factor LHC Model

The default intensity of the one-factor LHC model, $m = 1$, has autonomous dynamics of the form

$$d\lambda_t = \left(\lambda_t^2 + \beta\lambda_t + b\gamma\right) dt + \sigma\sqrt{\lambda_t(\gamma - \lambda_t)}\, dW_t.$$

The diffusion function of λ_t is the same as the diffusion function of a Jacobi process taking values in the compact interval $[0,\gamma]$. However, the drift of λ_t includes a quadratic term that is neither present in Jacobi nor in affine processes.[1] Conditions (3.22)–(3.23) in Theorem 3.3.1 rewrite

$$b \geq 0 \quad \text{and} \quad (\gamma + b + \beta) \leq 0.$$

That is, the drift of λ_t is nonnegative at $\lambda_t = 0$ and nonpositive at $\lambda_t = \gamma$. We can factorize the drift as

$$\lambda_t^2 + \beta\lambda_t + b\gamma = (\lambda_t - \ell_1)(\lambda_t - \ell_2)$$

for some roots $0 \leq \ell_1 \leq \gamma \leq \ell_2$. This way, the default intensity mean-reverts to ℓ_1. The corresponding original parameters are $\beta = -(\ell_1 + \ell_2)$ and $b\gamma = \ell_1\ell_2$, so that the drift of the factor X_t reads

$$\beta S_t + BX_t = (\ell_1 + \ell_2)\left(\frac{\ell_1\ell_2}{\gamma(\ell_1 + \ell_2)} S_t - X_t\right). \tag{3.26}$$

As a sanity check we verify that the constant default intensity case $\lambda_t = \gamma$ is nested as a special case. This is equivalent to have $X_t = S_t$ which can be obtained by specifying the dynamics $dX_t = -\gamma X_t\, dt$ for the factor process and the initial condition $X_0 = 1$. This corresponds to the roots $\ell_1 = 0$ and $\ell_2 = \gamma$.

The dynamics of the standard one-factor affine model on $\mathbb{R}_+$ is

$$d\lambda_t = \ell_2(\ell_1 - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t,$$

where ℓ_2 is the mean-reversion speed and ℓ_1 the mean-reversion level. Figure 3.2 shows the drift and diffusion functions of the default intensity for the one-factor LHC and affine models. The drift function is affine in the affine model whereas it is quadratic in the LHC

---

[1] The Jacobi process has been used in (Delbaen and Shirakawa 2002) to model the short rate in which case the risk-free bond prices are given by weighted series of Jacobi polynomials in the short rate value.

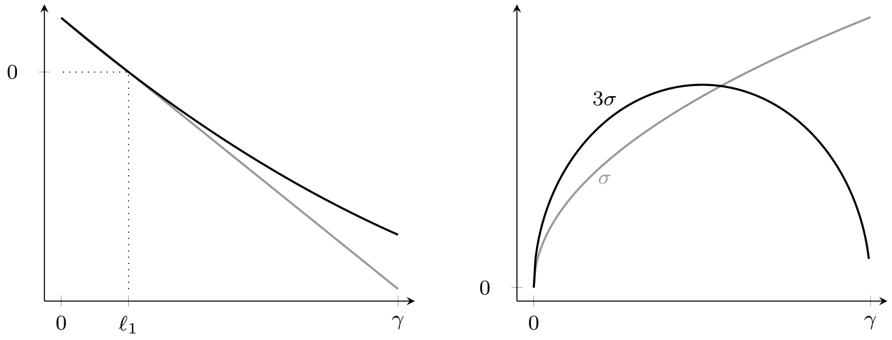


Figure 3.2 – Comparison of the one-factor LHC and CIR models.
Drift and diffusion functions of the default intensity for the one-factor LHC model (black line) and affine model (grey line). The parameter values are $\ell_1 = 0.05$, $\ell_2 = 1$, and $\gamma = 0.25$. s

model. However, for reasonable parameters values, the drift functions look similar when the default intensity is smaller than the mean-reversion level ℓ_1. On the other hand, the force of mean-reversion above ℓ_1 is smaller and concave in the LHC model. The diffusion function is strictly increasing and concave for the affine model whereas it has a concave semi-ellipse shape in the LHC model. The diffusion functions have the same shape on $[0, \gamma/2]$ but typically do not scale equivalently in the parameter σ. Note that the parameter γ can always be set sufficiently large so that the likelihood of λ_t going above $\gamma/2$ is arbitrarily small.

3.3.2 Option Price Approximation

We saw in Sections 3.2.4 and 3.2.5 that the pricing of a CDS option or a UCVA boils down to computing a $\mathscr{F}_t$-conditional expectation of the form

$$\Phi(f; t, T) = \mathbb{E}\left[f(S_T, X_T) \mid \mathscr{F}_t\right]$$

for some continuous function $f(y,x)$ on E. We now show how to approximate $\Phi(f; t, T)$ in analytical form by means of a polynomial approximation of $f(y,x)$. The methodology presented hereinafter applies to any linear credit risk model which has a compact the state space E and for which the $\mathscr{F}_t$-conditional higher moments of (Y_T, X_T) are computable.

To this end, we first recall how the factors moments can be efficiently computed as described in (Filipović and Larsson 2016). Denote by $\mathrm{Pol}_n(E)$ the set of polynomials $p(s,x)$ on E of degree n or less. It is readily seen that the generator of (S_t, X_t),

$$\mathscr{G} f(s,x) = \left(-\gamma^\top x, (\beta s + Bx)^\top\right) \nabla f(s,x) + \frac{1}{2}\sum_{i=1}^{m} \frac{\partial^2 f(s,x)}{\partial x_i^2} \sigma_i^2 x_i (s - x_i), \tag{3.27}$$

is polynomial in the sense that

$$\mathscr{G}\mathrm{Pol}_n(E) \subset \mathrm{Pol}_n(E) \quad \text{for any } n \in \mathbb{N}.$$

Let $N_n = \binom{n+1+m}{n}$ denote the dimension of $\mathrm{Pol}_n(E)$ and fix a polynomial basis $\{h_1, \dots, h_{N_n}\}$ of $\mathrm{Pol}_n(E)$. We define the function of (s, x)

$$H_n(s,x) := (h_1(s,x), \dots, h_{N_n}(s,x))^\top$$

with values in $\mathbb{R}^{N_n}$. There exists a unique matrix representation G_n of $\mathscr{G}|_{\mathrm{Pol}_n(E)}$ with respect to this polynomial basis such that for any $p \in \mathrm{Pol}_n(E)$ we can rewrite

$$\mathscr{G} p(s,x) = H_n(s,x)^\top G_n \vec{p}$$

where $\vec{p}$ is the coordinate representation of p. (Filipović and Larsson 2016, Theorem (3.1)) then states that for any $t \le T$ we have

$$\mathbb{E}\left[p(S_T, X_T) \mid \mathscr{F}_t\right] = H_n(S_t, X_t)^\top \mathrm{e}^{G_n(T-t)} \vec{p}. \tag{3.28}$$

Remark 3.3.4. *The choice for the basis $H_n(x,s)$ of $Pol_n(E)$ is arbitrary and one may simply consider the monomial basis,*

$$H_n(s,x) = \{1,\, s,\, x_1,\, \dots,\, x_m,\, s^2,\, s x_1,\, x_1^2,\, \dots,\, x_m^n\}$$

in which G_n is block-diagonal. There are efficient algorithms to compute the matrix exponential $\mathrm{e}^{G_n(T-t)}$, see for example (Higham 2008). Note that only the action of the matrix exponential is required, that is $\mathrm{e}^{G_n(T-t)} \vec{p}$ for some $p \in Pol_n(E)$, for which specific algorithms exist as well, see for examples (Al-Mohy and Higham 2011) and (Sidje 1998) and references within.

Now let $\epsilon > 0$. From the Stone-Weierstrass approximation theorem (Rudin 1974, Theorem 5.8) there exists a polynomial $p \in \mathrm{Pol}_n(E)$ for some n such that

$$\sup_{(s,x)\in E} \left| f(s,x) - p(s,x) \right| \le \epsilon. \tag{3.29}$$

Combining (3.28) and (3.29) we obtain the desired approximation of $\Phi(f; t, T)$.

Theorem 3.3.5. *Let $p \in \mathrm{Pol}_n(E)$ be as in* (3.29)*. Then $\Phi(f; t, T)$ is uniformly approximated by*

$$\sup_{t \le T} \left\| \Phi(f; t, T) - H_n(S_t, X_t)^\top \mathrm{e}^{G_n(T-t)} \vec{p} \right\|_{L^\infty} \le \epsilon. \tag{3.30}$$

The approximating polynomial p in (3.29) needs to be found case by case. We illustrate this for the CDS option (3.16) in Section 3.5.2 and for the CDIS (3.17) option on an homogenous portfolio in Section 3.5.3.

Remark 3.3.6. *Approximating the payoff function $f(s,x)$ on a strict subset of the state space E is sufficient to approximate an option price. Indeed, for any times $t \le u \le T$ the process (S_u, X_u) takes values in*

$$\{(s,x) \in E : S_t \ge s \ge \mathrm{e}^{-\gamma^\top \mathbf{1}(T-t)} S_t\} \subset E.$$

A polynomial approximation on a compact set smaller than E can be expected to be more precise and, as a result, to produce a more accurate price approximation.

3.4 Extensions

In this section we describe several model extensions. We first construct multi-name models, then introduce stochastic interest rates, and finally include jumps to generate simultaneous defaults.

3.4.1 Multi-Name Models

We build upon the LHC model to construct multi-name models with correlated default intensities and which can easily accommodate the inclusion of new factors and firms. Yet this approach can be applied to other linear credit risk models, as long as they belong to the class of polynomial models. We consider d independent pairs of processes

$$(Y_t^1, X_t^1), \dots, (Y_t^d, X_t^d) \tag{3.31}$$

where each pair $(Y_t^j, X_t^j) \in \mathbb{R}_+^{1+m_j}$ is defined as (S_t, X_t) in Section 3.3. We therefore have that $Y_t = (Y_t^1, \dots, Y_t^d)$ with $Y_0 = \mathbf{1}$, and $X_t = (X_t^1, \dots, X_t^d)$ with $X_0 \in [0,1]^m$ where $m = \sum_{i=1}^d m_i$. In the following E denotes the state space of (Y_t, X_t) and E_Y its restriction to Y_t.

Denote $h_t = (h_t^1, \dots, h_t^d)$ the $\mathbb{R}_+^d$-valued process whose j-th component is given by the j-th factor implied intensity,

$$h_t^j = \frac{\gamma^{j\top} X_t^j}{Y_t^j}$$

where the vector $\gamma^j \in \mathbb{R}^{m_j}$ is the drift parameter of Y_t^j as (3.19).

Linear Construction

The survival process of the firm $i = 1, \dots, N$ can be defined as in (3.1), $S_t^i = a_i^\top Y_t$, for some vector $a_i \in \mathbb{R}_+^n$ satisfying $a^\top \mathbf{1} = 1$. The implied default intensity λ_t^i of firm i is a convex sum of the factor intensities,

$$\lambda_t^i = w_t^{i\top} h_t$$

with the weights $w_{jt}^i = a_{ij} Y_t^j / S_t^i > 0$ satisfying $\sum_{j=1}^d w_{jt}^i = 1$.

Polynomial Construction

We show that polynomial specifications in (Y_t, X_t) of the survival process are equivalent to the linear specification (3.1)–(3.3) with an extended factors representation. Fix an integer n and define the survival process of each firm $i = 1, \dots, N$ as follows,

$$S_t^i = p_i(Y_t) \tag{3.32}$$

for some polynomial $p_i(y) \in \mathrm{Pol}_n(E_Y)$ which is componentwise non-increasing and positive on E_Y, and such that $p_i(\mathbf{1}) = 1$. Let $H_n(y,x)$ be a polynomial basis of $\mathrm{Pol}_n(E)$ stacked in a row vector and of the form

$$H_n(y,x) = (H_n(y),\, H_n^*(y,x)) \tag{3.33}$$

where $H_n(y)$ is itself a polynomial basis of $\mathrm{Pol}_n(E_Y)$. Then,

$$S_t^i = a_i^\top \mathscr{Y}_t$$

with the finite variation process $\mathscr{Y}_t = H_n(Y_t)$, the factor process $\mathscr{X}_t = H_n^*(Y_t, X_t)$ and where the vector a_i is given by the equation $p_i(y) = H_n(y)\, a_i$. It follows from the polynomial property that the process $(\mathscr{Y}_t, \mathscr{X}_t)$ has a linear drift as in (3.2)–(3.3), where the explicit values for the vector a_i and for the matrix $\mathscr{A}$ defining its drift as in (3.5) will depend on the choice of polynomial basis $H_n(y,x)$.

Example 3.4.1. *Take $p(y) = y^\alpha = \prod_{i=1}^d y_i^{\alpha_i}$ for some non negative integers $\alpha \in \mathbb{N}^d$, then the implied default intensity of is a linear sum of the factor intensities $\lambda_t = \alpha^\top h_t$.*

Remark 3.4.2. *The dimension of $H_n(y,x)$ is $\binom{n+d+m}{n}$ and may be large depending on the values of $m+d$ and n. However, given that the pairs (Y_t^i, X_t^i) in (3.31) are independent, the conditional expectation of a monomial rewrites*

$$\mathbb{E}\left[\prod_{i=1}^{d} (Y_T^i)^{\alpha_i} (X_T^i)^{\beta_i} \mid \mathscr{F}_t\right] = \prod_{i=1}^{d} \mathbb{E}\left[(Y_T^i)^{\alpha_i} (X_T^i)^{\beta_i} \mid \mathscr{F}_t\right], \quad T > t,$$

for some $\alpha_i \in \mathbb{N}$ and $\beta_i \in \mathbb{N}^{m_j}$ for all $i = 1, \dots, d$. Hence, to compute bonds and CDSs prices we only need to consider d independent polynomial bases of total size at most equal to $\sum_{i=1}^d \binom{n+1+m_i}{n}$.

3.4.2 Stochastic Interest Rates

A stochastic short rate model can be used to model stochastic interest rates while preserving the model tractability. We denote the discount process $D_t = \exp(-\int_0^t r_s ds)$ where r_s is the short rate value at time s. Assuming that the short rate is nonnegative, the discount process has the same properties as the survival process and may thus be modeled in a similar way and, possibly, with the same factors Y_t. More precisely, we assume that $D_t = a_r^\top Y_t$ for some $a_r \in \mathbb{R}_+^n$. We follow Section 3.4.1 and let $H_2(y,x)$ be a polynomial basis of $\mathrm{Pol}_2(E)$ which defines a new linear credit risk model $(\mathscr{Y}_t, \mathscr{X}_t) = (H_2(Y_t), H_2^*(Y_t, X_t))$ whose linear drift is given by a matrix $\mathscr{A}$ as in (3.5).

Proposition 3.4.3. *The pricing vectors with stochastic interest rates are defined with respect to* $(\mathscr{Y}_t, \mathscr{X}_t)$ *and given by*

$$\psi_{\mathrm{Z}}(t,T)^\top = \begin{pmatrix} a_{\mathrm{Z}}^\top & 0 \end{pmatrix} \mathrm{e}^{\mathscr{A}(T-t)}$$

where the vector a_{Z} *is given by* $H_2(y)^\top a_{\mathrm{Z}} = (a_r^\top y)(a^\top y)$,

$$\psi_{\mathrm{D}}(t,T)^\top = a_{\mathrm{D}}^\top \left(\int_t^T \mathrm{e}^{\mathscr{A}(s-t)} ds \right), \quad \psi_{\mathrm{D}_*}(t,T)^\top = a_{\mathrm{D}}^\top \left(\int_t^T s\,\mathrm{e}^{\mathscr{A}(s-t)} ds \right),$$

where the vector a_{D} *is given by* $H_2(y,x)\, a_{\mathrm{D}} = (a_r^\top y) \left(-a^\top \begin{pmatrix} c\,y & \gamma\,x \end{pmatrix} \right)$ *for the dynamics* (3.2). *The pricing formulas* (3.6), (3.7), *and* (3.9) *then apply with* $r = 0$, *and* $a_{\mathrm{Z}}^\top \mathscr{Y}_t$ *in place of* S_t.

In practice it can be sufficient to consider a basis strictly smaller than $H_2(y,x)$, as the following example suggests.

Example 3.4.4. *Consider two independent LHC models* (Y_t^j, X_t^j) *with* $m_j = 1$ *for* $j \in \{1,2\}$, *and consider the following linear credit risk model with stochastic interest rate,*

$$D_t = Y_t^1 \quad \text{and} \quad S_t = v\,Y_t^1 + (1-v)\,Y_t^2 \tag{3.34}$$

for some constant $v \in (0,1)$. *The calculation of bond and CDS prices only requires the subbases*

$$H_0(y,x) = \begin{pmatrix} y_1^2 & y_1 y_2 \end{pmatrix}, \quad H_1(y,x) = \begin{pmatrix} y_1 x_1 & y_1 x_2 & x_1 y_2 & x_1^2 & x_1 x_2 \end{pmatrix},$$

whose total dimension is $\dim(H(y,x)) = 7 < \dim(\mathrm{Pol}_2(E)) = 15$. *The drift term of the process* $H(Y_t, X_t)$ *is*

$$\mathscr{A} = \begin{pmatrix} 0 & 0 & -2\gamma_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\gamma_2 & -\gamma_1 & 0 & 0 \\ b_1 & 0 & \beta_1 & 0 & 0 & -\gamma_1 & 0 \\ 0 & b_2 & 0 & \beta_2 & 0 & 0 & -\gamma_1 \\ 0 & b_1 & 0 & 0 & \beta_1 & 0 & 0 \\ \sigma_1^2 & 0 & 2b_1 - \sigma_1^2 & 0 & 0 & 2\beta_1 & 0 \\ 0 & 0 & 0 & b_1 & b_2 & 0 & \beta_1 + \beta_2 \end{pmatrix}$$

where the subscripts indicate the LHC model identity. The pricing vectors in this basis are

$$a_{\mathrm{Z}} = \begin{pmatrix} v & 1-v \end{pmatrix} \quad \text{and} \quad a_{\mathrm{D}} = \begin{pmatrix} 0 & 0 & -v\gamma_1 & -(1-v)\gamma_2 & 0 & 0 & 0 \end{pmatrix}.$$

3.4.3 Jumps and Simultaneous Defaults

There are two ways to include jumps in the survival process dynamics that may result in simultaneous defaults. The first one is to let the martingale part of Y_t be driven by a jump process so that multiple survival processes may jump at the same time. The second is to let time run with a stochastic clock leaping forward hence producing synchronous jumps in the factors and the survival processes.

Let (Y_t, X_t) be a LHC model as in (3.19) whose parameters γ, β, B satisfy Theorem 3.3.1, and let Z_t be a nondecreasing Lévy process with Lévy measure $\nu^Z(d\zeta)$ and drift $b^Z \geq 0$ that is independent from the natural filtration of the Brownian motion W_t and the uniform random variables $U^1, \dots, U^N$.

Jump-Diffusion Model

Assume that $\Delta Z_t \leq 1$, we define the dynamics of the LHC model with jumps as follows

$$d\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} 0 & -\gamma^\top \\ b & \beta \end{pmatrix} \begin{pmatrix} Y_{t-} \\ X_{t-} \end{pmatrix} dt + \begin{pmatrix} 0 \\ \Sigma(Y_{t-}, X_{t-}) \end{pmatrix} dW_t - \begin{pmatrix} c\,Y_{t-} + \delta^\top X_{t-} \\ \operatorname{diag}(\nu) X_{t-} \end{pmatrix} dZ_t$$

for some $c > 0$, $\delta \in \mathbb{R}^m_+$, and $\nu \in \mathbb{R}^m_+$ such that

$$c + \delta^\top \mathbf{1} < 1, \quad c + \delta^\top \mathbf{1} \leq \nu_i \leq 1, \qquad i = 1, \dots, m \tag{3.35}$$

$$\text{and } \nu_i < 1 \text{ if (3.24) applies}, \quad i = 1, \dots, m \tag{3.36}$$

Conditions (3.35)–(3.36) ensure that the process always jumps inside its state space. The same process Z_t can affect the dynamics of different (Y_t, X_t) processes. Reciprocally, multiple jump components can be included in the same (Y_t, X_t).

The linear credit risk model representation of the model is then as follows,

$$d\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} -c & -(\gamma + \delta\,\mathbb{E}[Z_1])^\top \\ b & \beta - \operatorname{diag}(\nu)\,\mathbb{E}[Z_1] \end{pmatrix} \begin{pmatrix} Y_{t-} \\ X_{t-} \end{pmatrix} dt + \begin{pmatrix} 0 \\ \Sigma(Y_{t-}, X_{t-}) \end{pmatrix} dW_t - \begin{pmatrix} c\,Y_{t-} + \delta^\top X_{t-} \\ \operatorname{diag}(\nu) X_{t-} \end{pmatrix} dN_t$$

with the martingale $N_t = Z_t - \mathbb{E}[Z_1]\,t$.

Model with Stochastic Clock

We consider the time-changed process $(\bar{Y}_t, \bar{X}_t) = (Y_{Z_t}, X_{Z_t})$. The survival process in the time-changed LHC model is similarly defined by

$$S_t = \bar{Y}_t$$

with the factor dynamics

$$\begin{pmatrix} d\bar{Y}_t \\ d\bar{X}_t \end{pmatrix} = \bar{A} \begin{pmatrix} \bar{Y}_t \\ \bar{X}_t \end{pmatrix} dt + \begin{pmatrix} 0 \\ dM_t^{\bar{X}} \end{pmatrix}$$

where the $(m+n) \times (m+n)$-matrix $\bar{A}$ is now given by

$$\bar{A} = b^Z A + \int_0^\infty \left(\mathrm{e}^{A\zeta} - \mathrm{Id} \right) \nu^Z(d\zeta) \tag{3.37}$$

with the matrix A as in Equation (3.5), see (Sato 1999, Chapter 6) and (Filipović and Larsson 2017). The time-changed LHC model remains a linear credit risk model. The background filtration $\mathscr{F}_t$ is now the natural filtration of the process (Y_{Z_t}, X_{Z_t}). In general, the matrix $\bar{A}$ will have to be computed by numerical integration, but in the following example it admits an explicit expression.

Example 3.4.5. *Let Z_t be a Gamma process such that $\nu^Z(d\zeta) = \gamma_Z \zeta^{-1} \mathrm{e}^{-\lambda_Z \zeta} d\zeta$ for some constants $\lambda_Z, \gamma_Z > 0$ and $b^Z = 0$. If the eigenvalues of the matrix A have nonpositive real parts, the drift of the time changed process (Y_{Z_t}, X_{Z_t}) is then equal to*

$$\bar{A} = -\gamma_Z \log\left(\mathrm{Id} - A\lambda_Z^{-1}\right). \tag{3.38}$$

Survival processes built from independent LHC models can be time changed with the same stochastic clock Z_t in order generate simultaneous defaults and thus default correlation.

3.5 Case Studies

We show that LHC models are able to capture complex term structure dynamics and that option prices can be accurately approximated. First, we fit parsimonious multi-factor models to CDS data and discuss the estimated parameters and factors. Second, we accurately approximate the price of CDS options at different moneyness for the three-factor model. The methodology is then extended to approximate the payoff function of a CDS index option on homogeneous portfolios.

3.5.1 CDS Calibration

In this section we calibrate the multi-factor LHC model to a high yield firm, Bombardier Inc., and also to an investment grade firm, Walt Disney Co., in order to show that the model flexibly adjusts to different spread levels and dynamics. We also present a fast filtering and calibration methodology which is specific to linear credit risk models.

	all	1 yr	2 yrs	3 yrs	4 yrs	5 yrs	7 yrs	10 yrs
Mean	274.51	144.07	194.80	243.38	279.43	329.40	357.10	373.71
Vol	165.23	156.66	158.95	153.31	147.95	141.14	130.46	121.64
Median	244.76	94.79	145.71	189.55	232.44	295.51	353.01	376.58
Min	28.02	28.02	39.22	59.50	86.64	109.58	146.32	171.29
Max	1288.71	1288.71	1151.92	1092.74	1062.57	1048.33	960.16	887.06

(a) Bombardier Inc.

	all	1 yr	2 yrs	3 yrs	4 yrs	5 yrs	7 yrs	10 yrs
Mean	31.01	11.97	17.53	22.74	28.90	34.59	45.00	56.18
Vol	21.85	12.93	15.73	17.18	18.18	18.15	16.13	15.66
Median	26.30	7.70	12.42	17.39	24.31	30.45	42.98	55.58
Min	1.63	1.63	3.24	4.47	5.81	8.18	12.92	17.51
Max	133.02	79.38	102.20	115.19	120.62	126.43	127.22	133.02

(b) Walt Disney Co.

Table 3.1 – CDS spreads summary statistics.

The sample contains 552 weekly observations collected between January 1st 2005 and January 1st 2015 summing up to 3620 CDS spreads in basis point for each firm.

Data

The empirical analysis is based on composite CDS spread data from Markit which are essentially averaged quotes provided by major market makers. At each date we include the available spreads with the modified restructuring clause on contracts with maturities of 1, 2, 3, 4, 5, 7, and 10 years. The sample starts on January 1th 2005 and ends on January 1th 2015. The data set contains 552 weekly observations summing up to 3620 observed CDS spreads for each firm.

Time series of the 1-year, 5-year, and 10-year CDS spreads are displayed in Figure 3.3, as well as the relative changes on the 5-year versus 1-year CDS spread. The two term-structures of CDS spreads exhibit important fluctuations of their level, slope, and curvature. The time series can be split into three time periods. The first period, before the subprime crisis, exhibits low spreads in contango and low volatility. The second period, during the subprime crisis, exhibits high volatility with skyrocketing spreads temporarily in backwardation. The crisis had a significantly larger impact on the high yield firm for which the spreads have more than quadrupled. The more recent period is characterized by a steep contango and a lot of volatility. Figure 3.3 also shows that CDS spread changes are strongly correlated across maturities. Summary statistics are reported in Table 3.1.

Model Specifications

The risk neutral dynamics of each survival process is given by the LHC model of Section 3.3 with two and three factors. We set $\gamma = \gamma_1 \boldsymbol{e}_1$, for some $\gamma_1 \geq 0$, and consider a cascading structure

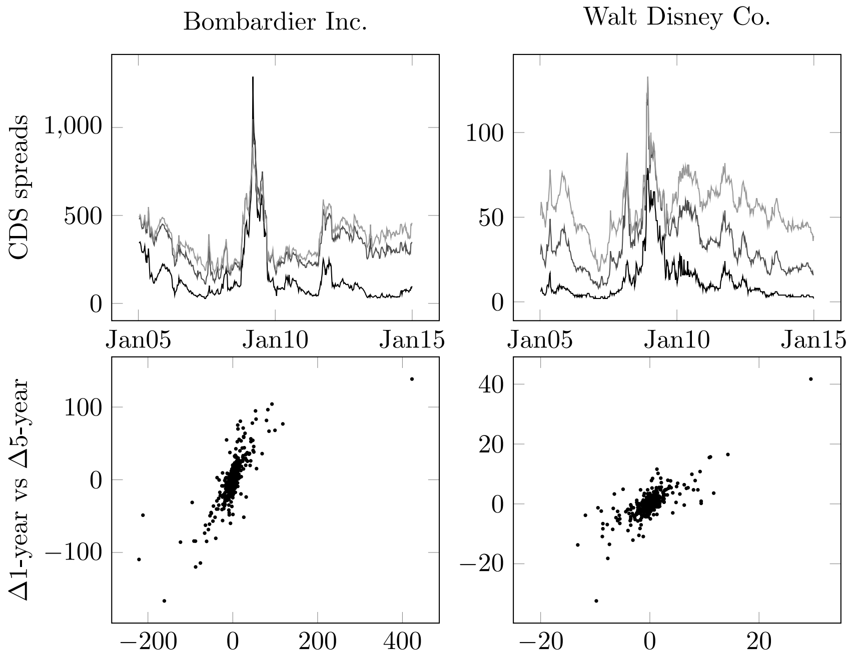


Figure 3.3 – CDS spreads data.
The first row displays the CDS spreads in basis points for the maturities 1 year (black), 5 years (grey), and 10 years (light-grey). The second row displays the weekly changes in 1-year versus 5-year CDS spreads.

of the form

$$dX_{it} = \kappa_i(\theta_i X_{(i+1)t} - X_{it})\,dt + \sigma_i\sqrt{X_{it}(S_t - X_{it})}\,dW_{it} \tag{3.39}$$

for $i = 1,\dots,m-1$ and

$$dX_{mt} = \kappa_m(\theta_m S_t - X_{mt})\,dt + \sigma_m\sqrt{X_{mt}(S_t - X_{mt})}\,dW_{mt} \tag{3.40}$$

for some parameters $\kappa,\theta,\sigma \in \mathbb{R}^m_+$ satisfying

$$\theta_i \le 1 - \frac{\gamma_1}{\kappa_i} \tag{3.41}$$

for $i = 1,\dots,m$ so that conditions (3.22)-(3.23) are satisfied. Indeed, we have that $\beta_{ii} = -\kappa_i$, $\beta_{i,i+i} = \kappa_i\theta_i$, and $\beta_{ij} = 0$ otherwise, $b_m = \kappa_m\theta_m$ and $b_i = 0$ otherwise. It directly follows that

$$0 \le b_i - \sum_{j\neq i}\beta^-_{ij} = \mathbb{1}_{\{i=m\}}\kappa_m\theta_m = \mathbb{1}_{\{i=m\}}\beta_{mm}$$

and for $i = 1\dots,m$

$$0 \ge \gamma_i + \beta_{ii} + b_i + \sum_{j\neq i}(\gamma_j + \beta_{ij})^+ = \gamma_1 - \kappa_i + \kappa_i\theta_i = \gamma_1 + \beta_{ii} + \mathbb{1}_{\{i\neq m\}}\beta_{i,i+1} + \mathbb{1}_{\{i=m\}}b_m.$$

The parameter conditions (3.22)-(3.23) therefore boil down to standard linear parameter constraints when expressed with β and b. They are therefore compatible with a large number of optimization algorithms.

This specification allows default intensity values to persistently be close to zero over extended periods of time. It also allows to work with a multidimensional model parsimoniously as the number of free parameters is equal to $3m+1$ whereas it is equal to $3m+m^2$ for the generic LHC model. The default intensity is then proportional to the first factor and given by $\lambda_t = \gamma_1 X_{1t}/S_t$.

We denote the two-factor and three-factor linear hypercube cascade models by LHCC(2) and LHCC(3), respectively. In addition, we estimate a three-factor model with parameter γ_1 fixed that we denote by LHCC(3)*. This parameter value is fixed so as to be about twice as large as the estimated γ_1 from the LHCC(3) model. We estimate the constrained model in order to determine whether the choice of the default intensity upper bound is critical for the empirical results.

We set the risk-free rate equal to the average 5-year risk-free yield over the sample, $r = 2.52\%$. We make the usual assumption that the recovery rate is equal to $\delta = 40\%$. We also use Lemma 3.2.7 to compute efficiently the CDS spreads, which is justified by the following Proposition.

Proposition 3.5.1. *Assume that* $r > 0$, *then the matrix* $A^* = A - r\,\mathrm{Id}$ *with* A *as in* (3.5) *is invertible for the cascade LHCC model defined in* (3.39)–(3.40) *and with* $\gamma = \gamma_1 \mathrm{e}_1$.

Filtering and Calibration

We present an efficient methodology to filter the factors from the CDS spreads. We recall that the CDS spread $\mathrm{CDS}(t, T_0, T_M)$ is the strike spread that renders the initial values of the CDS contract equal to zero, we therefore obtain the affine equation

$$\psi_{\mathrm{cds}}(t, T_0, T_M, \mathrm{CDS}(t, T_0, T_M))^\top \begin{pmatrix} 1 \\ Z_t \end{pmatrix} = 0$$

under the assumption that $\tau > t$, and with the normalized factor $Z_t = X_t/S_t \in [0,1]^m$. Therefore, in principle we could extract the pseudo factor Z_t values from the observation of at least m spreads with different maturities. The factor (S_t, X_t) values can then be inferred, for example, by applying the Euler scheme to compute the survival process value, for example, and then rescaling the pseudo factor,

$$S_{t_i} = S_{t_{i-1}} - \gamma^\top X_{t_{i-1}} \Delta t \quad \text{and} \quad X_{t_i} = S_{t_i} Z_{t_i} \tag{3.42}$$

for all the observation dates t_i, and with $S_{t_0} = 1$. In practice, we consider all the observable spreads and minimize the following weighted mean squared error

$$\begin{aligned} \min_{z} \quad & \frac{1}{2} \sum_{k=1}^{N^i} \left(\psi_{\mathrm{cds}}(t_i, t_i, T_k^i, \mathrm{CDS}(t_i, t_i, T_k^i))^\top \begin{pmatrix} 1 \\ z \end{pmatrix} \Big/ \psi_{\mathrm{prem}}(t_i, t_i, T_k)^\top \begin{pmatrix} 1 \\ Z_{t_{i-1}} \end{pmatrix} \right)^2 \\ \text{s.t.} \quad & 0 \leq z_i \leq 1, \; i = 1, \ldots, m \end{aligned} \tag{3.43}$$

where $T_1^i, \ldots, T_{N^i}^i$ are the maturities of the N^i observed spreads at date t_i, and t_{i-1} is the previous observation date. Dividing the CDS price error by an approximation of the CDS premium leg value gives an accurate approximation of the CDS spread error when $Z_{t_i} \approx Z_{t_{i-1}}$. The above minimization problem is a linearly constrained quadratic optimization problem which can be solved virtually instantaneously numerically.

For any parameter set we can extract the observable factor process at each date by recursively solving (3.43) and applying (3.42). With the parameters and the factor process values we can in turn compute the difference between the model and market CDS spreads. Therefore, we numerically search the parameter set that minimizes the aggregated CDS spread root-mean-squared-error (RMSE) by using the gradient-free Nelder-Mead algorithm together with a penalty term to enforce the parameter constraints and starting from several randomized initial parameter sets. Note that we do not calibrate the volatility parameters σ_i for $i = 1, \ldots, m$ since CDS spreads do not depend on the martingale components with linear credit risk models and since the factor process is observable directly from the CDS spreads. The total number of parameters for LHCC(2), LHCC(3), and LHCC(3)* model is therefore equal to 5, 7, and 6 respectively. Equipped with a fast filter and a low dimensional parameter space, the calibration procedure is swift.

Remark 3.5.2. *Alternatively one could estimate the parameters using the generalized method of moments, performing a quasi-maximum likelihood estimation for example. This can be implemented in a straightforward manner with the LHC model if the market price of risk specification preserves the polynomial property of the factors as the real world conditional moments of* (S_t, X_t) *would also then be analytical, see Section 3.7. However this approach comes at the cost of more parameters and possibly more stringent conditions on them.*

Parameters, Fitted Spreads, and Factors

The fitted parameters are reported in Table 3.2. An important observation is that the parameter constraint in Equation (3.41) is binding for each dimension in all the fitted models. Besides that, for both firms, it seems that there is some consistency in the calibrated parameter values, and the calibrated default intensity upper bounds appear large enough to cover the high spread values observed during the subprime crisis.

The fitted factors extracted from the calibration are used as input to compute the fitted spreads. With the fitted spreads we compute the fitting errors for each date and maturity. Not

	LHCC(2)	LHCC(3)	LHCC(3)*
γ_1	0.205	0.201	*0.400*
κ_1	0.546	1.263	1.316
κ_2	0.421	0.668	0.884
κ_3		0.385	0.668
θ_1	0.624	0.841	0.696
θ_2	0.512	0.699	0.548
θ_3		0.478	0.401

(a) Bombardier Inc.

	LHCC(2)	LHCC(3)	LHCC(3)*
γ_1	0.056	0.064	*0.130*
κ_1	0.167	0.258	0.294
κ_2	0.165	0.229	0.280
κ_3		0.091	0.212
θ_1	0.666	0.753	0.558
θ_2	0.662	0.721	0.536
θ_3		0.298	0.387

(b) Walt Disney Co.

Table 3.2 – Fitted parameters for the LHC models.

surprisingly the more flexible specification LHCC(3) seems to perform the best. Estimating the default intensity upper bound γ_1 instead of setting an arbitrarily large value improves the calibration. Table 3.3 reports summary statistics of the errors by maturity. The LHCC(3) model has the smallest RMSE for each maturity. In particular, its overall RMSE is half the one of the two-factor model. The LHCC(3)* model faces difficulties in reproducing long-term spreads as, for example, its RMSE is twice as large as the one of the unconstrained LHCC(3) for the 10-year maturity spread for both firms. Figure 3.4 displays the fitted spreads and the RMSE time series. Again, the LHCC(3) appears to have the smallest level of errors over time. The two other models do not perform as good during the low spreads period before the financial crisis, and during the recent volatile period. Overall, the fitted models appear to reproduce relatively well the observed CDS spread values.

Figure 3.5 shows the estimated factors. They are remarkably similar across the different specifications. The default intensity explodes and the survival process decreases rapidly during the financial crisis. The m-th factor controls the long term default intensity level. The second factors controls the medium term behavior of the term-structure of credit risk in the three-factor models. The two-factor model requires an almost equal to zero default intensity to capture the steep contango of the term structure at the end of the sample period, even lower than before the financial crisis. This seems counterfactual and illustrates the limitations of the two-factor model in capturing changing dynamics. The m-th factor visits the second half of its support $[0, S_t]$ and appears to stabilize in this region for the three models.

		all	1 yr	2 yrs	3 yrs	4 yrs	5 yrs	7 yrs	10 yrs
LHCC(2)	RMSE	26.24	23.87	31.79	24.13	12.31	24.36	27.70	33.33
	Median	-0.22	-13.90	-3.16	-1.23	4.63	20.20	-0.17	-18.90
	Min	-83.96	-64.23	-83.96	-65.09	-22.09	-20.50	-38.64	-79.80
	Max	123.86	123.86	43.98	32.90	39.31	57.07	75.58	54.45
LHCC(3)	RMSE	16.10	8.90	19.63	19.46	11.01	17.35	15.93	16.94
	Median	-0.25	1.14	-7.69	-5.47	1.06	16.46	2.06	-9.42
	Min	-56.64	-24.62	-56.64	-52.93	-31.01	-0.66	-12.85	-46.56
	Max	107.23	107.23	23.86	15.42	20.38	41.61	49.57	31.94
LHCC(3)*	RMSE	21.87	9.07	23.52	24.01	12.67	16.56	25.15	32.37
	Median	-0.42	0.02	-4.22	-3.94	-3.12	14.22	-0.66	-4.80
	Min	-82.13	-24.32	-66.96	-68.24	-32.91	-31.95	-54.44	-82.13
	Max	67.51	24.43	25.10	26.16	22.24	42.51	67.51	59.33

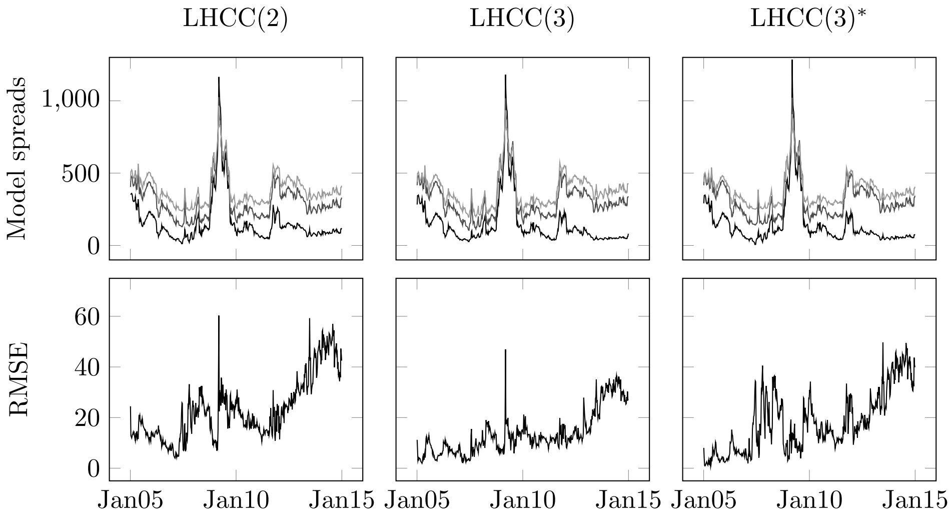
(a) Bombardier Inc.

		all	1 yr	2 yrs	3 yrs	4 yrs	5 yrs	7 yrs	10 yrs
LHCC(2)	RMSE	2.88	3.09	1.66	2.73	2.82	2.82	2.00	4.30
	Median	-0.33	-0.13	-0.86	-1.99	-1.40	-0.43	1.40	1.10
	Min	-12.65	-12.65	-4.15	-5.21	-4.34	-4.32	-5.54	-12.64
	Max	8.81	3.58	5.11	8.81	8.70	8.22	4.62	6.43
LHCC(3)	RMSE	1.06	0.85	1.09	1.02	0.89	1.31	1.33	0.75
	Median	-0.03	0.35	0.19	-0.55	-0.43	0.14	0.70	-0.26
	Min	-5.57	-4.87	-5.57	-3.53	-3.55	-4.34	-4.62	-1.97
	Max	4.94	2.74	4.94	3.58	4.34	3.85	3.53	2.68
LHCC(3)*	RMSE	1.17	1.02	1.11	0.98	1.15	1.62	1.07	1.12
	Median	0.01	0.47	0.35	-0.62	-0.60	-0.06	0.48	-0.02
	Min	-5.48	-5.45	-5.48	-3.49	-3.78	-4.83	-3.92	-4.65
	Max	4.63	2.68	4.49	3.28	4.63	3.98	2.98	4.15

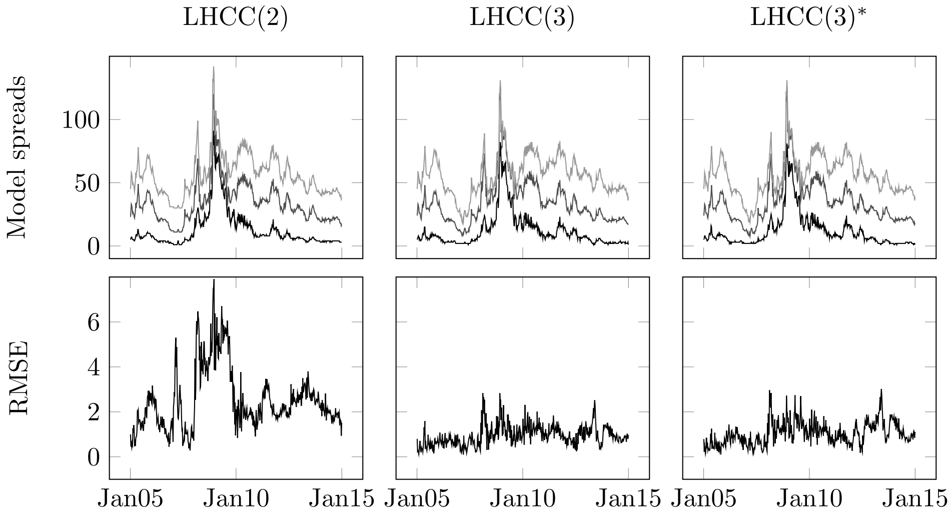
(b) Walt Disney Co.

Table 3.3 – Comparison of CDS spreads fits for the LHC models.

The tables report the minimal, maximal, median, and root mean squared errors in basis point by maturity over the entire time period for the three different specifications.

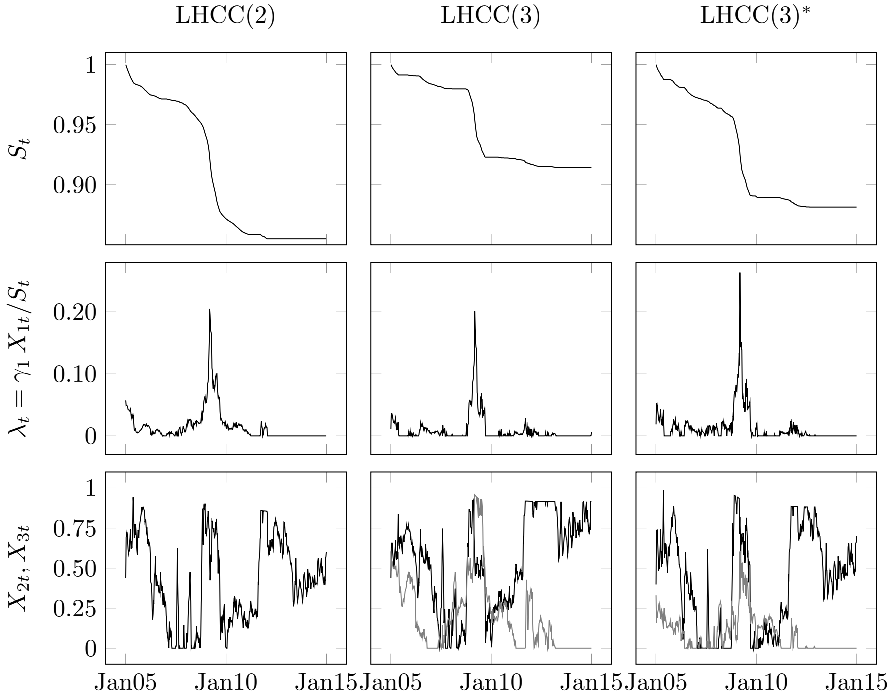


(a) Bombardier Inc.

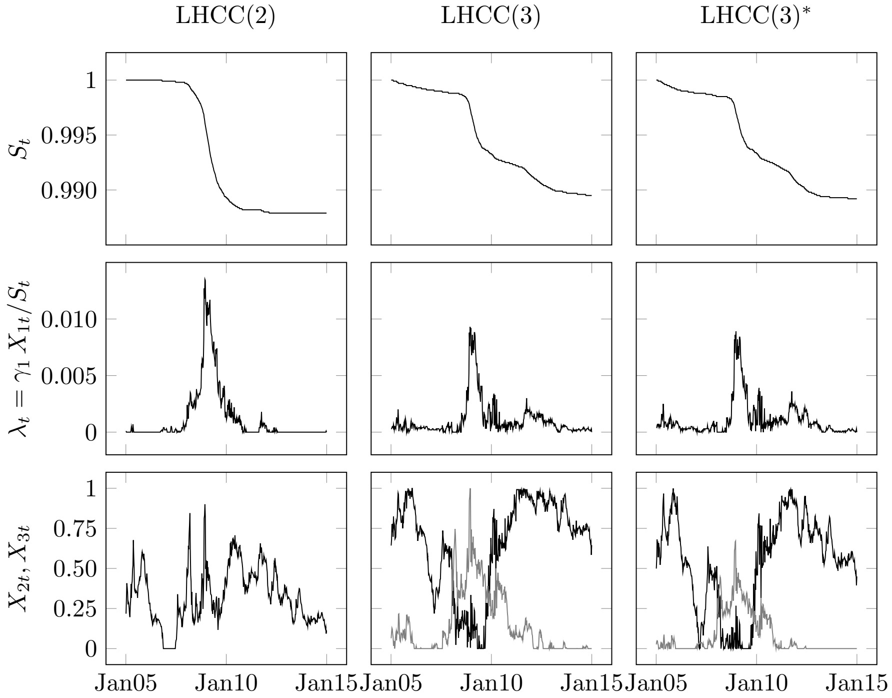


(b) Walt Disney Co.

Figure 3.4 – CDS spreads fits and errors.
The first row displays the fitted CDS spreads in basis points with maturities 1 year (black), 5 years (grey), and 10 years (light-grey) for the three specifications. The second row displays the root-mean-square error (in basis points) computed every day and aggregated over all the maturities.



(a) Bombardier Inc.



(b) Walt Disney Co.

Figure 3.5 – Factors fitted from CDS spreads.

The filtered factors of the three estimated specifications are displayed over time. The first row displays the drift only survival process, the second row the implied default intensity, and the last row the process X_{mt} in black and the process X_{2t} in grey for the three-factor models.

3.5.2 CDS Option

We describe an accurate and efficient methodology to price CDS options that builds on the payoff approximation approach presented in Section 3.3.2, and illustrate it with numerical examples. The model used for the numerical illustration is the one-factor LHC model from Section 3.3.1 with stylized but still realistic parameters $\gamma = 0.25$, $\ell_1 = 0.05$, $\ell_2 = 1$, $\sigma = 0.75$, $X_0 = 0.2$, and $r = 0$ unless otherwise stated.

From Section 3.2.4, we know that the time-t CDS option price with strike spread k is of the form

$$V_{\mathrm{CDSO}}(t, T_0, T_M, k) = \mathbb{1}_{\{\tau>t\}}\mathbb{E}\left[f(Z(T_0, T_M, k)) \mid \mathscr{F}_t\right]$$

with the payoff function $f(z) = \mathrm{e}^{-r(T_0-t)}z^+/S_t$ and where the random variable $Z(T_0, T_M, k)$ is defined by

$$Z(T_0, T_M, k) = \psi_{\mathrm{cds}}(T_0, T_0, T_M, k)^\top \begin{pmatrix} S_{T_0} \\ X_{T_0} \end{pmatrix} \tag{3.44}$$

with $\psi_{\mathrm{cds}}(T_0, T_0, T_M, k)$ as in (3.15). Furthermore, the random variable $Z(T_0, T_M, k)$ takes values in the interval $[a, b]$ with the LHC model which is given by

$$a = \sum_{i=1}^{m+1} \min(0, \psi_{\mathrm{cds}}(T_0, T_0, T_M, k)_i) \quad \text{and} \quad b = \sum_{i=1}^{m+1} \max(0, \psi_{\mathrm{cds}}(T_0, T_0, T_M, k)_i).$$

We now show how to approximate the payoff function f with a polynomial by truncating its Fourier-Legendre series, and then how the conditional moments of $Z(T_0, T_M, k)$ can be computed recursively from the conditional moments of (S_t, X_t).

Let $\mathscr{L}e_n(x)$ denotes the generalized Legendre polynomials taking values on the closed interval $[a, b]$ and given by

$$\mathscr{L}e_n(x) = \sqrt{\frac{1+2n}{2\sigma^2}}\, Le_n\left(\frac{x-\mu}{\sigma}\right)$$

where $\mu = (a + b)/2$, $\sigma = (b - a)/2$, and the standard Legendre polynomial $Le_n(x)$ on $[-1, 1]$ are defined recursively by

$$Le_{n+1}(x) = \frac{2n+1}{n+1} x\, Le_n(x) - \frac{n}{n+1} Le_{n-1}(x)$$

with $Le_0 = 1$ and $Le_1(x) = x$. The generalized Legendre polynomials form a complete orthonormal system on $[a, b]$ in the sense that the mean squared error of the Fourier-Legendre series approximation $f^{(n)}(x)$ of any piecewise continuous function $f(x)$, defined by

$$f^{(n)}(x) = \sum_{k=0}^{n} f_n \mathscr{L}e_n(x), \quad \text{where } f_n = \int_a^b f(x)\mathscr{L}e_n(x)\,dx, \tag{3.45}$$

converges to zero,

$$\lim_{n\to\infty}\int_a^b \left(f(x)-f^{(n)}(x)\right)^2\,dx=0.$$

The coefficients for the CDS option payoff are closed-form,

$$f_n=\mathbb{1}_{\{\tau>t\}}\frac{\mathrm{e}^{-r(T_0-t)}}{S_t}\int_0^b z\mathscr{L}e_n(z)\,dz,$$

since the integrands are polynomial functions. Note that a similar approach is followed in Chapter 1 on the unbounded interval $\mathbb{R}$ with a Gaussian weight function.

The $\mathscr{F}_t$-conditional moments of $Z(T_0,T_M,k)$ can be computed recursively from the conditional moments of (S_{T_0},X_{T_0}). Let $\pi:\mathscr{E}\mapsto\{1,\dots,N_n\}$ be an enumeration of the set of exponents with total order less or equal to n, $\mathscr{E}=\left\{\boldsymbol{\alpha}\in\mathbb{N}^{1+m}:\sum_{i=1}^{1+m}\alpha_i\le n\right\}$. Denote the polynomials $h_{\pi(\boldsymbol{\alpha})}(s,x)=s^{\alpha_1}\prod_{i=1}^m x_i^{\alpha_{1+i}}$ which form a basis of $\mathrm{Pol}_n(E)$, $\mathbf{1}$ the $(1+m)$-dimensional vector of ones, and e_i the $(1+m)$-dimensional vector whose i-th coordinate is equal to one and zero otherwise.

Lemma 3.5.3. *For all $n\ge 2$ we have*

$$\mathbb{E}\left[Z(T_0,T_M,k)^n\mid\mathscr{F}_t\right]=\sum_{\boldsymbol{\alpha}^\top\mathbf{1}=n}c_{\pi(\boldsymbol{\alpha})}\,\mathbb{E}\left[h_{\pi(\boldsymbol{\alpha})}(S_{T_0},X_{T_0})\mid\mathscr{F}_t\right]$$

where the coefficients $c_{\pi(\boldsymbol{\alpha})}$ are recursively given by

$$c_{\pi(\boldsymbol{\alpha})}=\sum_{i=1}^{1+m}\mathbb{1}_{\{\alpha_i-1\ge0\}}\,c_{\pi(\boldsymbol{\alpha}-\mathrm{e}_i)}\,\psi_{\mathrm{cds}}(T_0,T_0,T_M,k)_i.$$

We now report the main numerical findings. We take $T_0=1$, $T_M=T_0+5$, and three reference strike spreads $k\in\{250,300,350\}$ basis points meant to represent in, at, and out of the money CDS options. The first row in Figure 3.6 shows the payoff approximation $f^{(n)}(z)$ in (3.45) for the polynomial orders $N\in\{1,5,30\}$ and the strike spreads $k\in\{250,300,350\}$. A more accurate approximation of the hockey stick payoff function is naturally obtained by increasing the order N, especially around the kink. The width of the support $[a,b]$ increases with the strike spread k, hence the uniform error bound should be expected to be larger for out of the money options. This is confirmed by the second row of Figure 3.6 that shows the error bound as a function of the approximation order N. Note that this error bound is typically non tight. The third row of Figure 3.6 shows the price approximation as a function of the polynomial order, up to $N=30$. The price approximations stabilize rapidly such that a price approximation using the first $N=10$ moments appear to be accurate up to a basis point.

We recall that the volatility parameter σ of the LHC model does not affect the CDS spreads, and can therefore be used to improve the joint calibration of CDS and CDS options. We illustrate this in left panel of Figure 3.7 where the CDS option price is displayed as a function of the

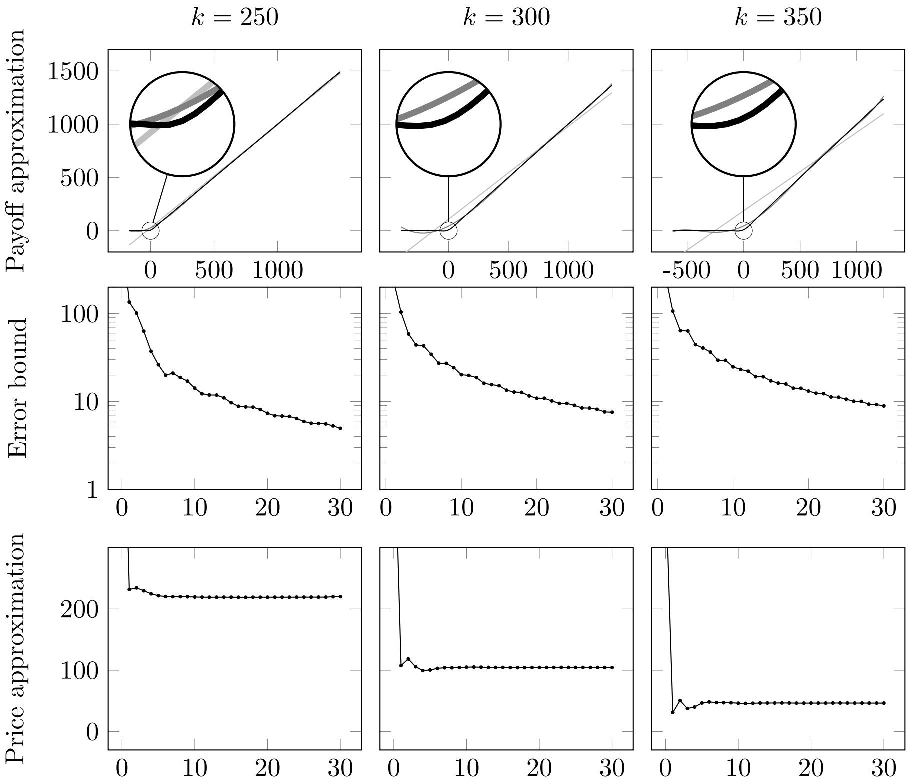


Figure 3.6 – Payoffs and prices approximations of CDS options.

The first row displays the polynomial interpolation of the payoff function approximation with the Fourier-Legendre approach at the order 1 (light-grey), 5 (grey), and 30 (black). The second and third rows display the price error bound and price approximation, respectively, as functions of the polynomial interpolation order. The first (second and third) column corresponds to a CDS option with a strike spread of 250 (300 and 350) basis points. All values are reported in basis points.

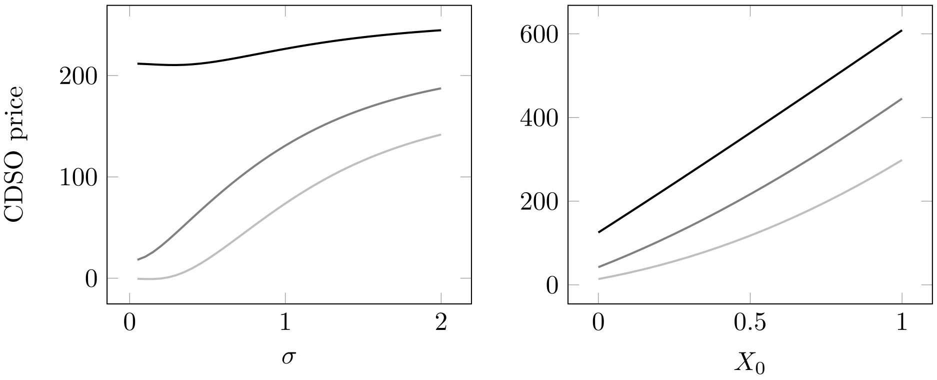


Figure 3.7 – CDS option price sensitivities.
The figure on the left (on the right) display the CDS option price as a function of the volatility parameter (the initial risk factor position) for the strike spread 250 (black), 300 (grey), and 350 (light-grey). All values are reported in basis points.

volatility parameter for different strike spreads. As expected, the option price is an increasing function of the volatility parameter. The right panel on same figure also shows that the X_0 has an almost linear impact on the CDS option price.

Note that the dimension of the polynomial basis $\binom{1+m+N}{N}$ becomes a programming and computational challenge when both the expansion order N and the number of factors $1+m$ are large. For example, for $N = 20$ and $1+m = 2$ the basis has dimension 231 whereas it has dimension 10'626 when $1+m = 4$. In practice, we successfully implemented examples with $1+m = 4$ and $N = 50$ on a standard desktop computer, in which case the basis dimension is 316'251.

Remark 3.5.4. *The CDS option payoff function can also be interpolated by means of Chebyshev polynomials. Details for this approach are given in Section 3.6. However, as discussed in the Section 3.5.3, the price approximation will typically exhibit more pronounced oscillations.*

3.5.3 CDIS Option

We discuss the approximation of the payoff function by means of Chebyshev polynomials for a CDIS option on a homogeneous portfolio. The reference parameters are the same as in Section 3.5.2.

Let $N_t = \sum_{i=0}^{N} \mathbb{1}_{\{\tau_i \le t\}}$ denote the number of firms which have defaulted by time t. Consider a CDIS option on an homogeneous portfolio such that $S_t^i = S_t$ for all $i = 1, \ldots, N$ for some reference survival process S_t. From Proposition 3.2.9 it follows that the time-t price of the

CDIS option is given by

$$
\begin{aligned}
V_{\mathrm{CDISO}}(t, T_0, T_M, k) = {} & \frac{\mathrm{e}^{-r(T_0-t)}}{N S_t^{N-N_t}} \sum_{n=0}^{N-N_t} \binom{N-N_t}{n} \\
& \mathbb{E}\left[\left(\frac{n}{S_{T_0}} \psi_{\mathrm{cds}}(T_0, T_0, T_M, k)^\top \begin{pmatrix} S_{T_0} \\ X_{T_0} \end{pmatrix} + (1-\delta)(N-n) \right)^+ (S_{T_0})^n (S_t - S_{T_0})^{N-N_t-n} \mid \mathscr{F}_t \right]
\end{aligned}
$$

with the notable difference that the summation contains at most $N+1$ terms because the defaults are symmetric and thus interchangeable. Define the random variable

$$
X(T_0, T_M, k) = \sum_{i=1}^{m} \psi_{\mathrm{cds}}(T_0, T_0, T_M, k)_{1+i} \, X_{i,T_0}
$$

the CDIS option price then rewrites

$$
V_{\mathrm{CDISO}}(t, T_0, T_M, k) = \mathbb{E}\left[f(S_{T_0}, X(T_0, T_M, k)) \mid \mathscr{F}_t \vee N_t \right]
$$

where the bivariate payoff function $f(s,x)$ is given by

$$
\begin{aligned}
f(s,x) = {} & \frac{\mathrm{e}^{-r(T_0-t)}}{N S_t^{N-N_t}} \Bigg[(1-\delta) N (S_t - s)^{N-N_t} \\
& + \sum_{n=1}^{N-N_t} \binom{N-j}{n} \Big(n \psi_{\mathrm{cds}}(T_0, T_0, T_M, k)_1 s + x + s(1-\delta)(N-n) \Big)^+ s^{n-1} (S_t - s)^{N-N_t-n} \Bigg].
\end{aligned}
$$

The $\mathscr{F}_t$-conditional moments of $(S_{T_0}, X(T_0, T_M, k))$ are also linear in the same order moments of (S_{T_0}, X_{T_0}) and where the coefficients can be computed recursively in a similar way as in Section 3.5.2.

The CDIS payoff function is displayed on the left graph in Figure 3.8 for the strike spread $k = 350$ bps and $N = 125$ firms with no initially observed defaults. Note that, since $m = 1$, we display directly the factor value X_{T_0} on the horizontal axis. A surprising feature of the payoff function is that it looks relatively smooth, especially in comparison to the hockey stick payoff function of the single-name CDS option. It consists of two mostly linear parts which appear to be curving toward each other at their intersection, in particular near the S_{T_0} axis.

We interpolate the payoff function using Chebyshev polynomials and nodes, the details are given in Section 3.6. The corresponding pricing error upper bound is reported on the right graph Figure 3.8, on which we see it decreasing overall with some oscillations. This is because this approach is a polynomial interpolation of the payoff function at the Chebyshev nodes which are reset for each N but independently from the payoff function. This is in contrario to the Fourier-Legendre approximation which aims to minimize deviation from the payoff function uniformly over the support. As a consequence, price approximation using Chebyshev interpolation of the payoff function will typically oscillate as well.

Remark 3.5.5. *The payoff function $f(s,x)$ could be approximated by truncating at a finite*

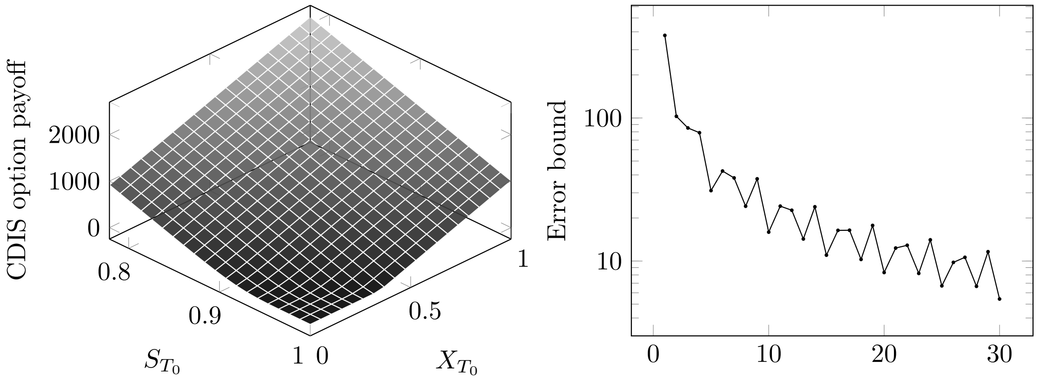


Figure 3.8 – CDS index option payoff function and Chebyshev error bound.
The figure on the left displays the $(\mathscr{F}_{T_0} \vee \mathscr{G}_t)$-conditional CDIS option payoff as a function of the time-T_0 survival process and factor values. The figure on the right reports the corresponding price approximation error bound as a function of the Chebyshev interpolation order. The payoff and error values are reported in basis points.

order n its two-dimensional Fourier-Legendre series representation as in Section 3.5.2. The series coefficients are again closed-form since they boil down to sums of integrated polynomials over compact supports. We used at different method in the example in order to illustrate that different approaches can result in different price series properties. The choice of Chebyshev interpolation is motivated by its popularity which partially originates from its straightforward implementation.

3.6 Chebyshev Interpolation

This section describes how to perform a Chebyshev interpolation of an arbitrary function on an interval $[a,b]\times[c,d]\subset\mathbb{R}$. The Chebyshev polynomials of the first kind take values in $[-1,1]$ but can be shifted and scaled so as to form a basis of $[a,b]$. In this case they are given by the following recursion formula,

$$\begin{aligned}
T_0^{a,b}(x) &= 1\\
T_1^{a,b}(x) &= \frac{x-\mu}{\sigma}\\
T_{n+1}^{a,b}(x) &= \frac{2(x-\mu)}{\sigma}T_n^{a,b}(x) - T_{n-1}^{a,b}(x)
\end{aligned}$$

with $\mu=(a+b)/2$ and $\sigma=(b-a)/2$. The Chebyshev nodes for the interval $[a,b]$ are then given by

$$x_j^{a,b} = \mu + \sigma\cos\left(z_j\right), \quad z_j = \frac{(1/2+j)\pi}{N+1}, \quad \text{for } j=0,\dots,N.$$

The polynomial interpolation of order N is

$$p_N(s,x) = \sum_{n=0}^{N}\sum_{m=0}^{N} c_{n,m}\, T_n^{a,b}(s)\, T_m^{c,d}(x)$$

where the coefficients are explicit

$$c_{n,m} = 2^{\mathbb{1}_{\{i\neq 0\}}+\mathbb{1}_{\{j\neq 0\}}} \sum_{i=0}^{N}\sum_{j=0}^{N} \frac{f\big(x_i^{a,b}, x_j^{c,d}\big)\cos(n\,z_i)\cos(m\,z_j)}{(N+1)^2}.$$

The coefficients can be computed in an effective way by applying Clenshaw's method, or by applying discrete cosine transform. This straightforward interpolation has the advantage to prevent the Runge's phenomena. We refer to (Gaß, Glau, Mahlstedt, and Mair 2015) for more details on the multidimensional Chebyshev interpolation, and for an interesting financial application of multivariate function interpolation in the context of fast model estimation or calibration.

3.7 Market Price of Risk Specifications

We discuss market price of risk (MPR) specifications such that X_t has a linear drift also under the real-world measure $\mathbb{P}\sim\mathbb{Q}$. This may further facilitate the empirical estimation of the LHC model.

Let $\Lambda(S_t,X_t)$ denote the MPR such that the drift of X_t under $\mathbb{P}$ becomes

$$\mu_t^{\mathbb{P}} = bS_t + \beta X_t + \Sigma(S_t,X_t)\Lambda(S_t,X_t).$$

It is linear in (S_t,X_t) of the form

$$\mu_t^{\mathbb{P}} = b^{\mathbb{P}} S_t + \beta^{\mathbb{P}} X_t,$$

for some vector $b^{\mathbb{P}}\in\mathbb{R}^m$ and matrix $\beta^{\mathbb{P}}\in\mathbb{R}^{m\times m}$, if and only if

$$\Lambda_i(s,x) = \frac{\big((b^{\mathbb{P}}-b)s + (\beta^{\mathbb{P}}-\beta)x\big)_i}{\sigma_i\sqrt{x_i(s-x_i)}}, \quad i=1,\dots,m. \tag{3.46}$$

In order that $\Lambda(S_t,X_t)$ is well defined and induces an equivalent measure change, that is, the candidate Radon–Nikodym density process

$$\exp\left(\int_0^t \Lambda(S_u,X_u)\,dW_u - \frac{1}{2}\int_0^t \|\Lambda(S_u,X_u)\|^2\,du\right) \tag{3.47}$$

is a uniformly integrable $\mathbb{Q}$-martingale, we need that (S_t,X_t) does not attain all parts of the boundary of E. This is clarified by the following theorem, which follows from (Cheridito, Filipović, and Yor 2005).

Theorem 3.7.1. *The MPR* $\Lambda(S_t, X_t)$ *in* (3.46) *is well defined and induces an equivalent measure* $\mathbb{P} \sim \mathbb{Q}$ *with Radon-Nikodym density process* (3.47) *if, for all* $i = 1, \dots, m$, $X_{i0} \in (0, S_0)$ *and* (3.24)–(3.25) *hold for the* $\mathbb{Q}$*-drift parameters* β, b *and for the* $\mathbb{P}$*-drift parameters* $\beta^{\mathbb{P}}, b^{\mathbb{P}}$ *in lieu of* β, b.

If, for some $i = 1, \dots, m$, $\beta^{\mathbb{P}}_{ij} = \beta_{ij}$ *for all* $j \neq i$ *and*

1. $b^{\mathbb{P}}_i = b_i$, *such that*

   $$\Lambda_i(s, x) = \frac{(\beta^{\mathbb{P}}_{ii} - \beta_{ii})\sqrt{x_i}}{\sigma_i \sqrt{s - x_i}},$$

   then it is enough if $X_{i0} \in [0, S_0)$ *instead of* $X_{i0} \in (0, S_0)$ *and* (3.22) *instead of* (3.24) *holds for* β_{ij}, b_i, *and thus for* $\beta^{\mathbb{P}}_{ij}, b^{\mathbb{P}}_i$.

2. $b^{\mathbb{P}}_i - b_i = \beta^{\mathbb{P}}_{ii} - \beta_{ii}$, *such that*

   $$\Lambda_i(s, x) = \frac{(\beta^{\mathbb{P}}_{ii} - \beta_{ii})\sqrt{s - x_i}}{\sigma_i \sqrt{x_i}},$$

   then it is enough if $X_{i0} \in (0, S_0]$ *instead of* $X_{i0} \in (0, S_0)$ *and* (3.23) *instead of* (3.25) *holds for* β_{ij}, b_i, *and thus for* $\beta^{\mathbb{P}}_{ij}, b^{\mathbb{P}}_i$.

The assumption of linear-drift preserving change of measure is often made for parsimony and to facilitate the empirical estimation procedure. For example, the specification of MPRs that preserve the affine nature of risk-factors has been theoretically and empirically investigated in (Duffee 2002), (Duarte 2004), and (Cheridito, Filipović, and Kimmel 2007) among others.

3.8 Proofs

This Section contains the proofs of all theorems and propositions in the main text.

Proof of (3.4)

This follows as in (Filipović, Larsson, and Trolle 2017, Lemma 3).

Proof of Example 3.2.3

The autonomous process X_t admits a solution taking values in $[-e^{-\epsilon t}, e^{-\epsilon t}]$ at time t with $\epsilon > 0$ and $X_0 \in [-1, 1]$ if and only if $\kappa > \epsilon$, see (Filipović and Larsson 2016, Theorem 5.1). The coordinates of Y_t are lower bounded by X_t, indeed for $i = 1, 2$ we have

$$dY_{it} = -\frac{\epsilon}{2}(Y_{it} \pm X_t)dt \geq -\frac{\epsilon}{2}(Y_{it} + \mathrm{e}^{-\epsilon t})dt$$

The solution of $dZ_t = -(\epsilon/2)(Z_t + \mathrm{e}^{-\epsilon t})dt$ with $Z_0 = 1$ is given by $Z_t = \mathrm{e}^{-\epsilon t}$ which proves that $Y_{it} \geq Z_t \geq |X_t|$ for $i = 1, 2$. Finally, by applying Ito's lemma we obtain

$$d\langle \lambda^1, \lambda^2 \rangle_t = -\frac{\epsilon^2}{4} \frac{\sigma^2(\mathrm{e}^{-\epsilon t} - X_t)(\mathrm{e}^{-\epsilon t} + X_t)}{Y_{1t} Y_{2t}} \leq 0, \quad t \geq 0.$$

Proof of Proposition 3.2.4

Proposition 3.2.4 is an immediate consequence of (3.4) and the following lemma.

Lemma 3.8.1. *Let Y be a nonnegative $\mathscr{F}_\infty$-measurable random variable. For any $t \leq T < \infty$,*

$$\mathbb{E}\left[\mathbb{1}_{\{\tau > T\}} Y \mid \mathscr{G}_t\right] = \mathbb{1}_{\{\tau > t\}} \frac{1}{S_t} \mathbb{E}\left[S_T Y \mid \mathscr{F}_t\right].$$

Note that $T < \infty$ is essential unless we assume that $S_\infty = 0$.

Lemma 3.8.1 follows from (Bielecki and Rutkowski 2002, Corollary 5.1.1). For the convenience of the reader we provide here a sketch of its proof. As in (Bielecki and Rutkowski 2002, Lemma 5.1.2) one can show that, for any nonnegative random variable Z, we have

$$\mathbb{E}\left[\mathbb{1}_{\{\tau > t\}} Z \mid \mathscr{H}_t \vee \mathscr{F}_t\right] = \mathbb{1}_{\{\tau > t\}} \frac{1}{S_t} \mathbb{E}\left[\mathbb{1}_{\{\tau > t\}} Z \mid \mathscr{F}_t\right].$$

Setting $Z = \mathbb{1}_{\{\tau > T\}} Y$ we can now derive

$$\begin{aligned} \mathbb{E}\left[\mathbb{1}_{\{\tau > T\}} Y \mid \mathscr{G}_t\right] &= \mathbb{E}\left[\mathbb{1}_{\{\tau > t\}} Y \mathbb{1}_{\{\tau > T\}} \mid \mathscr{G}_t\right] = \mathbb{1}_{\{\tau > t\}} \frac{1}{S_t} \mathbb{E}\left[\mathbb{1}_{\{\tau > T\}} Y \mid \mathscr{F}_t\right] \\ &= \mathbb{1}_{\{\tau > t\}} \frac{1}{S_t} \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\{\tau > T\}} \mid \mathscr{F}_\infty\right] Y \mid \mathscr{F}_t\right] = \mathbb{1}_{\{\tau > t\}} \frac{1}{S_t} \mathbb{E}\left[S_T Y \mid \mathscr{F}_t\right]. \end{aligned}$$

Proof of Proposition 3.2.5

The subsequent proofs build on the following lemma that follows from (Bielecki and Rutkowski 2002, Proposition 5.1.1).

Lemma 3.8.2. *Let Z_t be a bounded $\mathscr{F}_t$-predictable process. For any $t \leq T < \infty$,*

$$\mathbb{E}\left[\mathbb{1}_{\{t < \tau \leq T\}} Z_\tau \mid \mathscr{G}_t\right] = \mathbb{1}_{\{t < \tau\}} \frac{1}{S_t} \mathbb{E}\left[\int_{(t,T]} -Z_u dS_u \mid \mathscr{F}_t\right].$$

Note that $T < \infty$ is essential unless we assume that $S_\infty = 0$.

We can now proceed to the proof of Proposition 3.2.5. The value of the contingent cash flow is given by the expression

$$C_{\mathrm{D}}(t, T) = \mathbb{E}\left[\mathrm{e}^{-r(\tau - t)} \mathbb{1}_{\{t \leq \tau \leq T\}} \mid \mathscr{G}_t\right]$$

By applying Lemma 3.8.2 we get

$$\begin{aligned}
C_{\mathrm{D}}(t,T) &= \frac{\mathbb{1}_{\{\tau>t\}}}{S_t}\mathbb{E}\left[\int_t^T -\mathrm{e}^{-r(s-t)}dS_s \mid \mathscr{F}_t\right] = \frac{\mathbb{1}_{\{\tau>t\}}}{S_t}\int_t^T \mathrm{e}^{-r(s-t)}\mathbb{E}\left[-a^\top(cY_s+\gamma X_s)\mid \mathscr{F}_t\right]ds \\
&= \frac{\mathbb{1}_{\{\tau>t\}}}{S_t}\int_t^T \mathrm{e}^{-r(s-t)} - a^\top \begin{pmatrix} c & \gamma \end{pmatrix} \mathrm{e}^{A(s-t)}\begin{pmatrix} Y_t \\ X_t \end{pmatrix} ds
\end{aligned}$$

where the second equality comes from the fact that $\int_0^t \mathrm{e}^{-ru}\, dM_u^S$ is a martingale. The third equality follows from (3.4).

Proof of Corollary 3.2.6

The value of this contingent bond is given by

$$C_{\mathrm{D}_*}(t,T) = \mathbb{E}\left[\tau\,\mathrm{e}^{-r(\tau-t)}\mathbb{1}_{\{t<\tau\le T\}}\mid \mathscr{G}_t\right] = \frac{\mathbb{1}_{\{\tau>t\}}}{S_t}\mathbb{E}\left[\int_t^T -s\,\mathrm{e}^{-r(s-t)}dS_s \mid \mathscr{F}_t\right]$$

and the result follows as in the proof of Proposition 3.2.5.

Proof of Lemma 3.2.7

Observe that for any matrix A and real r we have $\mathrm{e}^r\mathrm{e}^A = \mathrm{e}^{\mathrm{diag}(r)+A}$, and that the matrix exponential integration can be computed explicitly as follows

$$\begin{aligned}
\int_0^u \mathrm{e}^{As}ds &= \int_0^u (I + As + A^2\frac{s^2}{2} + \ldots)ds = Iu + A\frac{u^2}{2} + A^2\frac{u^3}{6} + \ldots \\
&= A^{-1}\left(\mathrm{e}^{Au} - I\right).
\end{aligned}$$

By change of variable $u = s - t$ we obtain

$$\int_t^T s\mathrm{e}^{A_*(s-t)}ds = \int_0^{T-t} u\mathrm{e}^{A_*u}du + t\int_0^{T-t}\mathrm{e}^{A_*u}du,$$

where the second term on the RHS is given in Lemma 3.2.5. The first term can be derived using integration by parts

$$\int_0^{T-t} u\mathrm{e}^{A_*u}du = (T-t)A_*^{-1}\mathrm{e}^{A_*(T-t)} - A_*^{-1}A_*^{-1}\left(\mathrm{e}^{A_*(T-t)} - I\right).$$

Proof of Proposition 3.2.8

The calculations of the protection leg $V_{\mathrm{prot}}^i(t,T_0,T_M)$ and the coupon part $V_{\mathrm{coup}}^i(t,T_0,T_M)$ follows from Propositions 3.2.4 and 3.2.5. The accrued interest $V_{\mathrm{ai}}^i(t,T_0,T_M)$ is given by the sum of contingent cash flows and of weighted zero-recovery coupon bonds, and thus its calculation follows from Propositions 3.2.5 and 3.2.6. The series of contingent cash flow is in fact equal to a single contingent payment paying τ at default,

$$C_{\mathrm{D}_*}(t,T_M)=\sum_{j=1}^{M}\mathbb{E}\left[\tau\,\mathrm{e}^{-r(\tau-t)}\mathbb{1}_{\{T_{j-1}<\tau\le T_j\}}\mid\mathscr{G}_t\right]=\mathbb{E}\left[\tau\,\mathrm{e}^{-r(\tau-t)}\mathbb{1}_{\{t<\tau\le T_M\}}\mid\mathscr{G}_t\right].$$

Using the identity $\mathbb{1}_{\{T_{j-1}<\tau\le T_j\}}=\mathbb{1}_{\{\tau>T_{j-1}\}}-\mathbb{1}_{\{\tau>T_j\}}$ we obtain that the second term of $V^i_{\mathrm{ai}}(t,T_0,T_M)$ is given by

$$\begin{aligned}-\mathbb{E}\left[\sum_{j=1}^{M}\mathrm{e}^{-r(\tau-t)}T_{j-1}\mathbb{1}_{\{T_{j-1}<\tau\le T_j\}}\mid\mathscr{G}_t\right]&=\sum_{j=1}^{M}T_{j-1}\left(C_{\mathrm{D}}(t,T_j)-C_{\mathrm{D}}(t,T_{j-1})\right)\\&=T_{M-1}C_{\mathrm{D}}(t,T_M)-T_0C_{\mathrm{D}}(t,T_0)-\sum_{j=1}^{M-1}(T_j-T_{j-1})C_{\mathrm{D}}(t,T_j).\end{aligned}$$

Proof of Proposition 3.2.9

The payoff at time T_0 of the CDIS option can always be decomposed into 2^N terms by conditioning on all the possible default events

$$q(\alpha)=\prod_{i=1}^{N}\left[(\mathbb{1}_{\{\tau_i>T_0\}})^{\alpha_i}+(\mathbb{1}_{\{\tau_i\le T_0\}})^{1-\alpha_i}\right]\tag{3.48}$$

for $\alpha\in\mathscr{C}=\{0,1\}^N$, and with the convention $0^0=0$, so that the payoff function rewrites

$$\begin{aligned}&\left(\sum_{i=1}^{N}\frac{\mathbb{1}_{\{\tau_i>T_0\}}}{S^i_{T_0}}\psi^i_{\mathrm{cds}}(T_0,T_0,T_M,k)^\top\begin{pmatrix}Y_{T_0}\\X_{T_0}\end{pmatrix}+(1-\delta)\mathbb{1}_{\{\tau_i\le T_0\}}\right)^+=\\&\qquad\sum_{\alpha\in\mathscr{C}}\left(\sum_{i=1}^{N}\frac{\alpha_i}{S^i_{T_0}}\psi^i_{\mathrm{cds}}(T_0,T_0,T_M,k)^\top\begin{pmatrix}Y_{T_0}\\X_{T_0}\end{pmatrix}+(1-\delta)(1-\alpha_i)\right)^+q(\alpha).\end{aligned}$$

We can apply (Bielecki and Rutkowski 2002, Lemma 9.1.3) to compute the probability

$$\mathbb{E}\left[\mathbb{1}_{\{\tau_1>T_0,\dots,\tau_N>T_0\}}\mid\mathscr{F}_{T_0}\vee\mathscr{G}_t\right]=\prod_{i=1}^{N}\mathbb{1}_{\{\tau_i>t\}}\frac{S^i_{T_0}}{S^i_t}$$

so that by writing (3.48) as a linear combination of indicator functions we obtain

$$\mathbb{E}\left[q(\alpha)\mid\mathscr{F}_{T_0}\vee\mathscr{G}_t\right]=\prod_{i=1}^{N}\left(\frac{(S^i_{T_0})^{\alpha_i}(S^i_t-S^i_{T_0})^{1-\alpha_i}}{S^i_t}\mathbb{1}_{\{\tau_i>t\}}+(\mathbb{1}_{\{\tau_i\le t\}})^{1-\alpha_i}\right)$$

which completes the proof.

Proof of Theorem 3.3.1

We define the bounded continuous map $(\mathscr{S},\mathscr{X}) : R^{1+m} \to R^{1+m}$ by

$$\mathscr{S}(s,x) = s^+ \wedge 1, \quad \mathscr{X}_i(s,x) = x_i^+ \wedge s^+ \wedge 1, \quad i = 1,\dots,m,$$

such that $(\mathscr{S},\mathscr{X})(s,x) = (s,x)$ on E. In a similar vein, extend the dispersion matrix $\Sigma(s,x)$ to a bounded continuous mapping $\Sigma((\mathscr{S},\mathscr{X})(s,x))$ on $\mathbb{R}^{1+m}$. The stochastic differential equation (3.19) then extends to $\mathbb{R}^{1+m}$ by

$$\begin{aligned} dS_t &= -\gamma^\top \mathscr{X}(S_t,X_t)\,dt \\ dX_t &= \left(b\mathscr{S}(S_t) + \beta\mathscr{X}(S_t,X_t)\right)dt + \Sigma\left((\mathscr{S},\mathscr{X})(S_t,X_t)\right)dW_t. \end{aligned} \tag{3.49}$$

Since drift and dispersion of (3.49) are bounded and continuous on $\mathbb{R}^{1+m}$, there exists a weak solution (S_t,X_t) of (3.49) for any initial law of (S_0,X_0) with support in E, see (Karatzas and Shreve 1991, Theorem V.4.22).

We now show that any weak solution (S_t,X_t) of (3.49) with $(S_0,X_0)\in E$ stays in E,

$$(S_t,X_t) \in E \text{ for all } t \geq 0. \tag{3.50}$$

To this end, for $i = 1,\dots,m$, note that

$$\Sigma_{ii}\left((\mathscr{S},\mathscr{X})(s,x)\right) = 0 \text{ for all } (s,x) \text{ with } x_i \leq 0 \text{ or } x_i \geq s. \tag{3.51}$$

Conditon (3.22) implies that

$$\left(b\mathscr{S}(s) + \beta\mathscr{X}(s,x)\right)_i \geq 0 \text{ for all } (s,x) \text{ with } x_i \leq 0. \tag{3.52}$$

For $\delta,\epsilon > 0$ we define

$$\tau_{\delta,\epsilon} = \inf\{t \geq 0 \mid X_{it} \leq -\epsilon \text{ and } -\epsilon < X_{is} < 0 \text{ for all } s \in [t-\delta,t)\}.$$

Then on $\{\tau_{\delta,\epsilon} < \infty\}$ we have, in view of (3.51) and (3.52),

$$0 > X_{i\tau_{\delta,\epsilon}} - X_{i\tau_{\delta,\epsilon}-\delta} = \int_{\tau_{\delta,\epsilon}-\delta}^{\tau_{\delta,\epsilon}} \left(b\mathscr{S}(S_u) + \beta\mathscr{X}(S_u,X_u)\right)_i du \geq 0,$$

which is absurd. Hence $\tau_{\delta,\epsilon} = \infty$ a.s. and therefore $X_{it} \geq 0$ for all $t \geq 0$. Similarly, conditon (3.23) implies that

$$-\gamma^\top \mathscr{X}(s,x) - \left(b\mathscr{S}(s) + \beta\mathscr{X}(s,x)\right)_i \geq 0 \text{ for all } (s,x) \text{ with } x_i \geq s. \tag{3.53}$$

Using the same argument as above for $S_t - X_{it}$ in lieu of X_{it}, and (3.53) in lieu of (3.52), we see that $S_t - X_{it} \geq 0$ for all $t \geq 0$. Finally, note that $0 \leq \gamma^\top \mathscr{X}(s,x) \leq \gamma^\top \mathbf{1} s^+$ for all (s,x), and thus $1 \geq S_t \geq \mathrm{e}^{-\gamma^\top \mathbf{1} t} > 0$ for all $t \geq 0$. This proves (3.50) and thus the existence of an E-valued solution of (3.19).

Uniqueness in law of the E-valued solution (S_t, X_t) of (3.19) follows from (Filipović and Larsson 2016, Theorem 4.2) and the fact that E is relatively compact.

The boundary non-attainment conditions (3.24)–(3.25) follow from (Filipović and Larsson 2016, Theorem 5.7(i) and (ii)) for the polynomials $p(s,x) = x_i$ and $s - x_i$, for $i = 1,\dots,m$.

Proof of Remark 3.3.2

The claim in Remark 3.3.2 follows from the following lemma.

Lemma 3.8.3. *The process* (S_t, X_t) *with drift as in* (3.19) *and dispersion matrix given by*

$$\operatorname{diag}\left(\sigma_1\sqrt{x_1(L_1 s - x_1)}, \dots, \sigma_m\sqrt{x_m(L_m s - x_m)}\right)$$

for some positive constants $L_1, \dots, L_m$ *is observationally equivalent to the one with* $L_i = 1$ *for all* $i = 1,\dots, m$.

Proof. This directly follows by applying the change of variable $X'_{it} = X_{it}/L_i$ for each $i = 1,\dots, m$. The dynamics of (S_t, X'_t) then rewrites

$$\begin{aligned} dS_t &= -\gamma'^{\top} X'_t \, dt \\ dX'_t &= (b' S_t + \beta' X'_t)\, dt + \Sigma(S_t, X'_t)\, dW_t \end{aligned}$$

with

$$\begin{aligned} \gamma' &= \operatorname{diag}(L_1, \dots, L_m)\, \gamma, \\ b' &= \operatorname{diag}(L_1, \dots, L_m)\, b, \\ \beta' &= \operatorname{diag}(1/L_1, \dots, 1/L_m)\, \beta \operatorname{diag}(L_1, \dots, L_m), \end{aligned}$$

and where the diffusion matrix is given by Equation (3.20). □

Proof of Proposition 3.4.3

The time-t price of the zero-coupon zero-recovery bond is now given by

$$\begin{aligned} B_Z(t,T) &= \mathbb{E}\left[\frac{D_T}{D_t} \mathbb{1}_{\{\tau > T\}} \mid \mathscr{G}_t\right] = \frac{\mathbb{1}_{\{\tau>t\}}}{D_t S_t} \mathbb{E}\left[D_T S_T \mid \mathscr{F}_t\right] = \frac{\mathbb{1}_{\{\tau>t\}}}{(a_r^{\top} Y_t)(a^{\top} Y_t)} \mathbb{E}\left[(a_r^{\top} Y_T)(a^{\top} Y_T) \mid \mathscr{F}_t\right] \\ &= \frac{\mathbb{1}_{\{\tau>t\}}}{a_Z^{\top} \mathscr{Y}_t} \begin{pmatrix} a_Z^{\top} & 0 \end{pmatrix} \mathrm{e}^{\mathscr{A}(T-t)} \begin{pmatrix} \mathscr{Y}_t \\ \mathscr{X}_t \end{pmatrix} \end{aligned}$$

by applying Lemma 3.8.1. Similarly for contingent cash flows by Lemma 3.8.2 we have

$$\begin{aligned}\mathbb{E}\left[\mathrm{e}^{-r(\tau-t)}\mathbb{1}_{\{t\leq\tau\leq T\}}\,|\,\mathscr{G}_t\right] &= \frac{f(\tau)\mathbb{1}_{\{\tau>t\}}}{S_tD_t}\mathbb{E}\left[\int_t^T -f(s)D_s dS_s\,|\,\mathscr{F}_t\right]\\ &= \frac{\mathbb{1}_{\{\tau>t\}}}{(a_r^\top Y_t)(a^\top Y_t)}\int_t^T f(s)\mathbb{E}\left[-(a_r^\top Y_s)(cY_s+\gamma X_s)\,|\,\mathscr{F}_t\right]ds\\ &= \frac{\mathbb{1}_{\{\tau>t\}}}{a_Z^\top \mathscr{Y}_t}\int_t^T f(s)\,a_{\mathrm{D}}^\top \mathrm{e}^{\mathscr{A}(s-t)}ds\begin{pmatrix}\mathscr{Y}_t\\ \mathscr{X}_t\end{pmatrix}\end{aligned}$$

with $f(s)$ being equal to s or 1, which completes the proof.

Proof of Equation (3.38)

The matrix $\bar{A}$ in Equation (3.37) rewrites

$$\begin{aligned}\bar{A} &= \int_0^\infty (\mathrm{e}^{At}-\mathrm{Id})\gamma_Z t^{-t}\mathrm{e}^{-\lambda_Z t}dt = \gamma_Z\sum_{k=1}^\infty \frac{A^k}{k!}\int_0^\infty t^{k-1}\mathrm{e}^{-\lambda_Z t}dt = \gamma_Z\sum_{k=1}^\infty\frac{A^k}{k!}\frac{\Gamma(k)}{\lambda_Z^k}\\ &= \gamma_Z\sum_{k=1}^\infty\frac{\left(A\lambda_Z^{-1}\right)^k}{k} = -\gamma_Z\log\left(\mathrm{Id}-A\lambda_Z^{-1}\right)\end{aligned}$$

where the second line follows from the definition of the matrix exponential, the third from the definition of the Gamma function and its explicit values for integers, and the last one from the definition of the matrix logarithm.

Proof of Proposition 3.5.1

The matrix A_* in the LHCC model is given by

$$A_* = \begin{pmatrix} -r & -\gamma_1 & 0 & 0 & \\ 0 & -(\kappa_1+r) & \kappa_1\theta_1 & 0 & \cdots \\ \vdots & & & \ddots & \\ \theta_m & & & 0 & -(\kappa_m+r)\end{pmatrix}$$

and its determinant is therefore equal to

$$|A_*| = -r\begin{vmatrix} -(\kappa_1+r) & \kappa_1\theta_1 & 0 & \cdots \\ \vdots & & \ddots & \\ 0 & & 0 & -(\kappa_m+r)\end{vmatrix} + (-1)^m\begin{vmatrix} -\gamma_1 & 0 & 0 & \\ -(\kappa_1+r) & \kappa_1\theta_1 & 0 & \cdots \\ \vdots & & \ddots & \\ 0 & & -(\kappa_m+r) & \kappa_m\theta_m\end{vmatrix}.$$

With $r > 0$, the first element on the right hand side is nonzero with sign equal to $(-1)^{1+m}$ and the second element also has a sign equal to $(-1)^{1+m}$. This is because the determinant of a triangular matrix is equal to the product of its diagonal elements. As a result, the determinant of A_* is nonzero which concludes the proof.

Proof of Lemma 3.5.3

We n-th power of $Z(T_0,T_M,k)$ rewrites

$$\begin{aligned}
Z(T_0,T_M,k)^n &= \left(\psi_{\mathrm{cds}}(T_0,T_0,T_M,k)^\top \begin{pmatrix} S_{T_0} \\ X_{T_0} \end{pmatrix}\right)^n \\
&= \psi_{\mathrm{cds}}(T_0,T_0,T_M,k)^\top \begin{pmatrix} S_{T_0} \\ X_{T_0} \end{pmatrix} \sum_{\alpha^\top \mathbf{1}=n-1} c_{\pi(\boldsymbol{\alpha})}\, h_{\pi(\boldsymbol{\alpha})}(S_{T_0},X_{T_0}) \\
&= \sum_{i=1}^{1+m} \sum_{\alpha^\top \mathbf{1}=n-1} c_{\pi(\boldsymbol{\alpha})} \psi_{\mathrm{cds}}(T_0,T_0,T_M,k)_i\, h_{\pi(\boldsymbol{\alpha}+\mathrm{e}_i)}(S_{T_0},X_{T_0})
\end{aligned}$$

which is a polynomial containing all and only polynomials of degree n, the lemma follows by rearranging the terms.

3.9 Conclusion

We introduce the class of linear credit risk models in which the background survival probability of a firm and its factors have a linear drift. The prices of defaultable bonds and CDSs become linear-rational in the factors. We define the single-name linear hypercube (LHC) model with diffusive factor process that has quadratic diffusion function and takes values in a compact state space. These features are employed to develop an efficient European option pricing methodology. We build upon the LHC model to construct parsimonious and versatile multi-name models, and to accommodate for stochastic interest rates. We also introduced jumps in the factors dynamics as well as stochastic clocks to generate simultaneous defaults. An empirical analysis shows that the LHC model is able to capture the complex CDS term structure dynamics. The prices of CDS options at different moneyness are accurately approximated.

4 Dependent Defaults and Losses with Factor Copula Models

We present a class of flexible and tractable static factor models for the joint term structure of default probabilities, the factor copula models. These high dimensional models remain parsimonious with pair copula constructions, and nest numerous standard models as special cases. With finitely supported random losses, the loss distributions of credit portfolios and derivatives can be exactly and efficiently computed. Numerical examples on collateral debt obligation (CDO), CDO squared, and credit index swaption illustrate the versatility of our framework. An empirical exercise shows that a simple model specification can fit credit index tranche prices.

4.1 Introduction

This chapter introduces factor copulas to model dependent default times and losses. We directly specify the joint probability of default times, taking as given the marginal default probabilities. Specifically, the default times are assumed to be independent conditional on a latent factor. The joint default probability is given by an explicit expression in terms of conditional copulas. We show that this specification nests all the standard factor models, such as the Gaussian, Archimedean, and stochastic correlation models. In addition, our framework has two main advantages over the existing models. First, the types of dependence between the default times and the latent factor can be highly heterogeneous across entities. Second, new simple and flexible models can be constructed using mixtures and cascades of pair copulas.

We present a new approach to compute efficiently and exactly the loss distribution of credit portfolios and derivatives on these portfolios. In particular, this allows us to compute the exact payoff distribution of credit portfolio derivatives such as portfolio tranche, collateralized debt obligation (CDO) squared, and credit index swaption. Conditional on the latent factor, the realized individual losses are assumed to be independent from each others and from the default times. This enables us to retrieve the exact portfolio loss distribution using discrete Fourier transform methods. This contrasts with existing approaches that either compute the exact loss distribution using slow recursive methods, or compute an approximate loss distribution by discretizing its support and then applying Fourier techniques. We also suggest the Beta-binomial distribution as a flexible mean to specify the loss given default of each entity.

We explore the versatility of our setup and discuss the impact of different dependence hypothesis on the loss distribution with numerical examples. The discrete Fourier method is shown to be significantly faster than the recursive method, especially when the dimension of the latent factor or of the loss support size is large. We construct a simple model for which the total number of defaults distribution exhibits the features of both highly and little dependent defaults, namely a bump and a fat tail. We compute the loss distribution of a CDO tranche and show that the loss distribution of a portfolio of tranches, also known as CDO squared, may have dramatically different profiles depending on the dependence structure between the underlying tranches. We illustrate the flexibility of the Beta-binomial models for individual loss amounts, and show that the specification of individual losses may also critically affect the portfolio loss distribution.

As an application, we use our approach to fit the market tranche prices on the North America investment grade credit index series 21. We start by exploring various model specifications (from standard copulas to multi-factor models), and introduce a mixture with two Gaussian copulas, parametrized with two correlations and a weight balancing each component. In a static analysis, we find that the mixture outperforms the other models, as it is the only one reproducing the prices of both the junior and senior tranches. Fitting this model for all days in our sample, we further find that the parameters are stable over time. Furthermore, one of the correlations being almost always equal to one, we repeat the exercise by fixing it to 0.999. Interestingly, we find similar results, therefore achieving an almost perfect calibration to all tranches using only two parameters (i.e. the other correlation and the weight).

Although motivated by credit risk applications, we present a generic framework to model dependent defaults and losses in high dimensions that may be useful in other areas of survival analysis such as contingent claim pricing in insurance.

We now review some of the related literature. Our approach builds on recent advances on the high-dimensional modeling of random variables. When dealing with multivariate data, copulas are attractive, allowing to model separately the marginal distributions and the dependence structure. Unfortunately, few copulas remain practically useful in high-dimensional settings, because common parametric families are often either too flexible, or not enough. An example of the former is the elliptical family, whose members have a number of parameters that grows quadratically with the dimension. Conversely, members of the archimedean family have a small and fixed number of parameters, independently of the dimension. Recently, high-dimensional copulas using a factor structure have been constructed independently by (Oh and Patton 2013; Oh and Patton 2017) and (Krupskii and Joe 2013; Krupskii and Joe 2015). Such approaches alleviate the curse of dimensionality by considering a smaller set of latent variables, conditional upon which the random variables of interest are assumed independent. Arguably the main difference between the methods presented in (Oh and Patton 2013; Oh and Patton 2017) and (Krupskii and Joe 2013; Krupskii and Joe 2015) is that copulas proposed in the former can only be simulated, whereas those in the latter admit closed form expressions. In fact, it can be shown the factor copulas from (Krupskii and Joe 2013; Krupskii and Joe 2015) are a special case of pair-copula constructions (PCCs). One of the hot topics of multivariate analysis over the last couple of years, PCCs are flexible representations of the dependence structure underlying a multivariate distribution. Introduced by (Bedford and Cooke 2001; Bedford and Cooke 2002) and popularized by (Aas, Czado, Frigessi, and Bakken 2009), PCCs are decompositions of a joint distribution by considering pairs of conditional random variables. For a given joint distribution, such a construction is not unique, but all possible decompositions can be organized as graphical structures, the so-called PCCs. Assuming the copula linking default times as in (Krupskii and Joe 2013; Krupskii and Joe 2015), an interesting aspect of our approach is that it nests the standard models described for instance in (Li 2000; Burtschell, Gregory, and Laurent 2005; Hofert and Scherer 2011) as special cases. Although static by construction, our approach can be incorporated in a dynamic doubly stochastic framework as described in (Schönbucher and Schubert 2001) in order to generate stronger default correlation than in pure intensity based models.

To recover the loss distribution, recursive techniques with proportional loss given default have been studied by (Andersen, Sidenius, and Basu 2003; Hull and White 2004), and Fourier approximations are presented in (Gregory and Laurent 2003; Laurent and Gregory 2005). The computational performance of the latter approach has been improved for models with a large number of Gaussian factors in (Glasserman and Suchintabandid 2012) by using a quadratic approximation technique.

The calibration of tranches on credit portfolios is a daunting task, which is often solved in an ad-hoc way (e.g., by considering a specific model for each tranche). Significant effort have been made to develop consistent models, see (Giesecke 2008) for a comparison between top down and bottom up approaches. Standard copula models generally had limited empirical success and other frameworks have been developed, see (Hull and White 2006), (Brigo, Pallavicini, and Torresetti 2007), (Kalemanova, Schmid, and Werner 2007), (Cousin and Laurent 2008), (Herbertsson 2008), (Fouque, Sircar, and Sølna 2009), (Burtschell, Gregory, and Laurent 2009), (Filipović, Overbeck, and Schmidt 2011). In this chapter, we develop bottom-up models that are both simple to calibrate and successful at reproducing all the tranche spreads. Furthermore, while the valuation of CDO squared has been considered with simulations in (Hull and White 2010),(Guillaume, Jacobs, and Schoutens 2009), this work is the first to derive explicitly the loss density of a CDO squared in a factor copula framework. We refer to (Brigo, Pallavicini, and Torresetti 2010) for a technical analysis of valuation methods for structured credit products.

The realized loss at default on corporate loans and bonds is known to be stochastic, volatile, and negatively correlated with the business cycle. The recovery rates volatility and correlation with default risk is studied, for example, in (Altman, Resti, and Sironi 2004). These important properties and their impact on the valuation of credit derivatives have been investigated by (Andersen and Sidenius 2004),(Krekel 2008), (Amraoui and Hitier 2008).

The remainder of the chapter is structured as follows. Section 4.2 presents the factor copula framework. Section 4.3 describes the construction of the individual loss amounts and the computation of the loss distributions. Section 4.4 contains numerical examples illustrating the performance of our setup and the impact of different dependence hypothesis. The empirical analysis is in Section 4.5. Additional results on standard factor copula models are summarized in Section 4.6. Pricing formulas for tranches, credit index swaps, and credit index swaption are given in Section 4.7. Section 4.8 contains the proofs. Section 4.9 concludes.

4.2 The Factor Copula Framework

We consider N entities. For each $j = 1, \dots, N$ let $p_{j,t}$ be a non-decreasing deterministic function satisfying $p_{j,0} = 0$ and $\lim_{t\to\infty} p_{j,t} = 1$ for all $0 < t < \infty$. We define the default time τ_j of entity j as follows

$$\tau_j := \inf\{t \geq 0 : U_j \leq p_{j,t}\},$$

where U_j is a uniform random variable on the unit interval. Hence, the function $p_{j,t}$ is equivalent to the marginal default probability of entity j

$$\mathbb{P}\left[\tau_j \leq t\right] = \mathbb{P}\left[U_j \leq p_{j,t}\right] = p_{j,t}.$$

When $p_{j,t}$ is absolutely continuous with respect to time, it has the following representation

$$p_{j,t} = 1 - e^{-\int_0^t \lambda_{j,s} ds} \tag{4.1}$$

for some non-negative default intensity function $\lambda_{j,s}$.

Note that, in this setup, the random vector $U = (U_1, \dots, U_N)$ is the only stochastic object. We recall that its probability distribution is by construction a copula.

Definition 4.2.1. *A copula C_U is the probability distribution of a random vector U taking values on the hypercube $[0,1]^N$ and having uniform marginal distributions.*

In other words, if for any vector $(u_1, \dots, u_N) \in [0,1]^N$ the random vector $U \in [0,1]^N$ is such that $\mathbb{P}\left[U_j \leq u_j\right] = u_j$ for each j, then its joint distribution is called a copula and we write

$$C_U(u_1, \dots, u_N) = \mathbb{P}\left[U_1 \leq u_1, \dots, U_N \leq u_N\right]. \tag{4.2}$$

The following lemma shows that for $(t_1, \dots, t_N) \in \mathbb{R}_+^N$, there exists a simple expression linking joint to marginal default probabilities using the copula C_U of U.

Lemma 4.2.2. *The joint default probability is given by*

$$\mathbb{P}\left[\tau_1 \le t_1, \ldots, \tau_N \le t_N\right] = C_U\left(p_{1,t_1}, \ldots, p_{N,t_N}\right). \tag{4.3}$$

A direct construction of high-dimensional copulas amounts at trading-off model complexity and tractability. This is somewhat problematic, because the usual parametric families contain either too many (e.g., in the case of implicit copulas extracted from known multivariate distributions), or too few (e.g., in the case of Archimedean copulas built using a continuous and nonincreasing N-monotone generator) parameters. Furthermore, as we will show in Section 4.3 when pricing complex financial derivatives, the notion of conditional independence (on a set of latent factors) allows us to obtain a flexible yet tractable class of models. Hereinafter we therefore focus on the so-called factor copulas.

4.2.1 One-Factor Copulas

A one-factor copula model is constructed by assuming that there exists a latent factor V such that, conditional on the realization of V, the coordinates of the random vector U are independent. Further assuming that V is uniformly distributed on the unit interval[1], it means that

$$\mathbb{P}\left[U_1 \le u_1, \ldots, U_N \le u_N \mid V = v\right] = \prod_{j=1}^{N} \mathbb{P}\left[U_j \le u_j \mid V = v\right] \tag{4.4}$$

for any vector $(u_1, \ldots, u_N) \in [0,1]^N$ and for any $v \in [0,1]$. The following proposition shows that such an assumption yields a simple decomposition in terms of bivariate copulas for C_U. We refer to such copula as one-factor copula.

Proposition 4.2.3 (One-factor copula)**.** *For $j = 1, \ldots, N$, let $C_{U_j,V}$ denote the joint distribution of U_j and V, that is $\mathbb{P}\left[U_j \le u_j, V \le v\right] = C_{U_j,V}(u_j, v)$. If the coordinates of U are independent conditionally on V, then*

$$C_U(u_1, \ldots, u_N) = \int_{[0,1]} \prod_{j=1}^{N} C_{U_j|V}\left(u_j \mid v\right) dv, \tag{4.5}$$

where, for all $j = 1, \ldots, N$,

$$C_{U_j|V}\left(u_j \mid v\right) = \frac{\partial C_{U_j,V}(u_j, v)}{\partial v}$$

are the so-called h-functions.

The h-functions have been introduced by (Aas, Czado, Frigessi, and Bakken 2009) while studying the pair-copula decomposition of a general multivariate distribution: if $C_{U_j,V}(u_j,v) = \mathbb{P}\left[U_j \leq u_j, V \leq v\right]$, then $C_{U_j|V}\left(u_j \mid v\right) = \mathbb{P}\left[U_j \leq u_j \mid V = v\right]$.

---

[1] This is without loss of generality as the latent factor could be mapped to such a V using the probability integral transform if it was not the case.

Note that $C_{U_j,V}(u,v) = uv$ implies $C_U(u_1,\ldots,u_N) = \prod_{j=1}^N u_j$. In other words, if U_j is independent from V, then it is also independent from U_k for all $k \in \{1,\ldots,j-1,j+1,\ldots,N\}$, which means that the coordinates of U depend on each other only through the factor V.

Example 4.2.4. *The Gaussian model of (Li 2000) is a one-factor copula obtained by using for all j*

$$C_{U_j,V}(u_j,v;\rho) = \Phi_2\left(\Phi^{-1}(u_j),\Phi^{-1}(v);\rho\right),$$

which implies

$$C_{U_j|V}(u_j \mid v;\rho) = \Phi\left(\frac{\Phi^{-1}(u_j)-\rho\Phi^{-1}(v)}{1-\rho^2}\right),$$

and

$$C_U(u_1,\ldots,u_N;\rho) = \int_0^1 \prod_{j=1}^N \Phi\left(\frac{\Phi^{-1}(u_j)-\rho\Phi^{-1}(v)}{1-\rho^2}\right) dv,$$

where $\Phi(\cdot)$ is the standard normal distribution and $\Phi_2(\cdot,\cdot;\rho)$ is the bivariate normal distribution with correlation ρ.

Observe that the specification in Proposition 4.2.3 is far more flexible, since one could build a model using, for each entity, a different bivariate copula, for which countless well-studied parametric families exist, see (Schepsmeier and Stöber 2014).

Beyond such parametric families, a simple way to increase the modeling flexibility while preserving analytical tractability is to combine different bivariate copulas. Through the following definition, mixture distributions enrich considerably the one-factor copulas.

Definition 4.2.5. *Let K be a positive integer, $C_{U_j,V}$ is a mixed bivariate copula if there exists K copulas $C^k_{U_j,V}$, K positive weights $w_k > 0$ such that $\sum_{k=1}^K w_k = 1$, and*

$$C_{U_j,V}(u_j,v) = \sum_{k=1}^K w_k C^k_{U_j,V}(u_j,v). \tag{4.6}$$

One way to interpret this expression is Bayesian, namely assuming that the dependence between the random variable U_j and the factor V is uncertain and follows the distribution $C^k_{U_j,V}$ with probability w_k. The corresponding h-function still has a simple expression, as we have

$$C_{U_j|V}(u_j \mid v) = \sum_{k=1}^K w_k C^k_{U_j|V}(u_j \mid v).$$

Of particular interest for risk management applications, the joint distribution of default times conditional on a subset of realized default times is obtained as a simple modification of Equation (4.5). Let $\mathscr{I} = \{1,\dots,N\}$ and $\mathscr{D} \subset \mathscr{I}$ denote respectively the entire set and a subset of entities. The following proposition shows that the joint default distribution conditional on the defaults of all the entities in $\mathscr{D}$ also has a simple representation.

Proposition 4.2.6. *In a one-factor copula model, the joint default distribution conditional on $\tau_k = t_k$ for $k \in \mathscr{D}$ is*

$$\mathbb{P}\left[\tau_1 \le t_1,\dots,\tau_N \le t_N \mid \tau_k = t_k : k \in \mathscr{D}\right] = \frac{\int_{[0,1]} \prod_{j\in\mathscr{I}\setminus\mathscr{D}} C_{U_j|V}\left(p_{j,t_j} \mid v\right) \prod_{k\in\mathscr{D}} c_{U_k,V}\left(p_{k,t_k}, v\right) dv}{\int_{[0,1]} \prod_{k\in\mathscr{D}} c_{U_k,V}\left(p_{k,t_k}, v\right) dv} \tag{4.7}$$

where

$$c_{U_j,V}(u,v) = \frac{\partial^2 C_{U_j,V}(u,v)}{\partial u \partial v}$$

is the density of the bivariate copula $C_{U_j,V}$.

Although the default times are correlated, conditioning on a subset of defaulted entities does not significantly complexify the expression for the joint distribution of the surviving entities. The denominator on the right hand side in (4.7) is the copula density of the defaulted entities evaluated at the default times. This result may be of particular interest to compute the loss distribution of a credit portfolio conditional on the default time of a specific entity which in turn could be used to compute the Credit Valuation Adjustment with respect to this entity.

4.2.2 Multi-Factor Copulas

In this section we consider a d-dimensional random vector of latent factors $V = (V_1,\dots,V_d)$. We assume that V takes values on the hypercube $[0,1]^d$ and has uniform marginal distributions. The joint distribution of V is by definition a copula that we denote C_V. The following proposition shows that the one-factor framework extends to a multi-factor one.

Proposition 4.2.7 (Multi-factor copula)**.** *For $j = 1,\dots,N$, let $C_{U_j,V}$ denote the joint distribution of U_j and V, that is $\mathbb{P}\left[U_j \le u_j, V \le v\right] = C_{U_j,V}(u_j, v)$. If the coordinates of U are independent conditionally on V, then*

$$C_U(u) = \int_{[0,1]^d} \prod_{j=1}^{N} C_{U_j|V}\left(u_j \mid v\right) dC_V(v) \tag{4.8}$$

where, for all $j = 1,\dots,N$,

$$C_{U_j|V}(u_j \mid v) = \frac{\partial^d C_{U_j,V}(u_j, v)}{\partial v_1 \dots \partial v_d}.$$

Although (4.8) appears to be similar to (4.5), it is arguably more complicated. The reason is that, instead of being bivariate, each $C_{U_j,V}$ has dimension $d+1$. However, the multi-factor framework simplifies under the assumption of independent latent factors V as shown in the following proposition. We denote the function composition with the symbol $\circ$, that is $f(g(x)) = f \circ g(x)$ for any real valued functions f and g.

Corollary 4.2.8 (Copulas with independent factors). *If* $C_V(v) = \prod_{j=1}^d v_j$, *then*

$$C_U(u_1, \dots, u_n) = \int_{[0,1]^d} \prod_{j=1}^N C_{U_j|V_1}(\cdot|v_1) \circ \cdots \circ C_{U_j|V_d}(u_j|v_d)dv, \tag{4.9}$$

where C_{U_j,V_k} *is a bivariate copula for* $j \in 1, \dots, d$ *and* $k = 1, \dots, d$.

Note that the recursive decomposition (4.9) is a particular case of pair-copula constructions (PCCs), as noted in (Krupskii and Joe 2013), which are representations of flexible joint distributions as cascade products of bivariate copulas and marginals. For more details on the subject, we refer to (Bedford and Cooke 2001),(Bedford and Cooke 2002), which proposed a graphical model to help organizing PCCs, or (Aas, Czado, Frigessi, and Bakken 2009), which popularized them by developing efficient computational algorithms for their inference and simulation.

This construction is interesting for several reasons. First, it is a parsimonious way to model a complex multivariate dependencies. Second, the hierarchical structure, which can be represented as a graphical model, has a visual interpretation. Third, because the integrand in (4.9) is a simple recursion, it can be vectorized in a computationally efficient manner.

Finally, it should be noted that the number of latent factors is also the dimension of the hypercube on which the product of conditional copulas has to be integrated to retrieve the joint default probability. One should therefore balance between higher modeling flexibility and lower computational cost.

4.2.3 Comparison with Standard Factor Models

In this section, we show that standard static models can be rewritten explicitly as factor copula models. One usually considers a random vector $Y = (Y_1, \dots, Y_N) \in \mathbb{R}^N$ along with a deterministic and componentwise non-decreasing vector $y_t = (y_{1,t}, \dots, y_{N,t}) \in \mathbb{R}^N$. For instance, Y can represent the values of N firms and y_t the corresponding default barriers[2]. The default time τ_j of firm j is then defined as the first time its value is below its default barrier, that is

$$\tau_j = \inf\{t \geq 0 : Y_j \leq y_{jt}\}.$$

---

[2] While firm values and default barriers are usually positive, this can be resolved by using a monotonic transformation of Y and y_t without affecting the results that follow. For instance, with $\widetilde{Y} = e^Y$ and $\widetilde{y_t} = e^{y_t}$, it is clear that $\mathbb{P}\left[Y_j \leq y_{j,t}\right] = \mathbb{P}\left[\widetilde{Y}_j \leq \widetilde{y}_{j,t}\right]$ and that the copulas of Y and $\widetilde{Y}$ are the same

Additionally, standard factor models are constructed by decomposing the stochastic behavior of the firm value into a systemic and an idiosyncratic component. In other words, one assumes the existence of a random vector $X \in \mathbb{R}^d$ and N variables ϵ_j for $j \in \{1, \ldots, N\}$, such that Y_j is a function X and ϵ_j, that is

$$Y_j = f_j(X, \epsilon_j)$$

for some $(d+1)$-dimensional function f_j taking values on $\mathbb{R}_+$.

Let F_{Y_j}, F_X, respectively $F_{Y_j}^{-1}, F_X^{-1}$, denote the distributions of Y and X, respectively their inverse, and $F_{Y_j|X}$ denote the conditional distribution of Y_j given X. The following proposition shows that any standard factor model is equivalent to a specific factor copula model.

Theorem 4.2.9. *A standard factor model is a factor copula model with marginal default probabilities $p_{j,t} = F_{Y_i}(y_{j,t})$ and conditional copulas*

$$C_{U_j|V}(u \mid v) = F_{Y_j|X}(F_{Y_j}^{-1}(u) \mid (F_{X_1}^{-1}(v_1), \ldots, F_{X_N}^{-1}(v_N)),$$

for $j = 1, \ldots, N$, and where the copula of V is given by

$$C_V(v) = F_X(F_{X_1}^{-1}(v_1), \ldots, F_{X_N}^{-1}(v_N)).$$

Furthermore, if the functions F_X and F_{Y_j} for all $j = 1, \ldots, N$ are continuous, then the copulas C_V and $C_{U_j|V}$ for all $j = 1, \ldots, N$ are unique.

Example 4.2.10. *The Gaussian model described in Example 4.2.4 is obtained by writing, for $j \in \{1, \ldots, N\}$, $Y_j = \rho X + \sqrt{1-\rho^2} Z_j$ and $y_{j,t} = \Phi(p_{j,t})$ where $X, Z_1, \ldots, Z_N$ are i.i.d. $N(0,1)$ random variables.*

In Section 4.6, we derive the factor copula representation of other popular models such as the Stochastic correlation, the t-Student, the Archimedean models, and the Gaussian-Mixture. However, while $C_{U_j|V}$ and C_V sometimes admit such closed-form expressions, it is clear that the marginal distributions are irrelevant. Instead, working directly with copulas offers more modeling flexibility while ensuring tractability.

Having described the construction of the joint distribution of default times, we now turn our attention toward the second element of our framework: the modeling of the losses given default. In the next section, we introduce a class of discrete loss distributions which can be computed in quasi-closed form.

4.3 Discrete Loss Distributions

We define the time-t loss L_t on a portfolio composed of securities written on N different obligors as

$$L_t = \sum_{j=1}^{N} \ell_j \, \mathbb{1}_{\{\tau_j \leq t\}} = \sum_{j=1}^{N} \ell_j \, \mathbb{1}_{\{U_j \leq p_{jt}\}}, \tag{4.10}$$

where ℓ_j is the possibly random loss amount experienced when obligor j defaults, and $\mathbb{1}_{\{\tau_j \leq t\}}$ is the default indicator of obligor j. In this section, we make two assumptions on ℓ_j to preserve the tractability of the portfolio loss distribution, and to enable efficient numerical techniques.

First, we assume that ℓ_j is V-conditionally independent of both ℓ_k for $k \neq j$ and U (or equivalently τ), that is

$$\mathbb{P}\left[U \leq u, \ell \leq x \mid V = v\right] = \prod_{j=1}^{N} C_{U_j|V}(u_j \mid V = v)\mathbb{P}\left[\ell_j \leq l_j \mid V = v\right],$$

with $\ell = (\ell_1, \ldots, \ell_N)$, and for any $u \in [0,1]^N$, $v \in [0,1]^d$ and $l \in \mathbb{R}_+^N$. As in the case of the joint distribution of default times, the V-conditional probabilities can be arbitrarily specified. Hence, the conditional independence property does not preclude some dependence between default rates and loss given default.

Second, as in (Andersen, Sidenius, and Basu 2003),(Andersen and Sidenius 2004),(Hull and White 2004), we assume that the losses are discrete. More specifically, we let $\delta \in \mathbb{R}_+$ be the common loss unit, such that each ℓ_j has a discrete support starting at zero and with mesh δ, that is

$$\ell_j \in \{0, \delta, 2\delta, \ldots, m_j\delta\}, \quad j = 1, \ldots, N$$

for some integer $m_j \in \mathbb{N}$. Hence, the portfolio loss distribution also has a discrete support with the same mesh δ, that is

$$L_t \in \{0, \delta, 2\delta, \ldots, M\delta\}$$

where $M = \sum_{i=1}^{N} m_i$. Although δ is an arbitrary constant, it can be as fine as required in order to mimick the discreteness of real-world prices. For instance, assuming that the granularity of prices is in cents (i.e., $\delta = 0.01\$$) and that the notional of each contract is 1\$, then $m_j = 100$ and $M = N \times 100$.

In the next section, we describe our method to compute the distribution of L_t in quasi-closed form using discrete Fourier inversion.

4.3.1 Portfolio Loss Distribution

In this section, we show that the portfolio loss distribution has an almost closed-form expression that can be efficiently computed numerically. Recall that, for a discrete and finitely supported random variable $X \in \{0, 1, \ldots, M\}$ admiting a characteristic function $\phi_X(u) = \mathbb{E}\left[\mathrm{e}^{\mathrm{i}uX}\right]$, its distribution can be represented as a finite sum

$$\mathbb{P}[X = k] = \frac{1}{M+1} \sum_{m=0}^{M} \phi_X\left(\frac{2\pi m}{M+1}\right) e^{-\frac{2\pi \mathrm{i} k m}{M+1}}$$

Therefore, if the characteristic function of the loss distribution admits a closed-form expression, so does the loss distribution itself. Using the V-conditional independence, the following proposition shows that the characteristic function of the loss admits a simple expression. To improve the clarity of the formulas, we work with the normalized losses

$$\ell_j \delta^{-1} \in \{0, 1, \ldots, m_j\}$$

and normalized portfolio loss

$$L_t \delta^{-1} \in \{0, 1, \ldots, M\}.$$

Proposition 4.3.1. *The characteristic function of the normalized portfolio loss $L_t\delta^{-1}$ is given by*

$$\phi_{L_t}(u) = \mathbb{E}\left[\mathrm{e}^{\mathrm{i}uL_t\delta^{-1}}\right] = \int_{[0,1]^d} \prod_{j=1}^{N} \left(1 - p_{j,t}(v) + p_{j,t}(v)\phi_{\ell_j}(u, v)\right) dC_V(v), \tag{4.11}$$

for any time $t \geq 0$ and for $u \in \mathbb{R}$, where $p_{j,t}(v) = C_{U_j|V}(p_{j,t} \mid v)$ is the conditional default probability of j, $p_{j,t}$ is the unconditional default probability of j defined by Equation 4.1, and

$$\phi_{\ell_j}(u, v) = \sum_{k=0}^{n_j} \mathbb{P}\left[\ell_j = \delta\, k \mid V = v\right] \mathrm{e}^{\mathrm{i}uk} \tag{4.12}$$

the V-conditional characteristic function of $\ell_j\delta^{-1}$.

The characteristic function is therefore explicit, up to the integral over the compact set $[0,1]^d$ which can be efficiently computed for reasonably large d using, for example, Legendre quadrature. The following lemma is a reminder that, since the support of the portfolio loss distribution is discrete and finite, we can compute it without approximation as the discrete Fourier transform of its the characteristic function.

Lemma 4.3.2. *The probability distribution of the portfolio loss is given by*

$$\mathbb{P}[L_t = k\delta] = \frac{1}{M+1} \sum_{m=0}^{M} \phi_{L_t}\left(\mu m\right) \mathrm{e}^{-\mathrm{i}\mu k m} \text{ for } k \in \{0, \ldots, M\}, \tag{4.13}$$

with $\mu = 2\pi/(M+1)$ *and* $\phi_{L_t}(\cdot)$ *is the characteristic function of* $L_t\delta^{-1}$.

Note that calculating directly this distribution is a combinatorial problem whose complexity is increasing exponentially fast with M. Equipped with Lemma 4.3.2, the computation boils down to an application of the Fast Fourier Transform (FFT) algorithm, which is of significant practical importance as long as evaluating the characteristic function is efficient.

As mentioned above, the assumption of loss unit and discretely supported portfolio losses appears already in (Andersen, Sidenius, and Basu 2003),(Andersen and Sidenius 2004),(Hull and White 2004), where the distribution is computed without approximation by a recursive algorithm. However, as will be shown in Section 4.4.1, the computational cost of this recursion increases much faster with both the support size and the number of factors than that of our approach.

The discrete Fourier inversion in Lemma 4.3.2 differs from the continuous Fourier inversion described in (Laurent and Gregory 2005),(Burtschell, Gregory, and Laurent 2009) which aims to approximate a continuous loss distribution. Since our approach provides quasi-closed expressions for the loss distribution, its scope is much wider, allowing notably the pricing of CDO squared and Credit Index Options without simulations.

Remark 4.3.3. *When the default intensities are driven by a stochastic process similar expressions can be derived for the default probabilities, see for example (Schönbucher and Schubert 2001), however their computations generally require costly numerical techniques such as simulations. Yet, combining the linear credit risk models described in Chapter 3 and polynomial factor copulas will result in tractable polynomial models with dependent default times and stochastic default intensities. Indeed, in that case, the joint default probability rewrites as an integral over the expectation of a polynomial in a polynomial diffusion which is an analytical expression, see (Filipović and Larsson 2016). Some examples of polynomial copulas are the Farlie-Gumbel-Morgenstern copula and other small order polynomial copulas found in (Nelsen 1999), and Bernstein copulas which can also be used to approximate any copula as discussed in (Sancetta and Satchell 2004).*

In the next section, we show that our framework allows us to price in quasi-closed form products as complex as tranches on credit portfolios, portfolios of such tranches, or credit index swaptions. While the market for some is booming (e.g., credit index swaption), other may have fallen out of fashion (e.g., CDO squared). Therefore, we emphasize that such examples are meant to illustrate the potential of combining factor copulas with discretely suported losses given default.

4.3.2 Pricing Multi-Name Credit Derivatives

In this section, we show that the loss distribution of more complex portfolios can also be retrieved explicitly for any horizon of time. We start with reminders on tranches and CDO squared, and derive their loss distributions. We then derive the joint distribution of the total number of defaulted entities and of the total loss, which is a necessary ingredient to price credit index swaptions.

Tranches and CDO Squared

A tranche on a credit portfolio is a derivative that pays a fraction of the realized portfolio losses above the attachment point a and below the detachment point b with $0 \leq a < b$, in exchange of regular payments functions of the effective tranche width. Define the tranche loss as

$$\mathscr{T}_t^{a,b} := \min\{\max\{L_t - a, 0\}, b - a\}, \tag{4.14}$$

and denote $\epsilon_a := \delta - (a \mod \delta)$. The following shows that the knowledge of the probability distribution of L_t implies that of $\mathscr{T}_t^{a,b}$.

Proposition 4.3.4. *The tranche loss $\mathscr{T}_t^{a,b}$ has a discrete support and its probability mass function is given by*

$$\begin{aligned}
\mathbb{P}\left[\mathscr{T}_t^{a,b} = 0\right] &= \sum_{m=0}^{\lfloor a/\delta \rfloor} \mathbb{P}\left[L_t = m\delta\right], \\
\mathbb{P}\left[\mathscr{T}_t^{a,b} = b - a\right] &= \sum_{m=\lceil b/\delta \rceil}^{M} \mathbb{P}\left[L_t = m\delta\right], \\
\mathbb{P}\left[\mathscr{T}_t^{a,b} = \epsilon_a + k\delta\right] &= \mathbb{P}\left[L_t = (k + \lceil a/\delta \rceil)\delta\right],
\end{aligned}$$

for any $k \in \mathbb{N}$ such that $0 < \epsilon_a + k\delta < b - a$, and where $\lfloor x \rfloor$ (respectively $\lceil x \rceil$) denotes the closest integer smaller (respectively larger) than x.

Similarly, a portfolio composed of multiple tranches from (potentially different) portfolios is known as a CDO squared. As for the tranche, its loss distribution can be computed explicitly, even when the defaults of obligors composing the different portfolios are assumed to be dependent. More formally, let us consider K tranches on portfolios written on (potentially different) obligors. For $k \in \{1\ldots, K\}$, we denote by $\mathscr{T}_{k,t}^{a_k,b_k}$ and $L_{k,t}$ the k-th tranche and portfolio loss, with a_k and b_k the k-th tranche attachment and detachment points. The CDO-squared loss (or simply squared loss) is

$$\mathscr{L}_t = \sum_{k=1}^{K} \mathscr{T}_{kt}^{a_k,b_k}.$$

Assume that for all k, we have

$$a_k \mod \delta = 0 \quad \text{and} \quad b_k \mod \delta = 0. \tag{4.15}$$

Then, each of the tranche losses as well as the squared loss have a discrete state space

$$\begin{aligned} \mathscr{T}_{kt}^{a_k,b_k} &\in \{0, \delta, 2\delta, \dots, b_k - a_k\} \quad \text{for } k \in \{1\dots, K\}, \\ \mathscr{L}_t &\in \{0, \delta, 2\delta, \dots, M_K\delta\}, \end{aligned}$$

where $M_K = \sum_{k=1}^{K}(b_k - a_k)/\delta$.

Corollary 4.3.5. *If Equation* (4.15) *holds for $k = 1, \dots, K$, then the characteristic function of the squared loss is*

$$\phi_{\mathscr{L}_t}(u) = \int_{\mathbb{R}^d} \prod_{k=1}^{K} \phi_{\mathscr{T}_{kt}}(u, v) dC_V(v)$$

where

$$\phi_{\mathscr{T}_{kt}}(u, v) = \sum_{n=1}^{(b_k-a_k)/\delta} \mathbb{P}\left[\mathscr{T}_{kt}^{a_k,b_k} = n\delta \mid V = v\right] \mathrm{e}^{\mathrm{i}un}$$

is the V-conditional characteristic function of $\mathscr{T}_{kt}\delta^{-1}$.

To compute $\phi_{\mathscr{T}_{kt}}$ one may use Proposition 4.3.4 applied to the V-conditional portfolio loss distribution, namely $\mathbb{P}[L_{kt} = m\delta \mid V = v]$. Applying Lemma 4.3.2 with $\mathscr{L}_t$ replacing L_t, one finally obtains the distribution of the squared loss. With the distribution of the squared loss, one can then price derivatives such as tranches on a portfolio of tranches.

Credit Index Swaption

A credit index swaption is an option on a credit index swap. Whereas tranches could be priced using the portfolio loss distribution only, the optionality embedded in such a option necessitates the joint distribution of the total number of defaulted entities and of the total loss, see Section 4.7 for details on the pricing formulas. Letting N_t be the number of defaulted entities at time t, that is

$$N_t = \sum_{j=1}^{N} \mathbb{1}_{\{\tau_j \le t\}}, \tag{4.16}$$

we need $\mathbb{P}[N_t = n, L_t = \delta k]$ for all $n = 0, \dots, N$, $k = 0, \dots, M$, and $t > 0$. The following lemma provides a generic expression for the joint distribution of (N_t, L_t).

Proposition 4.3.6. *The joint distribution of (N_t, L_t) is given by*

$$\mathbb{P}[N_t = n, L_t = \delta k] = \sum_{j=0}^{N} \sum_{l=0}^{M} \frac{\phi_{N_t,L_t}(\mu j, \nu l)\, e^{-\mathrm{i}\mu n j} e^{-\mathrm{i}\nu k l}}{(1+N)(1+M)}$$

with $\mu = 2\pi/(M+1)$, $\nu = 2\pi/(N+1)$, *and*

$$\phi_{N_t,L_t}(x,y) = \int_{[0,1]^d} \prod_{j=1}^{N} \left(1 - p_{j,t}(v) + p_{j,t}(v)\phi(x,y,v)\right) dC_V(v)$$

where $p_{j,t}(v)$ *is as in Proposition 4.3.1, and*

$$\phi(x,y,v) = \sum_{k=0}^{n_j} \mathbb{P}\left[\ell_j = \delta\, k \mid V = v\right] \mathrm{e}^{\mathrm{i}(x+yk)}.$$

Note that this results requires a two-dimensional discrete Fourier transform inversion as described in the proof. One may observe that $\phi(x,y,v)$ is the V-conditional characteristic function of $x + y\ell_j\delta^{-1}$ evaluated at one, that is

$$\phi(x,y,v) = \mathbb{E}\left[\mathrm{e}^{\mathrm{i}(x+y\ell_j\delta^{-1})} \mid V = v\right],$$

and that $\phi_{N_t,L_t}(x,y)$ is the characteristic function of $(N_t, L_t\delta^{-1})$ evaluated at (x,y), that is

$$\phi_{N_t,L_t}(x,y) = \mathbb{E}\left[\mathrm{e}^{\mathrm{i}(xN_t+yL_t\delta^{-1})}\right].$$

Remark 4.3.7. *When the loss amounts* ℓ_j *are homogeneous and independent from* V, *then the following more direct calculation can be applied*

$$\mathbb{P}\left[N_t = n, L_t = \delta k\right] = \mathbb{P}\left[L_t = k\delta \mid N_t = n\right] \mathbb{P}\left[N_t = n\right]$$

where $\mathbb{P}\left[N_t = n\right]$ *can be computed as in Lemma 4.3.2, and where*

$$\mathbb{P}\left[L_t = k\delta \mid N_t = n\right] = \mathbb{P}\left[\sum_{j=1}^{n} \ell_j = k\delta\right]$$

may also be derived using the discrete Fourier transform.

Note that, up to this point, we left unspecified the V-conditional distribution of the loss amounts. In the next section, we suggest a flexible specification for $\mathbb{P}\left[\ell_j = \delta\, k \mid V = v\right]$ for $k \in \{0,\dots,m_j\}$ and $j \in \{1,\dots,N\}$, which is required to compute the characteristic function of the portfolio loss in proposition 4.11 .

4.3.3 Beta-Binomial Loss Amounts

In this section, we assume that the loss amount distribution of each obligor can be dependent on the default times and others loss amounts. For each $j = 1,\dots,N$, we let the loss amount ℓ_j take value in a set of the form

$$\ell_j \in \{b_j\delta, (a_j+b_j)\delta, \dots, (n_j a_j + b_j)\delta\} \subset \{0, \delta, 2\delta, \dots, m_j\delta\}$$

with the integers $a_j, b_j, n_j \in \mathbb{N}$ such that $n_j a_j + b_j = m_j > 0$. Note that the two sets are equivalent when $a_j = 1$ and $b_j = 0$. The Beta-binomial model is obtained by assuming that the V-conditional distribution of the loss amount increment $(\ell_j \delta^{-1} - b_j)/a_j$ is a Beta-binomial random variable.

Definition 4.3.8 (The Beta-Binomial model)**.** *The V-conditional probability of loss is*

$$\mathbb{P}\left[\ell_j = (a_j k + b_j)\delta \mid V = v\right] = \int_{[0,1]} \mathbb{P}\left[Z = k \mid p, n_j\right] \pi_j\left(p \mid V = v\right) dp$$

for any $k = 0, \dots, n_j$, where $Z \sim Bin(n_j, p)$, that is

$$\mathbb{P}\left[X = k \mid p, n_j\right] = \binom{n_j}{k} p^k (1-p)^{n_j - k}$$

and with the Beta distribution

$$\pi_j\left(p \mid V = v\right) = \frac{p^{\alpha(v)-1}(1-p)^{\beta(v)-1}}{\mathrm{B}(\alpha(v), \beta(v))}$$

for some functions $\alpha : [0,1]^d \to \mathbb{R}_{++}$ and $\beta : [0,1]^d \to \mathbb{R}_{++}$.

Conditional on V the number of loss units experienced upon default is the sum of a constant b_j and of k units a_j where k follows a Binomial distribution with parameter p and support $0, \dots, n_j$. In addition, the probability p is random and distributed according to a Beta distribution with parameters $\alpha(v)$ and $\beta(v)$. Note that the functions α and β may be obligor specific.

Although the Beta-binomial specification may look intimidating, it is a well-studied flexible distribution that nests a large spectrum of distributions such as the Bernouilli (see below), the discrete uniform (when $\alpha = \beta = 1$), and asymptotically the binomial (for large α and β). An additional important feature is that an explicit expression is available for its probability mass function

$$\begin{aligned}\mathbb{P}\left[\ell_j = (a_j k + b_j)\delta \mid V = v\right] = & \frac{\Gamma(n_j + 1)}{\Gamma(k+1)\Gamma(n_j - k + 1)} \frac{\Gamma(\alpha(v) + \beta(v))}{\Gamma(\alpha(v))\Gamma(\beta(v))} \\ & \times \frac{\Gamma(k + \alpha(v))\Gamma(n_j - k + \beta(v))}{\Gamma(n + \alpha(v) + \beta(v))}.\end{aligned}$$

for any $k = 0, \dots, n_j$ and where Γ denotes the gamma function. The V-conditional loss amount mean and variance therefore also have a explicit expression

$$\mathbb{E}\left[\ell_j \mid V = v\right] = \left(a_j \frac{n_j \alpha(v)}{\alpha(v) + \beta(v)} + b_j\right) \delta$$

and

$$\mathrm{Var}\left[\ell_j \mid V = v\right] = \frac{n_j\alpha(v)\beta(v)(\alpha(v)+\beta(v)+n_j)}{(\alpha(v)+\beta(v))^2(\alpha(v)+\beta(v)+1)} {a_j}^2\delta^2.$$

Remark that the mean loss amount is positively correlated with V when the function $v \mapsto \alpha(v)/(\alpha(v)+\beta(v))$ is increasing on $[0,1]$.

Example 4.3.9 (Bernouilli model)**.** *The loss amount distribution reduces to a Bernoulli when* $n_j = 1$ *with probability*

$$p(v) = \frac{\Gamma(\alpha(v)+\beta(v))}{\Gamma(\alpha(v))} \times \frac{\Gamma(1+\alpha(v))}{\Gamma(1+\alpha(v)+\beta(v))}$$

which can take any value in $(0,1)$ *and thus also be arbitrary close to the Dirac delta function.*

Example 4.3.10 (Linear Beta-Binomial model)**.** *Assume that* $d = 1$ *and that the functions* α, $beta$ *are linear such that* $\alpha(v) = m_1 + m_2 v$ *and* $\beta(v) = m_3 + m_4 v$ *where* $m_i > 0$ *for all* $i = 1,\ldots,4$. *This specification is discussed in further details in Section 4.4.4.*

4.4 Numerical Analysis

In this section we illustrate the computational performance of our approach, and numerically study the properties of selected models with different dependence and loss given default assumptions.

4.4.1 Computational Performance

We show here that the discrete Fourier transform (DFT) method proposed in Section 4.3 is significantly more efficient that the recursive methods suggested in (Andersen, Sidenius, and Basu 2003),(Hull and White 2004). Note that for a loss support of size M the DFT is computationally equivalent to the numerical inversion of (Laurent and Gregory 2005) with M discretization points, yet the DFT returns the exact loss distribution.

We consider the standard one-factor and two-factor copula models. Figure 4.1 displays the computing time necessary to retrieve the probability mass function with the DFT and with the recursive method. The calculations have been performed on a single CPU from a standard personal computer in the R programming language. The DFT method is significantly faster than the recursive method in both cases: it takes roughly the same amount of time to retrieve a distribution with 1000 points with DFT and a 100 points with recursion.

4.4.2 Dependent Defaults with a Mixed Copula

We investigate the joint default probability and the total number of defaults density in a one-factor copula model with a mixed bivariate copula specification as defined in Equation (4.6).

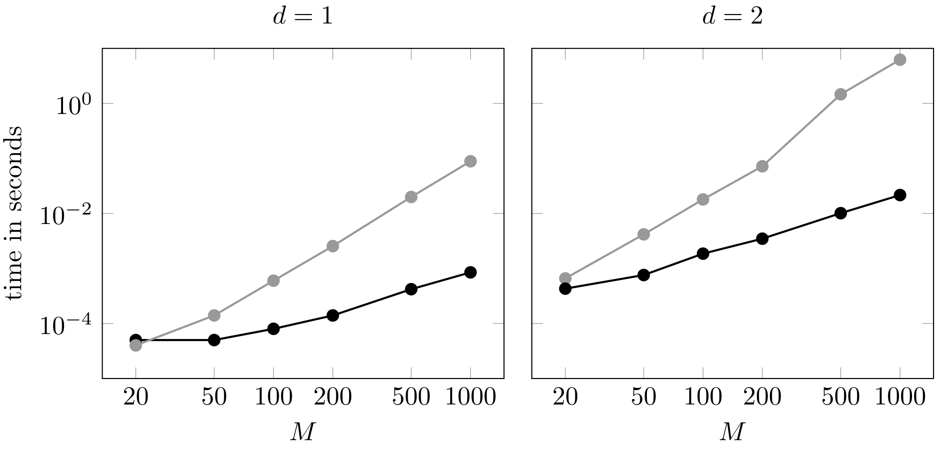


Figure 4.1 – Computation performance.
The time in seconds to compute the loss probability mass function is displayed against the loss support size M for the discrete Fourier transform (black line) and recursive (grey line) methods. The one-factor (left panel) and two-factor (right panel) standard Gaussian copula have been used under the assumption of constant loss given default.

We fix $K = 2$ and assume that $p_{j,t} = 1 - e^{-\lambda t}$ for $j \in \{1,2\}$ with $\lambda = 5\%$. Consider the following copula mixture

$$C_{U_j,V}(u_j, v) = wC^{\mathrm{C}}_{U_j,V}(u_j, v) + (1-w)C^{\mathrm{G}}_{U_j,V}(u_j, v)$$

for $j \in \{1,\dots,N\}$, for some $w \in [0,1]$, and where C^{C} denotes the Clayton copula with parameter 5 and C^{G} the Gaussian copula with parameter 25%. Figure 4.2 displays the probability and cumulative density functions of joint defaults of two entities for the times $0 \le t \le 20$, and for the weights $w \in \{0, 0.5, 1\}$. The two limit cases therefore correspond to the Gaussian and Clayton copulas. We observe that the joint probability of default also becomes a mixture of the two limit cases.

We set $N = 125$, Figure 4.3 displays the total number of defaults at a 5-years horizon. It is visually obvious that the distribution of the number of defaults is a mixture of the two limit components: it has the bump of the Gaussian with parameter $\rho = 25\%$ and the fat tail of the Clayton with parameter 5.

4.4.3 Credit Derivatives

We explore the loss distribution of a large portfolio, a tranche on this portfolio, and a portfolio of tranches when the underlying tranches are independent and when they depend on the same factor V. Let $N = 1000$ and assume that $\ell_j = 1$ and $\lambda_{jt} = 0.01$ for all $j \in \mathscr{I}$ and $t \ge 0$. The reference model is the standard one-factor Gaussian copula with correlation parameter $\rho = 25\%$. All the tranches have for attachment point $a_k = 100$ and detachment point $b_k = 200$. The CDO squared is composed of 10 tranches so as to have the same loss support as the

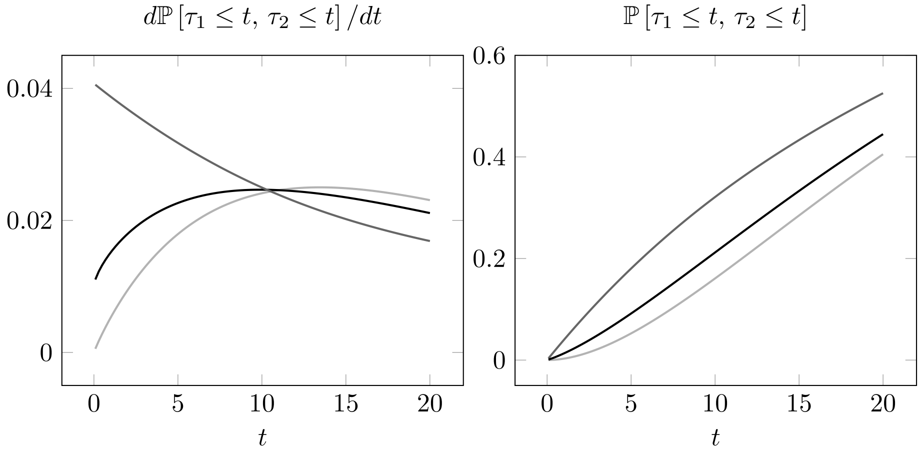


Figure 4.2 – Defaults dependence and copula mixture.
The probability (left panel) and cumulative (right panel) density functions of the joint default are displayed for time horizons ranging from 1 week to 20 years for three different one-factor models: an equiweighted copula mixture (black line) between a Gaussian copula with $\rho = 0.25$ (light-grey line) and a Clayton copula with parameter equal to 5 (grey line).

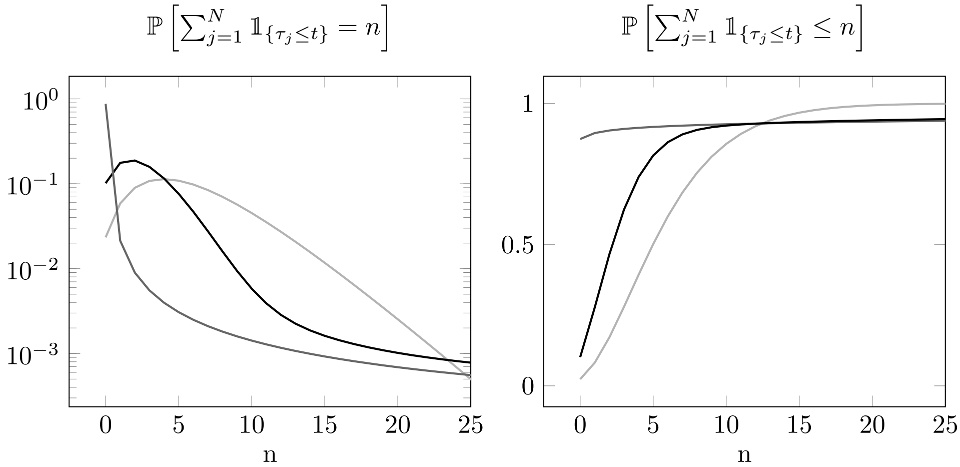


Figure 4.3 – Total number of defaults with copula mixture.
The Figure displays the probability (left panel) and cumulative (right panel) density functions of the total number of defaults on a portfolio of 125 homogeneous entities. We assume that $p_{j,t} = 5\%$ for all $j \in \mathscr{J}$ and consider an equiweighted copula mixture (black line) between a Gaussian copula with $\rho = 0.25$ (light-grey line) and a Clayton copula with parameter equal to 5 (grey line).

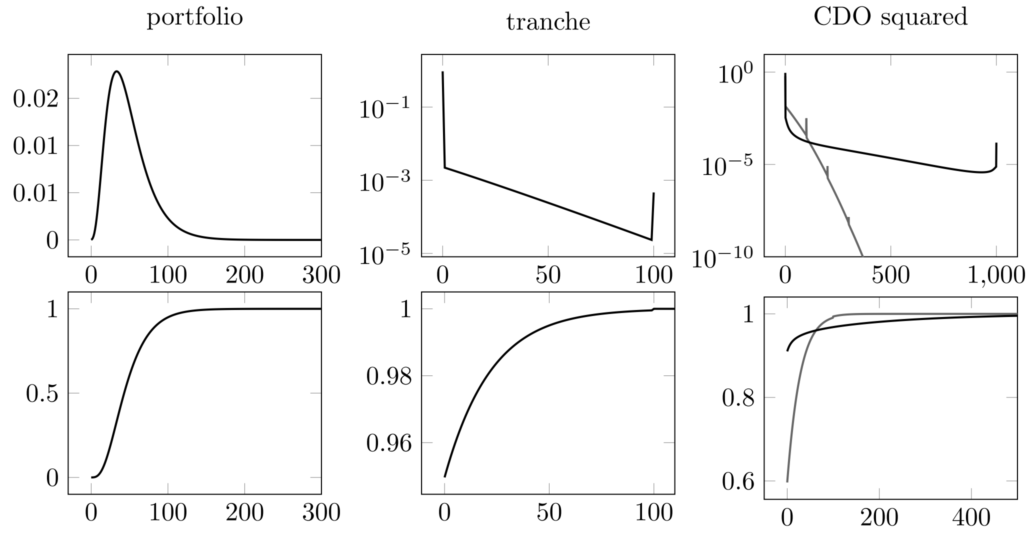


Figure 4.4 – Multi-name credit derivatives losses.
The probability (first row) and cumulative (second row) density functions of the loss distribution are displayed for three different derivatives. The first column is concerned with a portfolio of 1000 entities, the second column with a tranche on this portfolio with attachment point 100 and detachment point 200, and the third column with a portfolio of 10 such tranches coming from different portfolios with a unique risk factor (black line) and with independent risk factors (grey line).

portfolio.

Figure 4.4 displays the probability and cumulative mass functions of the portfolio, tranche, and portfolio of tranches at the 5-year horizon. Observe that the tranche loss distribution has two masses at the beginning and end of its support corresponding the probabilities of no loss and full loss respectively. These more concentrated masses combined and creates a spiky pattern in the portfolio of tranches loss distribution.

The CDO squared loss distribution has been computed under the assumption of unique factor and tranche specific factor. The two resulting loss distributions have dramatically different profiles. With independent factors the CDO squared appears even less exposed to losses than the vanilla portfolio. For example, the senior tranches on the pooled portfolio are virtually riskless. On the other hand, with a unique common factor the CDO squared has a fat tailed loss distribution and a large probability, about 91%, of having zero losses: when the risk driver behind all tranches is the same, the diversification benefit almost completely disappears. Similar results has been obtained (Hull and White 2010) using Monte Carlo simulations.

4.4.4 Stochastic and Correlated Loss Amounts

In this section we investigate the impact of introducing stochastic losses that may be correlated with the factor V on the loss distribution of a portfolio. We consider the linear Beta-Binomial model presented in Section 4.3.3 and always assume that $a_j = 1$ and $b_j = 0$.

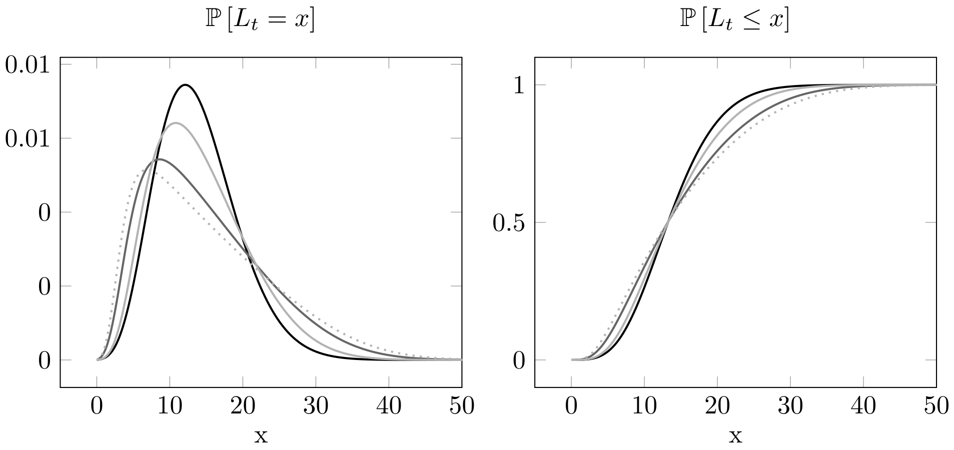


Figure 4.5 – Loss distribution and loss amounts dependence.
The probability (first row) and cumulative (second row) density functions of the loss distribution are displayed for three different loss amounts specifications. With a standard one-factor copula model with $\rho = 0.25$, constant marginal default intensity $\lambda_{jt} = 5\%$, and a 5-year horizon we consider the linear Beta-Binomial loss amounts with $m_1 = m_3$ and $m_2 = m_4$ for the values: $m_1 = 1$ and $m_2 = 0$ (black line), $m_1 = 1$ and $m_2 = 1$ (grey line), $m_1 = 3$ and $m_2 = 1$ (light-grey line), and $m_1 = 3$ and $m_2 = 5$ (dotted light-grey line).

Assume that $m_3 = m_1$ and $m_4 = m_2$ such that the V-conditional expected loss is

$$\mathbb{E}\left[\ell_j \mid V = v\right] = \frac{n_j a_j \delta(m_1 + m_2(1-v))}{2m_1 + m_2} + b_j \delta.$$

In this particular case the expected loss is constant is, that is

$$\mathbb{E}\left[\ell_j\right] = \int_0^1 \mathbb{E}\left[\ell_j \mid V = v\right] dv = \frac{1}{2},$$

for any m_1 and m_2 when $a_j = 1$, $b_j = 0$ and $n_j \delta = 1$. Consider the standard one-factor Gaussian copula with $\rho = 0.25$, $N = 125$, $\lambda_j = 0.05$ for all $j \in \mathscr{I}$, and with the same loss amount model as above having an expected loss one half. Figure 4.5 shows that the loss distribution is significantly affected by the choice of dependence parameters. Compared to the benchmark case of independent and equi-distributed loss amounts, increasing the dependence on the factor V also increases the portfolio average loss and tail risk.

4.4.5 Number of Defaults and Loss Dependence

We investigate how dependent individual losses affect the portfolio loss distribution given a number of realized defaults. We remind that this distribution is required to price credit swaptions. Consider the usual one-factor homogeneous Gaussian copula with $\rho = 0.25$, with default intensities $\lambda_{jt} = 0.05$, with $N = 125$ entities, and for a 5-year horizon. We assume that

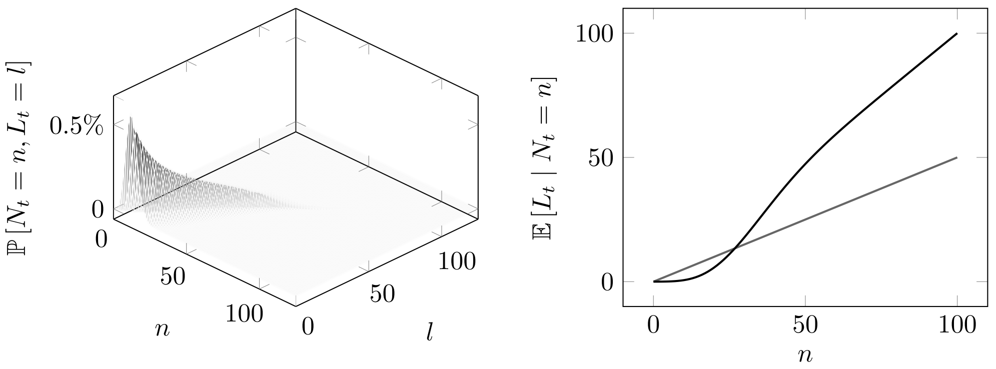


Figure 4.6 – Number of defaults and loss dependence.
The left panel displays the joint probability density of the number of default N_t and the loss L_t. The right panel displays the expected loss given n defaults, the loss amounts may not depend from the factor V (grey line) or may depend on the factor V (black line). The reference model is a one-factor homogeneous Gaussian copula with $\rho = 0.25$, default intensities $\lambda_{jt} = 0.05$, a 5-year horizon, and contains $N = 125$ entities. The loss amount ℓ_{jt} is zero or one and has an expected value of 50%.

the V-conditional loss amounts ℓ_j is given by

$$\ell_j = \begin{cases} 1 & \text{with probability } 1 - v \\ 0 & \text{with probability } v \end{cases}$$

such that $\mathbb{E}[\ell_j] = 0.5$ for all $j \in \mathscr{I}$. The left panel on Figure 4.6 displays the probability density of the joint probability distribution of the number of defaults and loss (N_t, L_t) computed as described in Proposition 4.2.2. We observe that most of the probability mass is concentrated on a diagonal band near the origin, and that there is little to no mass on the off diagonal parts.

The right panel of Figure 4.6 displays the expected loss given a certain number of default, that is $\mathbb{E}[L_t \mid N_t = n]$ for $n = 0, \ldots, 125$. This value is increasing with $N_t = n$ as the losses are expected to increase with the total number of defaults. Several interesting observations can be made. The marginal rate of losses starts from almost zero at the origin and increases rapidly, and the conditional expected loss converges to the maximal possible loss. This is in contrast with the case of independent loss amounts defined by $\mathbb{P}[\ell_j = 0] = \mathbb{P}[\ell_j = 1] = 0.5$ also displays on this Figure and where the relation between N_t and L_t is linear.

4.5 Empirical Analysis

In this section, we illustrate our approach by calibrating various factor copula models to credit index tranche prices.

Name	Attachement	Detachment	Spread
Equity	0	3	5
Mezzanine	3	7	1
Senior	7	15	1
Super-senior	15	100	0.25

Table 4.1 – Tranches structure on the CDX.NA.IG.21.
The attachment/detachment points and the spread per annum for each of the four tranches are given in percentage.

	Equity	Mezzanine	Senior	Super-senior
Mean	15.18	5.92	-0.29	-0.23
Vol	4.10	2.74	1.27	0.22
Min	8.59	1.34	-2.07	-0.53
Max	24.87	13.28	2.82	0.25

Table 4.2 – Summary statistics for the tranches on the CDX.NA.IG.21.
The statistics concern the upfront payments, which are quoted in percentage of the tranche width.

4.5.1 Data

We focus on tranches of the *CDX.NA.IG* index, which is composed of 125 investment grade North American companies. Historically, all tranches except the most junior were unfunded. Similarly as standard swaps, they were quoted with a spread and didn't include upfront payments. Since 2009 however, a new set of rules, known as the *Big Bang Protocol*, was amended to the International Swaps and Derivatives Association's master agreement (i.e., the standardized contract used between dealers and their counterparties). Arguably the most important was the *100/500 Credit Derivative Initiatives*: by standardizing coupons at 1% or 5% per annum[3] with quarterly payments, the rule made the upfront necessary to enter a contract on any tranche.

Based on liquidity, new series of the CDX.NA.IG index with tenors of 3, 5, 7, and 10 years are determined every 6 months (in March and September). The series 21, issued in September 2013 with a tenor of 5 years, came along with four standardized tranches, whose spreads and attachments/detachments points are detailed in Table 4.1. Our sample contains 405 daily upfront payments for the four tranches, which we summarize in Table 4.2 and display in Figure 4.7. By convention, the market quotes upfronts in percentage of the corresponding tranche width, which is about thirty times larger for the super-senior than for the equity. Furthermore, the sign of the upfront is also interesting: since it is negative, one most often receives money to buy protection on the super-senior tranche, as well as on the senior tranche at the end of the sample period.

---

[3] Although it has since been extended to also include coupons of 0.25% and 10%.

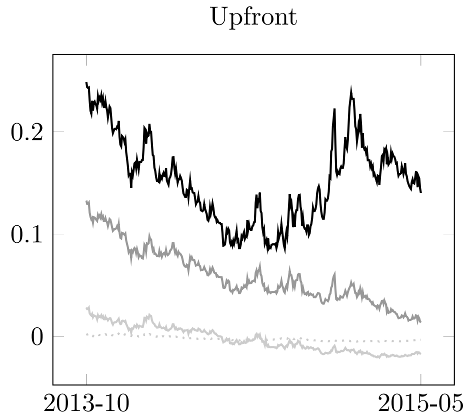


Figure 4.7 – Upfronts on CDX.NA.IG.21 tranches.
The time-series of quoted ufront payments are displayed for the equity (black line), mezzanine (grey line), senior (light-grey line) and super-senior (dotted light-grey line) tranches.

4.5.2 Calibration

Let P^{a_i,b_i}, a_i and b_i for $i \in \{1,\dots,4\}$ denote the quoted upfronts, and attachments/detachments points of each tranches. For a model parametrized with $\boldsymbol{\theta} \subseteq \Theta \subseteq \mathbb{R}^l$ (i.e., l is the number of parameters), we denote by $P^{a_i,b_i}(\boldsymbol{\theta})$ the model price, that is the quantity satisfying

$$P^{a_i,b_i}(\boldsymbol{\theta})(b_i - a_i) + V^{a_i,b_i}_{\text{prem}}(\boldsymbol{\theta}) = V^{a_i,b_i}_{\text{prot}}(\boldsymbol{\theta}),$$

where $b_i - a_i$ is the tranche width, and the premium and protection legs are defined as

$$V^{a_i,b_i}_{\text{prem}}(\boldsymbol{\theta}) = S^{a_i,b_i}\, \mathbb{E}_{\boldsymbol{\theta}}\left[\sum_{j=1}^{n} e^{-\int_0^{T_j} r_s ds}(T_j - T_{j-1}) \int_{T_{j-1}}^{T_j} \frac{b - a - \mathcal{T}^{a,b}_t}{T_j - T_{j-1}} dt\right],$$

and

$$V^{a_i,b_i}_{\text{prot}}(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{\theta}}\left[\int_0^T e^{-\int_0^t r_s ds} d\mathcal{T}^{a_i,b_i}_t\right],$$

with $0 = T_0 \le \cdots \le T_n = T$ the payment dates, T the maturity, and S^{a_i,b_i} the tranche spread. See Section 4.7 for more details.

Assuming $r = 0$ and a homogeneous portfolio with no recovery (i.e., $\delta_j = 0$), we let the default probability be

$$p_{j,t} = 1 - e^{-\lambda t},\ j \in \{1,\dots,125\}$$

where λ is the credit index swap spread, and the model is calibrated by minimizing the squared

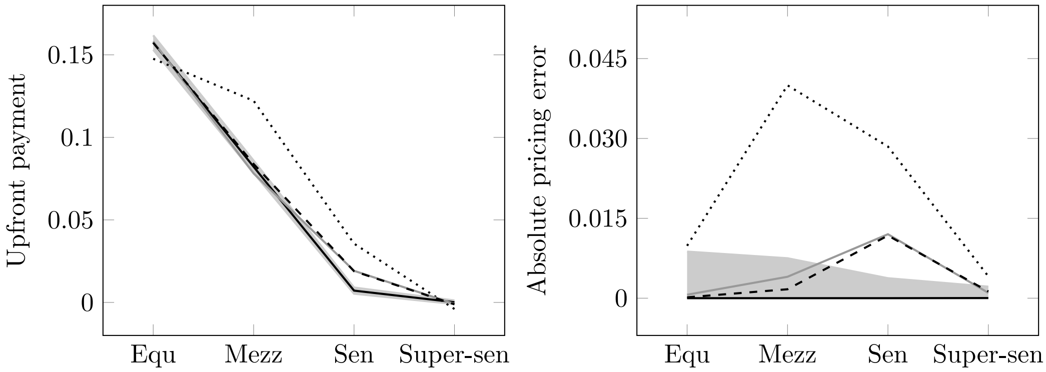


Figure 4.8 – Models calibration to tranches on the CDX.NA.IG.21.
Using the quoted upfronts on January 6th, 2014, various copulas models are calibrated: the one-factor Gaussian (dotted line), the one-factor t copula (dashed line), the two-factors Gaussian-Clayton copula (grey line), and the one-factor mixture with two Gaussians (black line). The shaded area is the bid-ask spread.

pricing error, that is

$$\widehat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta} \subseteq \Theta}{\operatorname{argmin}} \sum_{i=1}^{4} \left(P^{a_i,b_i} - P^{a_i,b_i}(\boldsymbol{\theta}) \right)^2. \tag{4.17}$$

In our current implementation, (4.17) is solved in two steps. First, we explore the parameter space to find a good starting value via a differential evolution algorithm. Second, we use the Nelder-Mead algorithm to refine the solution, enforcing the bounds by means of a parameter transformation.

4.5.3 Results

In Figure 4.8, we show calibration of various copulas to upfronts quoted on January 6th, 2014. While the one-factor Gaussian (dotted line) is completely off, both the one-factor t copula (dashed line) and the two-factors Gaussian-Clayton copula (grey line) perform better but miss the senior tranche. The only model achieving a perfect fit (i.e., the black line) is the following one-factor two-Gaussians mixture

$$C_{U_j,V}(u_j,v) = wC^{\rho_1}_{U_j,V}(u_j,v) + (1-w)C^{\rho_2}_{U_j,V}(u_j,v),\ j \in \{1,\dots,125\}$$

with $w \in [0,1]$ and C^{ρ_i} is a Gaussian copula with parameter ρ_i for $j \in \{1,2\}$ (i.e., $\boldsymbol{\theta} = (w,\rho_1,\rho_2)$ and $\Theta = [0,1] \times [-1,1] \times [-1,1]$).

Repeating (4.17) of the mixture for each day of the sample, we obtain time-series of calibrated parameters that we display as the plain lines in Figure 4.9. There are two interesting observations that can be made. First, the parameters do not vary much over time, which indicates that the model is not over-parametrized and can be reliably estimated. Second, the second parameter is very close to 1, which means the second component of the mixture describes

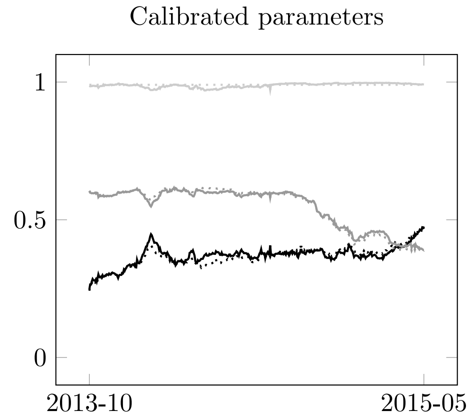


Figure 4.9 – Parameters calibrated on CDX.NA.IG.21 tranches.
The time-series of calibrated parameters are displayed for w (black lines) ρ_1 (grey lines) and ρ_2 (light-grey lines). The plain and dotted lines correspond to models with either three (plain) or two (dotted) parameters, that with either $\boldsymbol{\theta}_1 = (w, \rho_1, \rho_2)$ or $\boldsymbol{\theta}_2 = (w, \rho_1, 0.99)$.

a comonotonic relationship between the factor and the uniform random variables for each obligor. In other words, we have

$$\mathbb{P}\left[U_j \leq u_j \mid V = v\right] \approx \begin{cases} wC^{\rho_1}_{U_j|V}(u_j \mid v), & \text{if } u_j \leq v, \\ wC^{\rho_1}_{U_j|V}(u_j \mid v) + (1-w), & \text{otherwise} \end{cases}, \tag{4.18}$$

for $j \in \{1, \dots, 125\}$. When fixing $\rho_2 = 0.99$ such that (4.18) holds, and calibrating $\boldsymbol{\theta} = (w, \rho_1)$ only, similar results were obtained, and the parameters time series are the dotted lines in the upper-right panel of Figure 4.9.

In the four remaining panels of Figure 4.10, we display a model diagnostic for each of the four tranches. For each day in the sample period, the pricing errors, namely

$$P^{a_i,b_i} - P^{a_i,b_i}(\boldsymbol{\theta}_i) \text{ with } \begin{cases} \boldsymbol{\theta}_1 = (w, \rho_1, \rho_2) \\ \boldsymbol{\theta}_2 = (w, \rho_1, 0.99) \end{cases},$$

are the black and grey lines respectively for $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, and the bid-ask spread, that is

$$P^{a_i,b_i}_{ask} - P^{a_i,b_i}_{bid},$$

are the the light grey lines. As the pricing errors are much lower than the bid-ask spread, the equity and mezzanine tranches are perfectly calibrated by both models. For the senior tranche with $\boldsymbol{\theta}_2$ and the super-senior tranche however, the pricing errors and the bid-ask spread have the same order of magnitude. To alleviate this issue, we could switch the target of the minimization in the right-hand side of (4.17) from percentage of the tranche width to dollar amount. In other words, by weighting each term of the sum by $(b_i - a_i)^2$, we would increase the relative importance of the super-senior tranche in the objective function. Nonetheless, the

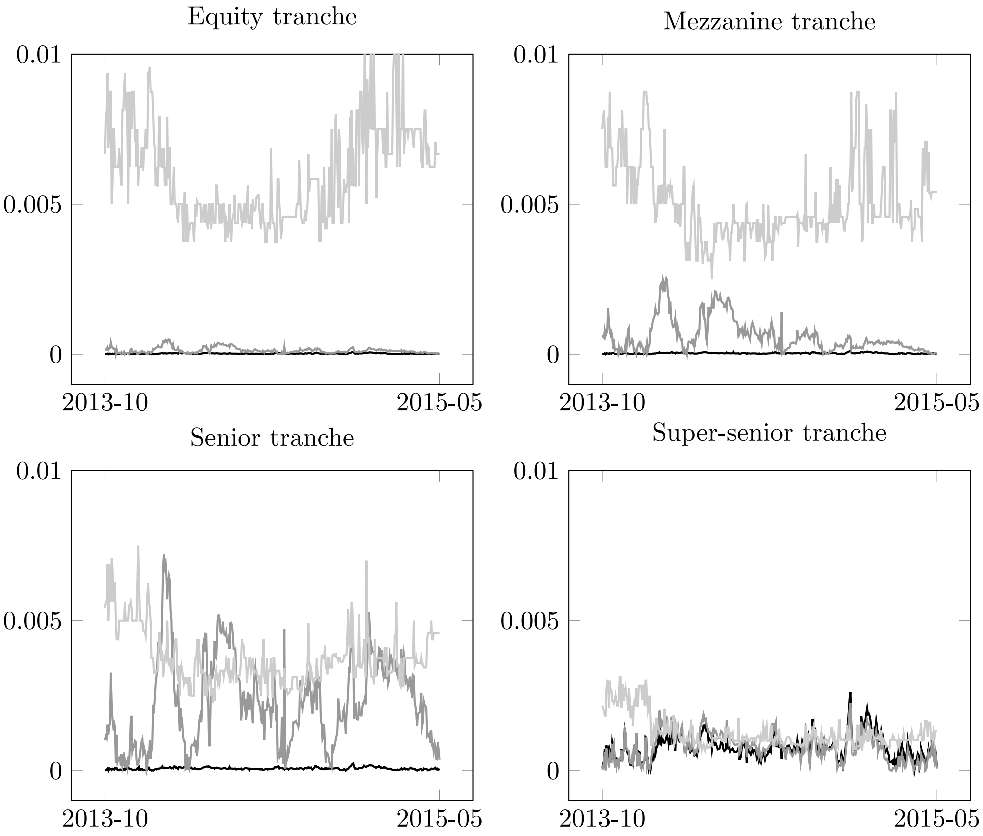


Figure 4.10 – Diagnostic of models calibrated on CDX.NA.IG.21 tranches.
Model diagnostic are displayed with the bid-ask spread (ligh-grey line) and the pricing errors for the model with either three (black line) or two (grey line) parameters, that with either $\boldsymbol{\theta}_1 = (w, \rho_1, \rho_2)$ or $\boldsymbol{\theta}_2 = (w, \rho_1, 0.99)$.

pricing error (and the bid-ask spread) are between 10 and 30 times smaller than the upfront itself.

To summarize, we achieve an almost perfect calibration to all tranches with only two parameters that remain stable over time.

4.6 Standard Copula Models

We derive in this section the factor copula representation of the most popular models that have been proposed in the literature on multi-name credit risk.

Gaussian copula models. Let us denote the Gaussian copula and h-function by

$$C^G_{U,V}(u, v; \rho) = \Phi_2\left(\Phi^{-1}(u), \Phi^{-1}(v); \rho\right)$$

and

$$C^G_{U;V}(u \mid v;\rho) = \Phi\left(\frac{\Phi^{-1}(u) - \rho\Phi^{-1}(v)}{1-\rho^2}\right),$$

where $\Phi(\cdot)$ is the standard normal distribution and $\Phi_2(\cdot,\cdot;\rho)$ is the bivariate normal distribution with correlation ρ. For instance, when $d = 1$ and all bivariate copulas are Gaussian, then a representation for the joint distribution of default times is the copula of a 1-factor model

$$Y_j = \beta_j X + \sqrt{1-\beta_j^2} Z_j,$$

where $X, Z_1, \dots, Z_N$ are i.i.d. $N(0,1)$ random variables. In this case, the correlation parameter for the bivariate copula linking the default of obligor j to the systematic factor is β_j. By considering a unique correlation parameter $\beta_j = \rho$ for $j \in \{1,\dots,N\}$, (Li 2000) is a special case of our formulation. Furthermore, when $d > 1$, then a representation for the joint distribution of default times is the copula of a d-factor model

$$Y_j = \sum_{i=1}^{p} \beta_{j,i} X_i + Z_j,$$

where $X_1,\dots,X_p, Z_1,\dots,Z_N$ are i.i.d. $N(0,1)$ random variables. In this case, the parameters for the second to d factors are partial correlations, namely

$$\rho_{U_j,V_k|X_1,\dots,V_{k-1}} = \frac{Cov(Y_j, X_k \mid X_1,\dots,X_{k-1})}{\sqrt{Var(Y_j \mid X_1,\dots,X_{k-1})}\sqrt{Var(X_k \mid X_1,\dots,X_{k-1})}} = \frac{\beta_{j,k}}{\sqrt{1-\beta_{j,1}^2 - \cdots - \beta_{j,k-1}^2}}.$$

Stochastic correlation models. It is straightforward to build more complex factor models, stochastic correlations models are obtained by writing

$$Y_j = \left(B_j\alpha_j + (1-B_j)\beta_j\right) X + \sqrt{1-\left(B_j\alpha_j + (1-B_j)\beta_j\right)^2} Z_j,$$

where B_j are i.i.d. Bernoulli(b_j) and $X, Z_1,\dots,Z_N$ as before. For this model, the bivariate copulas are convex sum of Gaussian copulas, that is

$$\begin{aligned} C^{SC}_{U_j,V}(u_j, v;\alpha_j,\beta_j,b_j) &= b_j^2 C^G_{U_j,V}(u_j,v;\alpha_j^2) + 2b_j(1-b_j)C^G_{U_j,V}(u_j,v;\alpha_j\beta_j) \\ &\quad + (1-b_j)^2 C^G_{U_j,V}(u_j,v;\beta_j^2), \end{aligned}$$

and deriving the h-function yields

$$C^{SC}_{U|V}(u \mid v;\alpha_j,\beta_j,b_j) = b_j C^G_{U|V}(u \mid v;\alpha_j) + (1-b_j)C^G_{U|V}(u \mid v;\beta_j).$$

The t-Student model. Usually, t-student models are specified by considering,

$$Y_j = \sqrt{W}\left(\beta_j X + \sqrt{1-\beta_j^2} Z_j\right)$$

where W is an i.i.d. random variable such than ν/W is $\chi^2(\nu)$ and $X, Z_1, \ldots, Z_N$ as before. Then the default times are independent conditional on (W, X) and their conditional probability distribution is easily derived (see e.g. (Burtschell, Gregory, and Laurent 2009)). Using our formulation, we obtain an equivalent t-student model by considering the copula and h-function directly, that is

$$C^t_{U,V}(u, v; \rho, \nu) = t_2(t_\nu^{-1}(u), t_\nu^{-1}(v); \rho, \nu)$$

and

$$C^t_{U;V}(u, v; \rho, \nu) = t_{\nu+1}\left(f(u, v)\right), \text{ with } f(u, v) = \frac{t_\nu^{-1}(u) - \rho t_\nu^{-1}(v)}{\sqrt{\frac{(1-\rho^2)\left(\nu+\left(t_\nu^{-1}(v)\right)^2\right)}{\nu+1}}},$$

where $t_\nu(\cdot)$ is the t-student distribution with ν degrees of freedom and $t_2(\cdot, \cdot; \rho, \nu)$ is the bivariate t-student distribution with correlation ρ and degrees of freedom ν. Compared to the other formulation, our alternative only require a one-dimensional integration. Furthermore, using different degrees of freedom for each bivariate copulas offers additional modeling flexibility without additional cost.

Archimedean models. One-parameter archimedean copulas are built by considering a continuous, strictly decreasing and convex generator $\psi : [0, 1] \times \Theta \to [0, \infty)$ such that $\psi(1; \theta) = 0$ for all $\theta \in \Theta$, where Θ represents the parameter space. Using this generator, a bivariate copula is obtained by writing

$$C^\psi_{U,V}(u, v; \theta) = \psi^{-1}\left(\psi(u; \theta) + \psi(v; \theta); \theta\right).$$

For such a copula, the h-function $C^\psi_{U;V}$ is usually straightforward to derive, and we summarize the most popular in Table 4.3.

Gaussian mixture models. The Gaussian mixture model from (Li and Liang 2005) is a one-factor copula mixture as described in Equation 4.6. The bivariate copulas are furthermore assumed to be Gaussian and equal $C^k_{U_j,V} = C^k_{U_l,V}$ for any $j, l = 1, \ldots, N$ and $k = 1, \ldots, K$.

4.7 Pricing Formulas

The credit contracts described in this section are composed of two cash-flows series. The contract buyer pays predefined coupons to the seller at the payments dates $0 = T_0 \leq \cdots \leq T_n =$

	Generator ψ	Inverse generator ψ^{-1}	Parameter space Θ
Clayton	$\frac{u^{-\theta}-1}{\theta}$	$(1+\theta u)^{-1/\theta}$	$(0,\infty)$
Gumbel	$\left(-\log(u)\right)^{\theta}$	$\exp\left(-u^{1/\theta}\right)$	$[1,\infty)$
Frank	$-\log\left(\frac{\exp(-\theta u)-1}{\exp(-\theta)-1}\right)$	$-\frac{1}{\theta}\log\left(1+\exp(-t)\left(\exp(-\theta)-1\right)\right)$	$(-\infty,\infty)\setminus\{0\}$
Joe	$-\log\left(1-(1-u)^{\theta}\right)$	$1-\left(1-\exp(-u)\right)^{1/\theta}$	$[1,\infty)$
Independence	$-\log(u)$	$\exp(-u)$	$\varnothing$

	Copula $C^{\psi}_{U,V}$	h-function $C^{\psi}_{U\vert V}$
Clayton	$\left(u^{-\theta}+v^{-\theta}-1\right)^{-1/\theta}$	$C^{\psi}_{U,V}(u,v;\theta)\,v^{-1-\theta}$
Gumbel	$e^{-\left((-\log(u))^{\theta}+(-\log(v))^{\theta}\right)^{1/\theta}}$	$C^{\psi}_{U,V}(u,v;\theta)\,\frac{\left((-\log(u))^{\theta}+(-\log(v))^{\theta}\right)^{1/\theta-1}(-\log(v))^{\theta}}{v\log(v)}$
Frank	$-\frac{1}{\theta}\log\left(\frac{1-e^{-\theta}-\left(1-e^{-u\theta}\right)\left(1-e^{-v\theta}\right)}{1-e^{-\theta}}\right)$	$\frac{e^{\theta}\left(e^{\theta u}-1\right)}{e^{\theta u+\theta v}-e^{\theta u+\theta}-e^{\theta v+\theta}+e^{\theta v}}$
Joe	$1-\left((1-u)^{\theta}+(1-v)^{\theta}-(1-u)^{\theta}(1-v)^{\theta}\right)^{1/\theta}$	$\left(C^{\psi}_{U,V}(u,v;\theta)\right)^{1-\theta}(1-v)^{\theta-1}(1-(1-u)^{\theta})$
Independence	uv	u

Table 4.3 – Archimedean copulas

Describes the generator ψ, the inverse generator ψ^{-1}, the parameter space Θ, the copula $C^{\psi}_{U,V}$, and the h-function $C^{\psi}_{U|V}$.

T where T is the contract maturity, we call this series of cash-flow the premium leg V_{prem}. The contract seller pays default contingent cash-flows to the buyer at the defaults dates when losses materialize, we call this series of cash-flow the protection leg V_{prot}. The contract value for the buyer is then given by $V_{\text{prot}} - V_{\text{prem}}$. We denote r_t the time-t risk-free rate and derive the contracts values at the initial date.

Tranche. A credit swap on a tranche $\mathscr{T}^{a,b}$ with attachment point a and detachment point b is a protection insuring the loss experienced on the tranche, in exchange of scheduled premium payments proportional to the remaining size of the tranche. The value of the protection leg is

$$V_{\text{prot}} = \mathbb{E}\left[\int_0^T e^{-\int_0^t r_s ds} d\mathscr{T}^{a,b}_t\right]$$

where the tranche density $\mathscr{T}^{a,b}_t$ is defined in Equation (4.14). The value of the premium leg is

$$V_{\text{prem}} = S^{a,b}\,\mathbb{E}\left[\sum_{j=1}^{n} e^{-\int_0^{T_j} r_s ds}(T_j - T_{j-1})\int_{T_{j-1}}^{T_j} \frac{b-a-\mathscr{T}^{a,b}_t}{T_j - T_{j-1}} dt\right]$$

where $S^{a,b}$ is the tranche spread. In practice, the above expressions for the two legs are necessarily approximated by replacing the integrals with sums where quadrature and trapezoid methods cab be used. Assuming that the short-rate and the default times are uncorrelated, (Mortensen 2006) used the parsimonious following discretization

$$V_{\text{prot}} \approx \sum_{j=1}^{n} B\left(\frac{t_j + t_{j-1}}{2}\right)\left(\mathbb{E}\left[\mathscr{T}^{a,b}_{t_j} - \mathscr{T}^{a,b}_{t_{j-1}}\right]\right)$$

and

$$V_{\text{prem}} \approx S^{a,b} \sum_{j=1}^{n} (t_j - t_{j-1}) B(t_j) \left(b - a - \frac{\mathbb{E}\left[\mathcal{T}_{t_j}^{a,b} + \mathcal{T}_{t_{j-1}}^{a,b}\right]}{2} \right)$$

where $B(t)$ denotes the risk-free bond price with maturity t and notional equal to one.

Credit swap. A credit swap on an index pays the realized losses in exchange for scheduled premium payments proportional to the number of non-defaulted entities. The value of the protection leg is

$$V_{\text{prot}} = \mathbb{E}\left[\sum_{j=1}^{N} e^{-\int_0^{\tau_j} r_s ds} \ell_j \right] = \mathbb{E}\left[\int_0^T e^{-\int_0^t r_s ds} dL_t \right]$$

where the total loss L_t is defined in Equation (4.10). The value of the premium leg is

$$V_{\text{prem}} = S\,\mathbb{E}\left[\sum_{j=1}^{n} e^{-\int_0^{T_j} r_s ds} (T_j - T_{j-1})(N - N_t) + \int_{T_{j-1}}^{T_j} e^{-\int_0^u r_s ds} (u - T_{j-1}) dN_u \right]$$

where S is the index spread and N_t is the time-t total number of default defined in Equation (4.16).

Credit swaption. A European credit swaption offers the right to enter a credit swap at a future time $0 < T_m < T$ at a predefined index spread S^*. In addition, the credit swaption typically provides default protection between the issuance date 0 and the option maturity T_m. Denote $V_{\text{CS}}(T_m)$ the time-T_m value of the credit swap which follows directly from above. The time-T_m value of the credit swaption is given by

$$V_{\text{CSO}} = \mathbb{E}\left[e^{-\int_0^{T_m} r_s ds} \Big(V_{\text{CS}}(T_m) + (L_{T_m} - L_0) \Big)^+ \right].$$

4.8 Proofs

This Section contains the proofs of all theorems and propositions in the main text.

Proof of Lemma 4.2.2

The joint probability of default rewrites

$$\mathbb{P}\left[\tau_1 \le t_1, \dots, \tau_N \le t_N\right] = \mathbb{P}\left[U_1 \le p_{1,t_1}, \dots, U_N \le p_{N,t_N}\right] = C_U\left(p_{1,t_1}, \dots, p_{N,t_N}\right)$$

where the second line follows by definition of C_U.

Proof of Proposition 4.2.3

Observe that for all $j = 1, \dots, N$ the random vector (U_j, V) takes values on $[0,1]^2$ and has uniform marginal densities, this implies that

$$\mathbb{P}\left[U_j \le u_j, V \le v\right] = C_{U_j,V}(u_j, v)$$

for some bivariate copulas $C_{U_j,V}$ and any $(u_j, v) \in [0,1]^2$. Therefore we have

$$\mathbb{P}\left[U_j \le u_j \mid V = v\right] = C_{U_j|V}(u_j \mid v)$$

and by plugging this into Equation (4.5) then integrating with respect to the density $f_V(v) = v$ of V we obtain

$$C_U(u_1, \dots, u_N) = \int_0^1 \prod_{j=1}^N \mathbb{P}\left[U_j \le u_j, V \le v\right] f_V(v) dv = \int_0^1 \prod_{j=1}^N C_{U_j,V}(u_j, v) dv.$$

The desired expression then follows from Lemma 4.2.2.

Proof of Proposition 4.2.6

We denote U_{I} the vector which contains the coordinates of U which are in $\mathscr{I} \setminus \mathscr{D}$, and U_{J} contains the ones which are in $\mathscr{D}$. For readability we assume that the coordinates are ordered according to $U = (U_{\mathrm{I}}, U_{\mathrm{J}})$. Similarly we group the marginal default probabilities into two vectors p_{I} and p_{J}. The size of U_{I} and U_{J} are respectively given by N_{I} and $N_{\mathrm{J}} = N - N_{\mathrm{I}}$. We directly write the proof for the multivariate case with $V \in [0,1]^d$. The joint default probability conditional on the default of the $k \in \mathscr{D}$ entities then rewrites

$$\begin{aligned}\mathbb{P}\left[\tau_1 \le t_1, \dots, \tau_N \le t_N \mid \tau_k = t_k : k \in \mathscr{D}\right] &= \mathbb{P}\left[U_1 \le p_{1,t_1}, \dots, U_N \le p_{N,t_N} \mid U_k = p_{k,t_k} : k \in \mathscr{D}\right] \\ &= \frac{\int_{[0,p_{\mathrm{I}}]} c_{U_{\mathrm{I}},U_{\mathrm{J}}}(u_{\mathrm{I}}, p_{\mathrm{J}}) du_{\mathrm{I}}}{\int_{[0,1]^{N_{\mathrm{I}}}} c_{U_{\mathrm{I}},U_{\mathrm{J}}}(u_{\mathrm{I}}, p_{\mathrm{J}}) du_{\mathrm{I}}} = \frac{\int_{[0,1]^d} C_{U_{\mathrm{I}}|V}(p_{\mathrm{I}} \mid v) c_{U_{\mathrm{J}},V}(p_{\mathrm{J}}, v) dC_V(v)}{\int_{[0,1]^d} c_{U_{\mathrm{J}},V}(p_{\mathrm{J}}, v) dC_V(v)}.\end{aligned}$$

The second equality comes from the definition of the conditional probability for measures. The third equality follows from the definition of the factor copula and Fubini's theorem,

$$\int_{[0,x]} c_{U_{\mathrm{I}},U_{\mathrm{J}}}(u_{\mathrm{I}}, p_{\mathrm{J}})\, du_{\mathrm{I}} = \int_{[0,1]^d} \prod_{U_j \in U_{\mathrm{I}}} \int_0^{x_j} c_{U_j,V}(u, v)\, du \prod_{U_j \in U_{\mathrm{J}}} c_{U_j,V}(u_j, v)\, dC_V(v),$$

along with $\int_0^{x_j} c_{U_j,V}(u, v) du = C_{U_j|V}(x_j \mid v)$ and $C_{U_j|V}(1 \mid v) = 1$.

Alternatively, the same result can be proved by first showing that

$$\begin{aligned}\mathbb{P}\left[\tau_1 \le t_1, \dots, \tau_N \le t_N \mid \{\tau_k = t_k : k \in \mathscr{D}\} \cup \{V = v\}\right] &= \frac{\int_{[0,p_{\mathrm{I}}]} c_{U_{\mathrm{I}},V}(x,v)dx\, c_{U_{\mathrm{J}},V}(p_{\mathrm{J}},v)\; c_V(v)}{\int_{[0,1]^{N_{\mathrm{I}}}} c_{U_{\mathrm{I}},V}(x,v)dx\, c_{U_{\mathrm{J}},V}(p_{\mathrm{J}},v)\; c_V(v)} \\ &= \prod_{j \in \mathscr{I} \setminus \mathscr{D}} C_{U_j|V}(p_{j,t_j} \mid v)\end{aligned}$$

and then integrating with respect to the following conditional density

$$\mathbb{P}[V \le v \mid \{U_k = p_{k,t_k} : k \in \mathscr{D}\}] = \frac{\int_{[0,v]} \prod_{j \in \mathscr{D}} c_{U_j,V}(p_{j,t_j}, x) dC_V(x)}{\int_{[0,1]^d} \prod_{j \in \mathscr{D}} c_{U_j,V}(p_{j,t_j}, x) dC_V(x)}.$$

Proof of Propostion 4.2.7

The V-conditional joint default probability as a similar expression as in Equation (4.4). The unconditional joint default probability follows by integrating with respect to the joint density $c_V(v)$ of V which gives the expression for C_U as $dC_V(v) = c_V(v)dv$. Observe now that the joint distribution of the random vector (U_j, V) is by construction given by a $(1+d)$-dimensional copula $C_{U_j,V}$ for all $j \in \mathscr{I}$. By definition we must have

$$\begin{aligned}C_{U_j,V}(u_j, v) = \mathbb{P}\left[U_j \le u_j,\, V \le v\right] &= \int_0^{v_1} \dots \int_0^{v_d} \mathbb{P}\left[U_j \le u_j \mid V = y\right] d\mathbb{P}\left[V \le y\right] \\ &= \int_0^{v_1} \dots \int_0^{v_d} C_{U_j|V}(u_j \mid y) dC_V(y)\end{aligned}$$

for all $(u_j, v) \in [0,1]^{1+d}$ which gives Equation (4.8).

Proof of Corollary 4.2.8

The density of V is given by $C_V(v) = \prod_{j=1}^d v_j$, and following (Joe 1996) the conditional copulas are given by

$$C_{U_j|V}(u_j \mid v) = \frac{\partial C_{U_j,V_k|V_{-k}}\left(C_{U_j|V_{-k}}(u_j \mid v_{-k}), v_k \mid v_{-k}\right)}{\partial v_k}$$

for any $k = 1, \dots, d$, and where $V_{-k} = (V_1, \dots, V_{k-1}, V_{k+1}, \dots, V_d)$ denotes the random vector V without its k-th coordinate. By iterating the previous equation, the conditional copula $C_{U_j|V}(u_j \mid v)$ can be rewritten as a recursive composition of bivariate linking copulas

$$C_U(u_1, \dots, u_n) = \int_{[0,1]^d} \prod_{j=1}^N C_{U_j|V_1}(\cdot|v_1) \circ \dots \circ C_{U_j|V_d}(u_j|v_d) dv$$

where C_{U_j,V_k} denotes a bivariate copula for $j = 1, \dots, N$ and $k = 1, \dots, d$.

Proof of Theorem 4.2.9

Observe that the random vector $U = (F_{Y_1}(Y_1), \dots, F_{Y_N}(Y_N))$ and $V = (F_{X_1}(X_1), \dots, F_{X_d}(X_d))$ have uniform margins by construction suggesting that their distributions are given by copulas. The following theorem proves the existence of C_V.

Theorem 4.8.1 (Sklar's Theorem 1959)**.** *F_V is a joint distribution with margins F_{X_i} for $i \in \{1, \cdots, d\}$ if and only if there exists a copula C_V, that is a distribution which is supported in the unit hypercube and has uniform margins, such that*

$$F_X(x_1, \dots, x_N) = C_V\left(F_{X_1}(x_1), \dots, F_{X_N}(x_N)\right)$$

for all $x \in \mathbb{R}^N$. Moreover, if the margins are continuous, then C_V is unique.

For all $v \in [0,1]^d$ the theorem implies that

$$\begin{aligned} C_V(v_1, \dots, v_d) &= F_X\left(F_{X_1}^{-1}(v_1), \dots, F_{X_d}^{-1}(v_d)\right) = \mathbb{P}\left[X_1 \le F_{X_1}^{-1}(v_1), \dots, X_d \le F_{X_d}^{-1}(v_d)\right] \\ &= \mathbb{P}\left[F_{X_1}(X_1) \le v_1, \dots, F_{X_d}(X_d) \le v_d\right] = \mathbb{P}\left[V_1 \le v_1, \dots, V_d \le v_d\right]. \end{aligned}$$

The copula C_V is thus the joint distribution of probability integral transforms. The X-conditional independence of Y implies that

$$\mathbb{P}\left[Y_1 \le y_{1t_1}, \dots, Y_N \le y_{Nt_N} \mid X = x\right] = \prod_{j=1}^{N} F_{Y_j|X}(y_{jt_j} \mid x),$$

where $F_{Y_j|X}$ denotes the distribution of Y_j conditional on X such that

$$\mathbb{P}\left[\tau_j \le t_j \mid X = x\right] = \mathbb{P}\left[U_j \le p_{j,t_j} \mid V = v\right],$$

where $v = \tilde{F}_X(x) := \left(F_{X_1}(x_1), \dots, F_{X_d}(x_d)\right)$. A copula representation of the above probability can finally be obtained by applying the conditional equivalent of Sklar's theorem:

Theorem 4.8.2 (Patton's Theorem 2002)**.** *$F_{Y|X}$ is a joint conditional distribution with conditional margins $F_{Y_i|X}$ for $i \in \{1, \cdots, N\}$ if and only if there exists a conditional copula $C_{U|V}$, that is a conditional distribution which is supported in the unit hypercube and has uniform conditional margins, such that*

$$F_{Y|X}\left(y_1, \dots, y_N \mid x\right) = C_{U|V}(F_{Y_1|X}(y_1 \mid x), \dots, F_{Y_N|X}(y_N \mid x) \mid F_X(x))$$

for all $y \in \mathbb{R}^N$ and $x \in \mathbb{R}$. Moreover, if the conditional margins are continuous, then $C_{U|V}$ is unique.

For all $u \in [0,1]^N$ and $v \in [0,1]^d$ the theorem implies

$$\begin{aligned} C_{U|V}(u_1, \dots, u_N \mid v) &= F_{Y|X}\left(F^{-1}_{Y_1|X}(u_1), \dots, F^{-1}_{Y_N|X}(u_N) \mid \tilde{F}^{-1}_X(v)\right) \\ &= \mathbb{P}\left[Y_1 \le F^{-1}_{Y_1|X}(u_1), \dots, Y_N \le F^{-1}_{Y_N|X}(u_N) \mid X = \tilde{F}^{-1}_X(v)\right] \\ &= \mathbb{P}\left[F_{Y_1|X}(Y_1 \mid X) \le u_N, \dots, F_{Y_N|X}(Y_N \mid X) \le u_N \mid \tilde{F}_X(X) = v\right] \\ &= \mathbb{P}\left[U_1 \le u_1, \dots, U_N \le u_N \mid V = v\right]. \end{aligned}$$

In other words, the copula $C_{U|V}$ is also the joint conditional distribution of the conditional probability integral transforms. As such, the joint conditional distribution of default times is given by

$$\begin{aligned} \mathbb{P}\left[\tau_1 \le t_1, \dots, \tau_N \le t_N \mid X = F^{-1}_X(v)\right] &= \mathbb{P}\left[U_1 \le p_{1,t_1}, \dots, U_N \le p_{N,t_N} \mid V = v\right] \\ &= C_{U|V}\left(p_{1,t_1}, \dots, p_{N,t_N} \mid v\right), \end{aligned}$$

which completes the proof.

Proof of Proposition 4.3.1

The default times and the loss amounts being independent conditional on V we have

$$\mathbb{E}\left[e^{iuL_t\delta^{-1}} \mid V = v\right] = \mathbb{E}\left[e^{iu\sum_{j=1}^N \mathbb{1}_{\{\tau_j \le t\}}\ell_j\delta^{-1}} \mid V = v\right] = \prod_{j=1}^N \mathbb{E}\left[e^{iu\mathbb{1}_{\{\tau_j \le t\}}\ell_j\delta^{-1}} \mid V = v\right]$$

Furthermore, by independence of the random variables $\mathbb{1}_{\{\tau_j \le t\}}$ and ℓ_j conditional on V we have

$$\mathbb{E}\left[e^{iu\mathbb{1}_{\{\tau_j \le t\}}\ell_j\delta^{-1}} \mid V = v\right] = 1 - \mathbb{P}\left[\tau_j \le t \mid V = v\right] + \mathbb{P}\left[\tau_j \le t \mid V = v\right]\phi_{\ell_j}(u, v)$$

where $\phi_{\ell_j}(u, v) := \mathbb{E}\left[e^{iu\ell_j\delta^{-1}} \mid V = v\right]$ denotes the V-conditional characteristic function of $\ell_j\delta^{-1}$. We finally apply the tower property

$$\begin{aligned} \phi_{L_t}(u) = \mathbb{E}\left[\mathbb{E}\left[e^{iuL_t\delta^{-1}} \mid V = v\right]\right] &= \int_{[0,1]^d} \mathbb{E}\left[e^{iuL_t\delta^{-1}} \mid V = v\right] dC_V(v) \\ = \int_{[0,1]^d} \left(1 - p_{j,t}(v) + p_{j,t}(v)\phi_{\ell_j}(u, v)\right) dC_V(v) \end{aligned}$$

where C_V is the density of X, and $p_{j,t}(v) = C_{U_j|V}(p_{j,t} \mid v)$.

Proof of Lemma 4.3.2

The proof is a straightforward application of discrete Fourier transform inversion. Observe that the random variable $L_t\delta^{-1}$ has state space $\{0, 1, \dots, M\}$. Its discrete Fourier transform is given by

$$F_m = \sum_{k=0}^{M} \mathbb{P}\left[L_t\delta^{-1} = k\right] \mathrm{e}^{-\mathrm{i}\frac{2\pi mk}{M+1}} = \phi_{L_t}\left(\frac{-2\pi m}{(M+1)}\right)$$

where ϕ_{L_t} as in Proposition 4.3.1 is the characteristic function of $L_t\delta^{-1}$. The probability mass function can be recovered as follows

$$\mathbb{P}\left[L_t = k\delta\right] = \frac{1}{M+1}\sum_{m=0}^{M} F_m \mathrm{e}^{\mathrm{i}\frac{2\pi mk}{M+1}}.$$

Equation (4.13) follows by observing that the signs can equivalently be switched between the complex weights.

Proof of Proposition 4.3.4

The proof of this proposition is immediate from the factor copula construction with proportional losses.

Proof of Corollary 4.3.5

This follows directly from Proposition 4.3.1 and Proposition 4.3.4.

Proof of Proposition 4.3.6

By construction we have

$$\begin{aligned} F_{x,y} := \phi_{N_t,L_t}(\mu x, \nu y) &= \mathbb{E}\left[\prod_{j=1}^{N} \exp\left\{\mathrm{i}\,\mathbb{1}_{\{\tau_j\le t\}}(\mu x + \nu y \ell_j \delta^{-1})\right\}\right] \\ &= \mathbb{E}\left[\exp\left\{\sum_{j=1}^{N} \mathrm{i}\,\mathbb{1}_{\{\tau_j\le t\}}(\mu x + \nu y \ell_j \delta^{-1})\right\}\right] = \mathbb{E}\left[e^{\mathrm{i}\mu x N_t + \mathrm{i}\nu y L_t \delta^{-1}}\right] \end{aligned}$$

Using this last expectation and the explicit expressions for μ and ν we obtain

$$F_{x,y} = \sum_{j=0}^{N}\sum_{k=0}^{M} \mathbb{P}\left[N_t = j, L_t = \delta k\right] e^{\mathrm{i}\frac{2\pi j}{N+1}x} e^{\mathrm{i}\frac{2\pi k}{M+1}y}.$$

This last expression is the two dimensional discrete Fourier transform of the density of the variable $(N_t, L_t\delta^{-1})$. The density can then immediately be retrieved by applying the inverse two-dimensional discrete Fourier transform inversion as follows

$$\mathbb{P}\left[N_t = j, L_t = \delta k\right] = \sum_{x=0}^{N}\sum_{y=0}^{M} F_{x,y}\, e^{-\mathrm{i}\frac{2\pi x}{N+1}j} e^{-\mathrm{i}\frac{2\pi y}{M+1}k}.$$

4.9 Conclusion

In this chapter we used factor copulas to construct flexible and tractable reduced form models for dependent default times. Using bivariate copulas as building blocks, we extend our framework from one-factor to multi-factor specifications, and we show that our approach nests most standard models as special cases. Furthermore, assuming the distribution of individual losses given default to be discrete, we propose a method to compute explicitly and efficiently the distribution of the portfolio loss. This allows us to price complex multi-name credit derivatives such as credit index swaptions, tranches on a portfolio of loans, and tranches on a portfolio of tranches.

We illustrate the versatility and computational efficiency of our approach with numerical examples. In particular, we investigate the impact on the portfolio loss distribution of different default dependence assumptions. We also examine how the loss distributions of credit derivatives, such as tranche and CDO squared, are affected. We calibrate multiple models to credit index tranche prices. We show that a particular specification achieve almost perfect calibration to all tranches using only two parameters that are stable over time.

References

Aas, K., C. Czado, A. Frigessi, and H. Bakken (2009, April). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics 44*(2), 182–198.

Abken, P. A., D. B. Madan, and B. S. Ramamurtie (1996). Estimation of risk-neutral and statistical densities by Hermite polynomial approximation: With an application to Eurodollar futures options. Working Paper 96-5, Federal Reserve Bank of Atlanta.

Ackerer, D. and D. Filipović (2016). Linear credit risk models. *Swiss Finance Institute Research Paper* (16-34).

Ackerer, D., D. Filipović, and S. Pulido (2016). The Jacobi stochastic volatility model. *Swiss Finance Institute Research Paper* (16-35).

Ackerer, D. and T. Vatter (2016). Dependent defaults and losses with factor copula models. *Swiss Finance Institute Research Paper* (16-59).

Ahdida, A. and A. Alfonsi (2013). A mean-reverting SDE on correlation matrices. *Stochastic Processes and their Applications 123*(4), 1472–1520.

Ait-Sahalia, Y. (2002). Maximum likelihood estimation of discretely sampled diffusions: A closed-form approximation approach. *Econometrica 70*(1), 223–262.

Al-Mohy, A. H. and N. J. Higham (2011). Computing the action of the matrix exponential, with an application to exponential integrators. *SIAM Journal on Scientific Computing 33*(2), 488–511.

Albrecher, H., P. Mayer, W. Schoutens, and J. Tistaert (2006). The little Heston trap. Technical report.

Altman, E., A. Resti, and A. Sironi (2004). Default recovery rates in credit risk modelling: A review of the literature and empirical evidence. *Economic Notes 33*(2), 183–208.

Amraoui, S. and S. G. Hitier (2008). Optimal stochastic recovery for base correlation. Technical report.

Andersen, L. and J. Sidenius (2004). Extensions to the Gaussian copula: Random recovery and random factor loadings. *Journal of Credit Risk 1*(1), 29–70.

Andersen, L., J. Sidenius, and S. Basu (2003). All your hedges in one basket. *Risk* (November), 67–72.

Andersen, L. B. and V. V. Piterbarg (2007). Moment explosions in stochastic volatility models. *Finance and Stochastics 11*(1), 29–50.

Backus, D. K., S. Foresi, and L. Wu (2004). Accounting for biases in Black-Scholes.

Bakshi, G. and D. Madan (2000). Spanning and derivative-security valuation. *Journal of Financial Economics 55*(2), 205–238.

Bedford, T. and R. M. Cooke (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence 32*, 245–268.

Bedford, T. and R. M. Cooke (2002). Vines - a new graphical model for dependent random variables. *Annals of Statistics 30*, 1031–1068.

Bernis, G. and S. Scotti (2017). Alternative to beta coefficients in the context of diffusions. *Quantitative Finance 17*(2), 275–288.

Bielecki, T. R., M. Jeanblanc, and M. Rutkowski (2006). Hedging of credit derivatives in models with totally unexpected default. *Stochastic Processes and Applications to Mathematical Finance*, 35–100.

Bielecki, T. R., M. Jeanblanc, and M. Rutkowski (2011). Hedging of a credit default swaption in the CIR default intensity model. *Finance and Stochastics 15*(3), 541–572.

Bielecki, T. R., M. Jeanblanc, M. Rutkowski, et al. (2008). Pricing and trading credit default swaps in a hazard process model. *The Annals of Applied Probability 18*(6), 2495–2529.

Bielecki, T. R. and M. Rutkowski (2002). *Credit Risk: Modeling, Valuation and Hedging*. Springer Science & Business Media.

Billingsley, P. (1995). *Probability and Measure*. Wiley Series in Probability and Statistics. John Wiley & Sons.

Black, F. and M. S. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy 81*(3), 637–654.

Brenner, M. and Y. Eom (1997). No-arbitrage option pricing: New evidence on the validity of the martingale property. NYU Working Paper No. FIN-98-009.

Brigo, D. and A. Alfonsi (2005). Credit default swap calibration and derivatives pricing with the SSRD stochastic intensity model. *Finance and Stochastics 9*(1), 29–42.

Brigo, D., A. Capponi, and A. Pallavicini (2014). Arbitrage-free bilateral counterparty risk valuation under collateralization and application to credit default swaps. *Mathematical Finance 24*(1), 125–146.

Brigo, D. and N. El-Bachir (2010). An exact formula for default swaptions' pricing in the SSRD stochastic intensity model. *Mathematical Finance 20*(3), 365–382.

Brigo, D. and M. Morini (2005). CDS market formulas and models. In *Proceedings of the 18th annual Warwick options conference*.

Brigo, D., A. Pallavicini, and R. Torresetti (2007). Calibration of CDO tranches with the dynamical generalized-Poisson loss model. Working Paper, Banca IMI, Milano.

Brigo, D., A. Pallavicini, and R. Torresetti (2010). *Credit Models and the Crisis: A Journey into CDOs, Copulas, Correlations and Dynamic Models.* John Wiley & Sons.

Broadie, M. and z. Kaya (2006). Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research 54*(2), 217–231.

Burtschell, X., J. Gregory, and J.-P. Laurent (2005). Beyond the Gaussian copula: Stochastic and local correlation. *Journal of Credit Risk.*

Burtschell, X., J. Gregory, and J.-P. Laurent (2009). A comparative analysis of CDO pricing models. *The Journal of Derivatives 16*(4), 9–37.

Carr, P. and D. Madan (1999). Option valuation using the fast Fourier transform. *Journal of Computational Finance 2*(4), 61–73.

Chen, H. and S. Joslin (2012). Generalized transform analysis of affine processes and applications in finance. *Review of Financial Studies 25*(7), 2225–2256.

Cheridito, P., D. Filipović, and R. L. Kimmel (2007). Market price of risk specifications for affine models: Theory and evidence. *Journal of Financial Economics 83*(1), 123 – 170.

Cheridito, P., D. Filipović, and M. Yor (2005). Equivalent and absolutely continuous measure changes for jump-diffusion processes. *The Annals of Applied Probability 15*(3), 1713–1732.

Corrado, C. J. and T. Su (1996). Skewness and kurtosis in S&P 500 index returns implied by option prices. *Journal of Financial research 19*(2), 175–192.

Corrado, C. J. and T. Su (1997). Implied volatility skews and stock index skewness and kurtosis implied by S&P 500 index option prices. *Journal of Derivatives 4*(4), 8–19.

Cousin, A. and J.-P. Laurent (2008). An overview of factor models for pricing CDO tranches. In *Frontiers In Quantitative Finance.*

Cuchiero, C., M. Keller-Ressel, and J. Teichmann (2012). Polynomial processes and their applications to mathematical finance. *Finance and Stochastics 16*(4), 711–740.

Delbaen, F. and H. Shirakawa (2002). An interest rate model with upper and lower bounds. *Asia-Pacific Financial Markets 9*(3-4), 191–209.

Demni, N. and M. Zani (2009). Large deviations for statistics of the Jacobi process. *Stochastic Processes and their Applications 119*(2), 518–533.

Di Graziano, G. and L. Rogers (2009). A dynamic approach to the modeling of correlation credit derivatives using markov chains. *International Journal of Theoretical and Applied Finance 12*(01), 45–62.

Drimus, G. G., C. Necula, and W. Farkas (2013). Closed form option pricing under generalized Hermite expansions.

Duarte, J. (2004, October). Evaluating an alternative risk preference in affine term structure models. *Review of Financial Studies 17*(2), 379–404.

Duffee, G. (2002). Term premia and interest rate forecasts in affine models. *The Journal of Finance 57*(1), 405–443.

Duffie, D., D. Filipović, and W. Schachermayer (2003). Affine processes and applications in finance. *Annals of Applied Probability 13*(3), 984–1053.

Duffie, D. and K. Singleton (1999). Modeling term structures of defaultable bonds. *Review of Financial Studies 12*(4), 687–720.

Duffie, D. and K. J. Singleton (2012). *Credit Risk: Pricing, Measurement, and Management.* Princeton University Press.

Dufresne, D. (2001). The integrated square-root process. Working Paper No. 90, Centre for Actuarial Studies, University of Melbourne.

Elliott, R. J., M. Jeanblanc, and M. Yor (2000). On models of default risk. *Mathematical Finance 10*(2), 179–195.

Erdélyi, A., W. Magnus, F. Oberhettinger, and F. G. Tricomi (1953). *Higher Transcendental Functions. Vols. I, II.* McGraw-Hill.

Eriksson, B. and M. Pistorius (2011). Method of moments approach to pricing double barrier contracts in polynomial jump-diffusion models. *International Journal of Theoretical and Applied Finance 14*(7), 1139–1158.

Ethier, S. N. and T. G. Kurtz (1986). *Markov Processes : Characterization and Convergence.* Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons.

Fang, F. and C. W. Oosterlee (2009). A novel pricing method for European options based on Fourier-cosine series expansions. *SIAM Journal on Scientific Computing 31*(2), 826–848.

Feller, V. (1960). *An Introduction to Probability Theory and Its Applications: Volume 1.* John Wiley & Sons.

Filipović, D. (2009). *Term-Structure Models: A Graduate Course.* Springer Finance.

Filipović, D. and M. Larsson (2016). Polynomial diffusions and applications in finance. *Finance and Stochastics 20*(4), 931–972.

Filipović, D. and M. Larsson (2017). Polynomial jump-diffusion models. *Working Paper.*

Filipović, D., M. Larsson, and A. B. Trolle (2017). Linear-rational term structure models. *The Journal of Finance 72*(2), 655–704.

Filipović, D., E. Mayerhofer, and P. Schneider (2013). Density approximations for multivariate affine jump-diffusion processes. *Journal of Econometrics 176*(2), 93–111.

Filipović, D., L. Overbeck, and T. Schmidt (2011). Dynamic CDO term structure modeling. *Mathematical Finance 21*(1), 53–71.

Fischer, B. and G. H. Golub (1992). How to generate unknown orthogonal polynomials out of known orthogonal polynomials. *Journal of Computational and Applied Mathematics 43*(1-2), 99–115.

Fouque, J.-P., R. Sircar, and K. Sølna (2009). Multiname and multiscale default modeling. *Multiscale Modeling & Simulation 7*(4), 1956–1978.

Gabaix, X. (2009). Linearity-generating processes: A modelling tool yielding closed forms for asset prices. Working Paper 13430, NBER.

Gaß, M., K. Glau, M. Mahlstedt, and M. Mair (2015). Chebyshev interpolation for parametric option pricing.

Giesecke, K. (2008). *Portfolio credit risk: Top-down versus bottom-up approaches*, pp. 251–267. John Wiley & Sons.

Glasserman, P. and K.-K. Kim (2011). Gamma expansion of the Heston stochastic volatility model. *Finance and Stochastics 15*(2), 267–296.

Glasserman, P. and S. Suchintabandid (2012). Quadratic transform approximation for CDO pricing in multifactor models. *SIAM Journal on Financial Mathematics 3*(1), 137–162.

Gourieroux, C. and J. Jasiak (2006). Multivariate Jacobi process with application to smooth transitions. *Journal of Econometrics 131*(1), 475–505.

Gregory, J. and J.-P. Laurent (2003). I will survive. *Risk* (June), 103–107.

Guillaume, F., P. Jacobs, and W. Schoutens (2009). Pricing and hedging of CDO-squared tranches by using a one factor Lévy model. *International Journal of Theoretical and Applied Finance 12*(05), 663–685.

Herbertsson, A. (2008). Pricing synthetic CDO tranches in a model with default contagion using the matrix-analytic approach. *Journal of Credit Risk 4*(4), 3–35.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies 6*(2), 327–343.

Heston, S. L. and A. G. Rossi (2016). A spanning series approach to options. *The Review of Asset Pricing Studies 7*(1), 2–42.

Higham, N. J. (2008). *Functions of Matrices: Theory and Computation.* SIAM.

Hochbruck, M. and C. Lubich (1997). On Krylov subspace approximations to the matrix exponential operator. *SIAM Journal on Numerical Analysis 34*(5), 1911–1925.

Hofert, M. and M. Scherer (2011). CDO pricing with nested archimedean copulas. *Quantitative Finance 11*(5), 775–787.

Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. *The Journal of Finance 42*(2), 281–300.

Hull, J. and A. White (2004). Valuation of a CDO and an n-th to default CDS without Monte Carlo simulation. *The Journal of Derivatives 12*(2), 8–23.

Hull, J. and A. White (2010). The risk of tranches created from mortgages. *Financial Analysts Journal 66*(5), 54–67.

Hull, J. C. and A. D. White (2003). The valuation of credit default swap options. *The Journal of Derivatives 10*(3), 40–50.

Hull, J. C. and A. D. White (2006). Valuing credit derivatives using an implied copula approach. *The Journal of Derivatives 14*(2), 8–28.

Jäckel, P. (2005). A note on multivariate Gauss-Hermite quadrature. Technical report.

Jacod, J. and A. N. Shiryaev (2013). *Limit Theorems for Stochastic Processes.* Springer Science & Business Media.

Jacquier, A. and P. Roome (2015). Asymptotics of forward implied volatility. *SIAM Journal on Financial Mathematics 6*(1), 307–351.

Jamshidian, F. (2004). Valuation of credit default swaps and swaptions. *Finance and Stochastics 8*(3), 343–371.

Jarrow, R. and A. Rudd (1982). Approximate option valuation for arbitrary stochastic processes. *Journal of Financial Economics 10*(3), 347–369.

Joe, H. (1996). Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. In L. Rüschendorf, B. Schweizer, and M. Taylor (Eds.), *Distributions with Fixed Marginals and Related Topics.*

Kahl, C. and P. Jäckel (2005). Not-so-complex logarithms in the Heston model. *Wilmott magazine 19*(9), 94–103.

Kahl, C. and P. Jäckel (2006). Fast strong approximation Monte Carlo schemes for stochastic volatility models. *Quantitative Finance 6*(6), 513–536.

Kalemanova, A., B. Schmid, and R. Werner (2007). The normal inverse Gaussian distribution for synthetic CDO pricing. *The Journal of Derivatives.*

Karatzas, I. and S. E. Shreve (1991). *Brownian Motion and Stochastic Calculus.* Graduate Texts in Mathematics. Springer.

Karlin, S. and H. Taylor (1981). *A Second Course in Stochastic Processes.* Academic Press.

Krekel, M. (2008). Pricing distressed CDOs with base correlation and stochastic recovery.

Krupskii, P. and H. Joe (2013). Factor copula models for multivariate data. *Journal of Multivariate Analysis 120,* 85 – 101.

Krupskii, P. and H. Joe (2015). Structured factor copula models: Theory, inference and computation. *Journal of Multivariate Analysis 138,* 53–73.

Kruse, S. and U. Nögel (2005). On the pricing of forward starting options in Heston's model on stochastic volatility. *Finance and Stochastics 9*(2), 233–250.

Küchler, U. and S. Tappe (2008). Bilateral gamma distributions and processes in financial mathematics. *Stochastic Processes and their Applications 118*(2), 261–283.

Lando, D. (1998). On Cox processes and credit risky securities. *Review of Derivatives Research 120,* 99–120.

Lando, D. (2009). *Credit Risk Modeling: Theory and Applications.* Princeton University Press.

Larsson, M. and S. Pulido (2017). Polynomial preserving diffusions on compact quadric sets. *Stochastic Processes and their Applications 127*(3), 901–926.

Laurent, J.-P. and J. Gregory (2005). Basket default swaps, CDOs and factor copulas. *The Journal of Risk 7*(4), 1.

Li, D. X. (2000). On default correlation: A copula function approach. *The Journal of Fixed Income 9*(4), 43–54.

Li, D. X. and M. H. Liang (2005). CDO squared pricing using Gaussian mixture model with transformation of loss distribution.

Li, H. and A. Melnikov (2012). On Polynomial-Normal model and option pricing. In *Stochastic Processes, Finance and Control: A Festschrift in Honor of Robert J Elliott*, Advances in Statistics, Probability and Actuarial Science, Chapter 12, pp. 285–302. Singapore: World Scientific Publishing Company.

Lipton, A. and A. Sepp (2008). Stochastic volatility models and Kelvin waves. *Journal of Physics A: Mathematical and Theoretical 41*(34), 344012.

Longstaff, F. (1995). Option pricing and the martingale restriction. *Review of Financial Studies 8*(4), 1091–1124.

Madan, D. B. and F. Milne (1994). Contingent claims valued and hedged by pricing and investing in a basis. *Mathematical Finance 4*(3), 223–245.

Mazet, O. (1997). Classification des semi-groupes de diffusion sur $\mathbb{R}$ associés à une famille de polynômes orthogonaux. In *Séminaire de Probabilités XXXI*, Volume 1655 of *Lecture Notes in Mathematics*, pp. 40–53. Berlin: Springer.

McNeil, A. J., R. Frey, and P. Embrechts (2015). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press.

Mortensen, A. (2006). Semi-analytical valuation of basket credit derivatives in intensity-based models. *Journal of Derivatives 13*(4), 8–26.

Mysovskikh, I. (1968). On the construction of cubature formulas with fewest nodes. In *Dokl. Akad. Nauk SSSR*, Volume 178, pp. 1252–1254.

Necula, C., G. G. Drimus, and W. Farkas (2015). A general closed form option pricing formula. Swiss Finance Institute Research Paper No. 15-53.

Nelsen, R. B. (1999). *An Introduction to Copulas*. Springer.

Oh, D. H. and A. J. Patton (2013). Simulated method of moments estimation for copula-based multivariate models. *Journal of the American Statistical Association 108*(502), 689–700.

Oh, D. H. and A. J. Patton (2017). Modeling dependence in high dimensions with factor copulas. *Journal of Business & Economic Statistics 35*(1), 139–154.

Pagès, G. and J. Printems (2003). Optimal quadratic quantization for numerics: The Gaussian case. *Monte Carlo Methods and Applications 9*(2), 135–165.

Revuz, D. and M. Yor (1999). *Continuous Martingales and Brownian Motion*. Grundlehren der mathematischen Wissenschaften. Springer.

Rogers, L. and D. Williams (2000). *Diffusions, Markov Processes, and Martingales: Volume 1, Foundations*. Cambridge Mathematical Library. Cambridge University Press.

Rogers, L. C. G. and Z. Shi (1995). The value of an Asian option. *Journal of Applied Probability 32*(4), 1077–1088.

Rudin, W. (1974). *Functional Analysis*. McGraw-Hill.

Sancetta, A. and S. Satchell (2004). The Bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econometric Theory 20*(03), 535–562.

Sato, K.-i. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.

Schepsmeier, U. and J. Stöber (2014, May). Derivatives and Fisher information of bivariate copulas. *Statistical Papers 55*(2), 525–542.

Schöbel, R. and J. Zhu (1999). Stochastic volatility with an Ornstein–Uhlenbeck process: An extension. *Review of Finance 3*(1), 23–46.

Schönbucher, P. (2000). A Libor market model with default risk.

Schönbucher, P. (2004, August). A measure of survival. *Risk*, 79–85.

Schönbucher, P. J. and D. Schubert (2001). Copula-dependent default risk in intensity models. Working paper, Department of Statistics, Bonn University.

Schoutens, W. (2012). *Stochastic Processes and Orthogonal Polynomials*, Volume 146. Springer Science & Business Media.

Sidje, R. B. (1998). Expokit: A software package for computing matrix exponentials. *ACM Transactions on Mathematical Software (TOMS) 24*(1), 130–156.

Stein, E. M. and J. C. Stein (1991). Stock price distributions with stochastic volatility: An analytic approach. *The Review of Financial Studies 4*(4), 727–752.

Wilcox, R. (1967). Exponential operators and parameter differentiation in quantum physics. *Journal of Mathematical Physics 8*(4), 962–982.

Xiu, D. (2014). Hermite polynomial based expansion of European option prices. *Journal of Econometrics 179*(2), 158–177.

Yamada, T. and S. Watanabe (1971). On the uniqueness of solutions of stochastic differential equations. *Journal of Mathematics of Kyoto University 11*(1), 155–167.

Yor, M. (2001). Bessel processes, Asian options, and perpetuities. In *Exponential Functionals of Brownian Motion and Related Processes*, pp. 63–92. Springer.

Curriculum Vitae

Damien Ackerer

Born on September 19th 1987

Education

2012–2017	Ph.D. in Finance, EPFL and Swiss Finance Institute
Fall 2016	Visiting Scholar, Columbia University in the City of New York
2009–2011	M.Sc. in Financial Engineering, EPFL
2006–2009	B.Sc. in Mathematics, Université Pierre et Marie Curie in Paris

Professional and Teaching Experiences

2013–2017	Teaching Assistant, EPFL
2011–2012	Research Engineer, Swissquote Bank in Geneva
Spring 2011	Intern, Swissquote Bank in Geneva

Working Papers

- The Jacobi Stochastic Volatility Model, 2016, with Damir Filipović and Sergio Pulido
- Linear Credit Risk Models, 2016, with Damir Filipović
- Dependent Defaults and Losses with Factor Copula Models, 2016, with Thibault Vatter

Conference Presentations

2017	Actuarial and Financial Mathematics conference, Brussels; Bernoulli Center workshop on Dynamical Models in Finance, Lausanne.
2016	Bachelier Society world congress, New-York; European Finance Association annual meeting, Oslo; University of Mannheim Mathematics seminar; ICMS workshop on Quantitative Finance, Edinburgh; AFFI Paris Finance meeting.
2015	7th General AMaMeF and Swissquote conference, Lausanne; CEQURA junior research workshop, Münich.