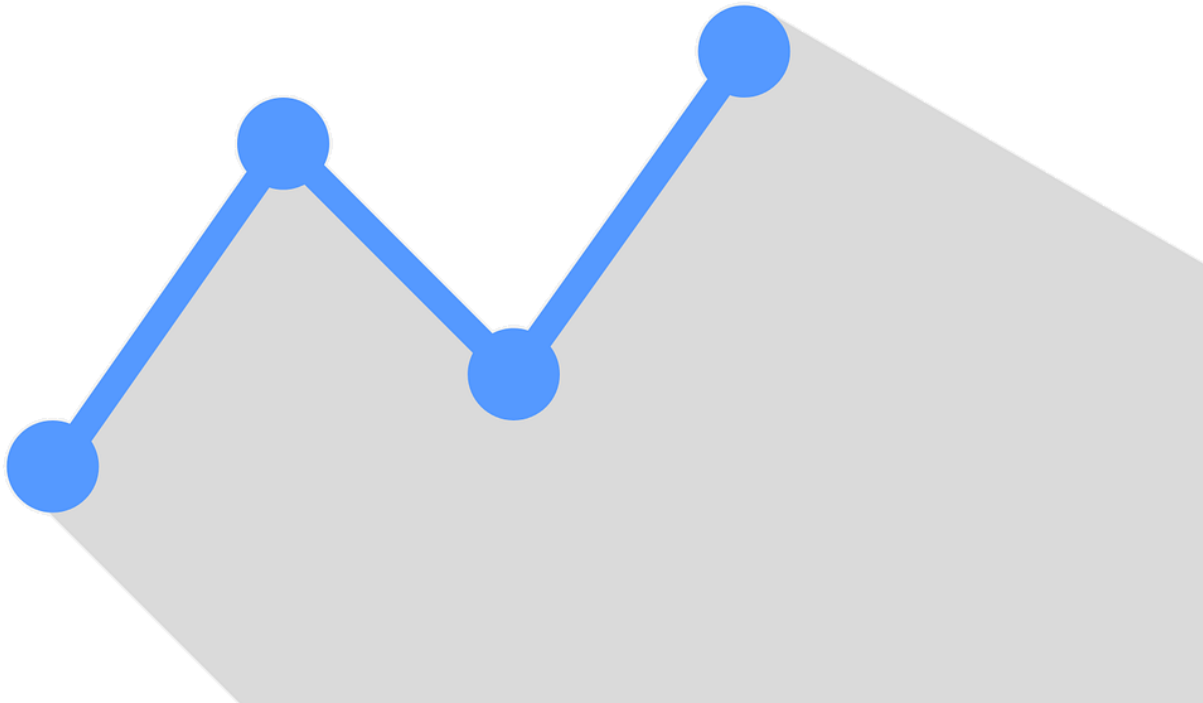
A Fourier-based approach for valuing bonds with embedded options

an analysis of convolution- and cosine-based approximation methods using jump-diffusion models and extended with stochastic interest rates

R. Roest

TU Delft

TRIPLE A RISK FINANCE



A Fourier-based approach for valuing bonds with embedded options

an analysis of convolution- and cosine-based approximation methods using jump-diffusion models and extended with stochastic interest rates

by

R. Roest

to obtain the degree of Master of Science

at the Delft University of Technology,

to be defended publicly on Thursday, October 13, 2022, at 14:00.

Student number:	4299086	
Project duration:	September 1, 2021 – August 31, 2022	
Thesis committee:	dr. ir. L. A. Grzelak,	University of Utrecht, daily supervisor
	Prof. dr. ir. C. L. Vuik,	TU Delft, supervisor
	Prof. dr. A. Papapantoleon,	TU Delft, head of chair Financial Engineering

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TU Delft

Abstract

Due to their attractive characteristics, convertible and callable bonds became a more important class of fixed income products within the financial market [5]. Therefore, the need for fair and accurate pricing of convertible and callable bonds increases. Where the convertible option can be considered as a right for the bondholder, the callable option is a right assigned to the bond issuer. Moreover, due to the hybrid nature of a convertible bond, it both contains characteristics of fixed income and equity products. As a consequence, both the risk of default and the possible equity profits need to be taken into account when valuing convertible bonds [7]. Furthermore, as many convertible bonds also include a call option, an extra early exercise feature due to this call option needs to be taken into account. This early exercise option further increases the complexity of the valuation problem, as a possible buyback of the bond before maturity needs to be considered in the value of the bond [3]. Computing a fair and accurate price for convertible and callable bonds, therefore gives rise to a complex valuation problem which leads to the need for further research.

In this thesis, a structural default model is used to value callable convertible bonds. Contrary to reduced-form models, structural default models are characterized by the rationale behind the default event [3]. Due to the better rationale and new insights into possible better market fits, structural default models have gained new interest in academic research [3, 15, 34]. Recent research is conducted on finding new numerical techniques used for approximating the prices of complex financial products [27]. This thesis discussed different valuation methods as proposed by Longstaff and Schwartz, Lord et al. and Oosterlee and Grzelak applied to the valuation of callable convertible bond. In particular, a convolution-based method (CONV) and a cosine-based (COS) method are discussed. Whereas the Monte Carlo methods and the CONV method are already used in different articles, the COS method was not yet applied to the callable convertible bond valuation [3, 1]. As the COS method has been proven to be an efficient algorithm for approximating the values of financial derivatives, this thesis uses the COS method to compare its convergence to seek more insights into the convergence of the CONV method. Monte Carlo methods are used to verify the obtained approximations.

For the problem considered under constant interest rates, this thesis shows that the COS method can be derived and applied. For small values of the grid size, the COS method showed to converge much faster than the CONV method. For larger values of the grid size, the CONV method showed to catch up with the COS method to become almost equally accurate. The results also showed that, contrary to the COS method, the CONV method was robust under the choice of the hyper-parameter concerning the integration grid. For the COS method, it was shown that a bad choice of the hyper-parameter could lead to a bad approximation. Although the two-dimensional CONV method shows to converge to the value obtained under Monte Carlo simulation, the results are still off for grid sizes of intermediate size. Results seem to indicate that the amount of grid points is not sufficient for the proposed integration interval and that therefore more grid points are needed. A greater amount of grid points, however, will also indicate a requirement for a greater amount of resources which may not always be available.

From the results of the zero-coupon case, the COS method showed a more rapid convergence towards the approximations obtained by the Monte Carlo methods than the approximations obtained using the CONV method. Furthermore, when taking the hyper-parameters into account, the CONV method showed less robust features than the COS method in the two-dimensional case. Only for a small range of the hyper-parameters convergence is obtained for the CONV method. On the other hand, the COS method clearly shows convergence for a much wider range of hyper-parameters.

Preface

This thesis has been submitted for the degree of Master of Science in Applied Mathematics from the Delft University of Technology. The thesis will discuss Fourier-based methods for the valuation of callable convertible bonds and is conducted in cooperation with Triple-A risk finance, which allowed me to experience the practical aspects of the financial industry.

I would like to thank my daily supervisor Dr.ir. L.A. Grzelak for his guidance and support in the development of this thesis. His feedback and ideas helped to improve my work on the subject and I am glad to have the opportunity to learn much of his knowledge on the subject of risk-neutral pricing. The exceptionally good knowledge of the industry and specific subject, helped me to get new insights and inspiration to tackle the challenges that came while conducting this research. Furthermore, I would like to thank Paul Kemper. His practical insights into the financial industry helped me to get a more solid view of the workings of the financial industry. Moreover, his positivity helped me to obtain much satisfaction from working on the subject. Additionally, I would like to thank Prof. dr. ir. C. L. Vuik and Prof. dr. A. Papapantoleon for being part of my thesis committee. I also like to thank the other colleagues at Triple-A, my friends, and my family for supporting me during this period.

R. Roest
Schiedam, August 2022

Contents

1 Introduction

1.1. Motivation and contribution

One way of financing investments is by issuing bonds. A bond is a contractual agreement between the bond issuer and the bondholder upon future payments from the issuer to the bondholder. At the issuance of a bond, the issuer receives the bond price which can be used to finance investments. At the maturity of the bond, the issuer pays the face value of the bond to the bondholder. Within financial markets, different parties, issue bonds. In general, these parties can be distinguished into two categories; governments and companies [32]. Bonds that are issued by companies are referred to as corporate bonds. Contrary to bonds issued by major stable governments, corporate bonds are more sensitive to credit risk [25]. When a company is not able to fulfill the obliged payments, the company is said to go into default. As a consequence, the bondholder will not receive the agreed payment and therefore miss out on the contractually agreed fixed income. To compensate for this default risk, corporate bonds often have higher returns than government bonds, which makes the corporate bond an interesting fixed-income product for investors.

Corporate bonds can be embedded with a convertible option, which enables the bondholder to exchange the bond for shares of the company's stock. Contrary to ordinary stocks, convertible bonds have a more stable income for the bondholder [37]. Moreover, in comparison to a plain corporate bond, convertible bonds still yield the possibility to profit from an increased performance of the company [7]. When a company performs well, the bondholder of a convertible bond may exercise the conversion right to profit from high share prices. Issuing convertible debt may also be attractive for the issuer of the bond. As the convertible option is equity for the bondholder, the costs for issuing a convertible bond are less than the costs for issuing a corporate bond without any options [4].

Convertible and corporate bonds may also be embedded with a callable option. Callable options provide a safety mechanism against interest rate changes for the issuer of the bond. The holder of a callable option owns the right to buy back the bond for a predetermined price. When interest rates decline, the issuer of the bond can exercise its right to buy back the bond and issue new bonds against a more favorable interest rate [5]. Moreover, bond yields are presumed to depend on the credit risk of the issuing company [29]. As the company acquires a higher credit score, it may be able to issue new debt against a lower yield. Hence, it may be favorable for the issuing company to buy back the bond and issue new debt against the lower yield [5]. As the callable option is a right that is assigned to the issuer of the bonds, the exercise right of the option is often represented in a higher return and a lower market price [13]. In particular, the high yields with lower market prices make callable bonds interesting for investors.

Due to their attractive characteristics, convertible and callable bonds became a more important class of fixed income products within the financial market [5]. Therefore, the need for fair and accurate pricing of convertible and callable bonds increases. Where the convertible option can be considered as a right for the bondholder, the callable option is a right assigned to the bond issuer. Moreover, due to the hybrid nature of a convertible bond, it both contains characteristics of fixed income and equity

products. As a consequence, both the risk of default and the possible equity profits need to be taken into account when valuing convertible bonds [7]. Furthermore, as many convertible bonds also include a call option, an extra early exercise feature due to this call option needs to be taken into account. This early exercise option further increases the complexity of the valuation problem, as a possible buyback of the bond before maturity needs to be considered in the value of the bond [3]. Computing a fair and accurate price for convertible and callable bonds, therefore gives rise to a complex valuation problem which leads to the need for further research.

1.1.1. Structural default models

Different academical approaches are proposed for valuing bonds with embedded options [12, 36, 9, 3, 21, 15, 29, 2]. The proposed models for modeling the default risk within financial products can be divided into two classes of models; structural default models and reduced-form default models. Within reduced-form models, it is assumed that the default event is driven by a stochastic process with a corresponding intensity rate. By modeling the default risk as an exogenous event that arrives by a predefined intensity (which may be stochastic), a credit spread can be derived [12]. This credit spread can be used as a discount factor for valuing a specific financial product subject to default risk [2]. The credit spread corresponding to the default risk can easily be obtained from market data. As a consequence, reduced-form models are popular as their calibration is relatively straightforward [3]. However, the exogenous approach of reduced-form models forms a less clear rationale for the default event, as it does not model the default event directly but uses a probability-based approach on internal and external events to determine a possible default event.

Contrary to reduced-form models, structural default models are characterized by the rationale behind the default event [3]. Structural default models assume that the default of a company is endogenously determined by the financial obligations of the company [4]. As the company can be defined in terms of its capital structure and its limited liability, structural models assume that the default event will appear when the company has insufficient assets to pay its liabilities. To model the financial obligations of the company, within structural models, not the share price but the value of the company is modeled by a stochastic process [21]. The market value of a company is defined by the capital structure of that company, which gives the relation between the equity and debt ratio that is used to finance the company [31]. For each financial obligation, the value of the company is compared to the financial obligations of the company to verify whether the company can fulfill the payments. When the capital is not sufficient, the company goes into default and is assumed no longer to be able to fulfill any payments.

Structural models show two major drawbacks. In the first place, structural models depend on the dynamics of the company value [3], which is not a publicly tradeable quantity and therefore cannot be observed in the market. As a consequence, structural models can not directly be calibrated on historical data. Secondly, more basic structural default models yield a bad approximation of credit spreads [34]. From market data, it can be seen that corporate bonds often trade higher yields compared to government bonds [20]. The difference between the higher yields for corporate bonds contrary to the lower yields for government bonds is often thought of as credit-risk compensation. As the investor buys a corporate bond instead of a government bond, she exposes herself to the default risk of the company issuing the bond [5]. The higher yields of corporate bonds form compensation for the default risk that comes with the corporate bond [5]. The part of the yield that is considered to be compensation is called the credit spread. Contrary to reduced-form models, the credit spread is not directly modeled by structural models, but, can be derived within the model. Huang and Huang showed that when basic models are considered, such as proposed by Merton, the derived credit spread compared to the actual credit spreads obtained from the market is off [20, 29].

Although basic models show bad approximations for the credit spreads, more advanced models tend to model the credit spreads better [3]. As general research within the modeling of financial assets evolved, more advanced models were invented. Recent research showed that using more advanced models within the structural models yields a better fit on the market data and, therefore, yields better credit spreads [15]. For instance, Fang et al. use a Lévy-process to model the value of the company when trying to numerically derive a price for CDSs. Fang et al. show that the obtained results closely resemble the market quotes and that the default probabilities obtained are following the market spreads.

On the calibration of structural default models
As the value of the company is not a publicly tradeable quantity and cannot be observed within the market, the calibration of structural default models is more difficult [4]. Different methods for the calibration of structural default models are proposed, which include calibration via inversion techniques, solving minimization problems, and bootstrapping techniques [6, 15, 16]. Though there exist calibration methods, they all depend on the availability of quoted credit default swaps (CDSs). The dependency on CDSs limits the use of structural default models to the selection of companies that have this information available, which is a limiting condition on the proposed model. To only partly rely on CDS spreads, Forte proposed an algorithm that is based both on the market capitalization of the equity of the company and the book value of the company and yields a proposal to obtain the value process of the company via an iterating process [16].

A calibration method that is in line with the given approach in this thesis is that proposed by Fang et al. [15]. Fang et al. propose to redefine the pricing algorithm to a minimization problem that can be numerically solved. The resulting method yields implied parameters for the structural default model obtained from market quotes of the CDS market. The minimization problem concerns retrieving the parameter set for which the difference between the market quotes of the CDSs and the risk-neutral value of the CDS is below a certain tolerance. The method relies on fast and efficient computations, which can be done by the cosine (COS) approximation method proposed. One parameter that is not obtained from this method, is the initial company value. However, this parameter can be obtained via the market value of the equity and the book values of the debt as proposed by Forte.

If the valuation is conducted at the issuance of the bond and did not issue any other bonds, a more naive approach can be considered. In the financial markets, it is common to trade bonds at part, that is selling the bond at its face value or very close to its face value. To be able to sell the bond at its face value, the features of the bond are set in such a way that the price investors are prepared to pay is equal to the face value. Hence, in this case, the initial value of the company is equal to the sum of the market capitalization of the equity and the face value of the bond. The other parameters can then be calibrated to a similar company that does trade CDSs.

1.1.2. Numerical valuation methods

Due to the better rationale and new insights into possible better market fits, structural default models have gained new interest in academic research [3, 15, 34]. More advanced models, such as jump-diffusion models, show much more interesting results and are therefore very useful within structural default models [15]. Moreover, recent research is conducted on finding new numerical techniques used for approximating the prices of complex financial products [27]. Two of the most commonly used classes of numerical techniques for the valuation of financial products are; Monte Carlo (MC) techniques and Fourier approximation.

MC techniques depend on simulation to realize a great number of simulated paths of the underlying asset. Monte Carlo paths contain realizations of the stochastic variable for each time step. These realizations can then be used to approximate the value of a financial product by computing the corresponding payoffs determined by the product. In particular, MC methods can be used to compute premature financial option values which makes the methods widely applicable. For example, Longstaff and Schwartz have introduced an algorithm that can be used for valuing early exercise options by use of MC simulation and least-squares regression [26]. In general, MC techniques are interesting because of their wide applicability and because they avoid the curse of dimensionality [32]. Some earlier used techniques, such as numerical partial differential solvers, show deficiencies when used on models that assume multiple stochastic factors. When systems of higher dimensional stochastic factors are computed, the number of integration points increases exponentially [24]. The exponential growth of integration points due to higher dimensional systems is often referred to as the curse of dimensionality [32].

Alternatively, Fourier techniques form a popular class of numerical valuation methods. Fourier methods exploit the phenomenon that some computations may become very complex in asset space but can be done easier in Fourier space [27]. Moreover, Fourier techniques are often very efficient approximation techniques, which makes them computationally fast [10]. To be able to perform Fourier methods, the characteristic function of the system is required. As models for asset classes may become more

complicated, their characteristic functions can often still be computed [32]. In particular, models that are included in the affine class have a characteristic function that frequently can be computed in closed form. Many popular advanced models are included within the class of affine models. The availability of the characteristic functions combined with efficient computational methods makes Fourier techniques attractive for financial computations.

Although Fourier methods are highly efficient, they are less widely applicable and less tractable compared to MC methods. When pricing more complex options with early exercise features, MC methods become more interesting [32]. On the other hand, Fourier methods are more attractive when fast, highly efficient computations are required. For instance, one way of calibrating a model is by optimizing the parameters of the model based on market quotes [15]. The obtained parameter estimations that follow from solving the optimization problem are referred to as the implied parameters. Often, the implied parameters are obtained numerically. As solving these numerical optimization problems to find the implied parameters may require fast and efficient computations, Fourier methods become very attractive for these calibration methods.

1.1.3. Contribution

Within this thesis, the COS method will be applied to the valuation problem of callable convertible bonds. The COS method has been proven useful in the valuation of different classes of financial derivatives. In particular, the COS method has been used for the valuation of both European as Bermudan types of derivatives. However, the COS method has not been applied to fixed-income products with a hybrid nature and multiple early exercise features. Within this thesis, a cosine-based algorithm will be derived to efficiently value the callable convertible bond under constant interest rates. Moreover, an application of the two-dimensional COS method will be presented to the non-coupon-paying convertible bond pricing problem. When the coupon payments are assumed to be constant, the convertible pricing problem where no early call optionality is included will be very similar to a European-type option valuation problem. The European type of valuation can be computed efficiently by the two-dimensional COS method. A derivation of this algorithm will be given in the thesis.

Furthermore, using the derived COS method for the constant interest rate case and the stochastic interest rate case, a convolution-based (CONV) algorithm as presented by Ballotta and Kyriakou will be analyzed. The CONV method presented by Ballotta and Kyriakou shows the potential to be an efficient algorithm for computing bond values. In the article presented by Ballotta and Kyriakou, only a brief discussion on the convergence of the CONV method is given, where the results are, among others, compared to the log-normal analytical results obtained by Ingersoll Jr [21]. In this thesis, a more elaborate discussion of convergence will be given. This discussion will include a comparison with the COS method for constant interest rates, a discussion on the hyper-parameters, and a discussion of the convergence under stochastic interest rates where the COS algorithm was used for the valuation of a non-coupon paying convertible bond.

As it is known that the CONV method does have a slower convergence than the COS method [14]. In this thesis, the convergence of the CONV method will be discussed based on numerically obtained results. As no closed-form solution for the value of a callable convertible bond with coupon payments under a structural default model is known, robust Monte-Carlo methods will be used to obtain a reference value. Using the obtained reference value, the COS and the CONV method will be tested for accuracy. Furthermore, the hyper-parameters of the methods will be considered, to obtain insights into the sensitivity of the hyper-parameters to the convergence of the method.

When efficient and fast computations methods are developed, the limited calibration condition for the structural default models on the availability of CDSs can be relaxed. A highly efficient Fourier method for computing bond values under structural default models will raise the possibility to further research on the application of Fourier methods in the calibration of structural default models based on bond data instead of CDS data. Being able to also incorporate bond data in the calibration process, will enhance the class of companies to which structural models can be applied to.

1.1.4. Thesis structure

The thesis will attain the following structure. This chapter will continue with a more elaborate discussion on corporate, convertible, and callable bonds. Section 1.2 will discuss the payoff structure and the early

exercise features of the (callable) convertible bond. In particular, also the optimal exercise policy of the convertible and callable feature will be discussed. Chapter 2 will be focused on the modeling of the risk factors. The risk factors both include the default risk, modeled by a structural default model, and interest rate risk modeled under the Hull-White process. The Hull-White model will be obtained under the HJM framework, which enables a direct fit on the current yield curve. The model proposed is included in the affine class of diffusion models, which enables the derivation of the characteristic function [32]. Both Monte Carlo and Fourier approaches will be discussed. In the third chapter, Monte-Carlo methods will be discussed for obtaining reference values that can be used in the convergence analysis.

Chapter 4 will be discussing the Fourier-based approaches. In particular, the CONV and the COS method will be discussed for valuing the callable convertible bond under the proposed jump-diffusion models. When constant interest rates are considered, both a CONV and COS algorithm will be discussed. In the case of stochastic interest rates, a discussion on the valuation problem of the callable convertible bond will be given based on the CONV method. When the pricing problem is reduced to a non-coupon paying convertible bond, also the COS method will be discussed under stochastic interest rates. The discussions given in Chapter 3 and Chapter 4 will be first given a valuation algorithm based on the less complex convertible bond, and then extend the algorithm to the more complex callable convertible bond. In Chapter 5, the numerical results will be discussed. This chapter will be elaborating on the convergence of both the CONV and the COS method concerning the different cases discussed. The last chapter will draw conclusions based on the numerical results and give suggestions for further research.

1.2. Characteristics corporate, convertible and callable bonds

A bond (both government and corporate) is a contractual agreement between two parties on future payments defined on a time horizon $[t_0, T]$, for $T \in \mathbb{R}_+$ and $0 \leq t_0 < T$. At issuance, the bond is sold on the financial market and the money that is raised can be used to finance new investments. By buying a bond, an investor tries to ensure a fixed income. At maturity T, the face value F will be paid to the bondholder and the bond will be terminated. One of the building blocks for financial derivatives is the risk-free zero-coupon bond. Risk-free zero-coupon bonds are a specific class of bonds, which surely pay a face value of $F = 1$ at maturity. This kind of bond will be referred to as the zero-coupon bond.

Definition 1.1. A zero-coupon bond is a risk-free bond with a face value of $F = 1$, which is paid at maturity [32]. The value of a zero-coupon bond with maturity T at time t is denoted by $P(t, T)$. Zero-coupon bonds do not yield any payments to the bondholder before maturity T.

The payoff scheme of a zero-coupon bond is composed of one payment at maturity T, as given in Figure 1.1.

Figure 1.1: Payment scheme of a zero-coupon bond. A zero-coupon bond has a sure payoff of $F = 1$ at maturity T.

Besides the payment of the face value F at maturity T, a bond may also provide intermediate payments, which are called coupon payments. The coupon payments are paid to the bondholder at a finite set $\mathcal{T}^{\text{coup}} = \{\tau_1, \ldots, \tau_{n-1}\} \subseteq [t_0, T]$ of $n - 1$ coupon dates. The payoff scheme of a fixed coupon bond is given in Figure 1.2.

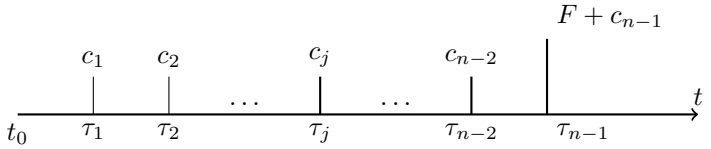


Figure 1.2: Schematic payoff of a bond paying $n-1$ fixed coupon payments at coupon payment dates $\mathcal{T}^{\text{coup}} = \{t_1, \dots t_{n-1}\}$. At maturity, T, both the last coupon payment as the face value are paid to the bondholder.

Coupon payments are often defined in terms of a percentage of the face value F of the bond, which is called the coupon rate. Coupon rates can be fixed or be based on market interest rates. Assume that the coupon rate of the bond is fixed and defined as $\alpha \in [0, 1]$, then each coupon payment can be written as

$$c_j = \alpha F,$$

for all $j \in \{1, \dots, n-1\}$. Hence, the final payment of the bond, in case the bond is not terminated before maturity, will be $F(1+\alpha)$.

1.2.1. Corporate bonds

Corporate bonds are subject to default risk [21]. The issuer of the bond may not be able to pay the liabilities by the bond. If the issuer is not able to pay any of the coupon payments at any coupon payment date $\tau_j \in \mathcal{T}^{\text{coup}}$ for $j \in \{1, \dots n-1\}$ or is not able to fulfill the payment of the face value at maturity, the issuer will go in default. It will be assumed that, if the issuer goes into default, the issuer will not recover and therefore will stay in default such that she is no longer able to fulfill any of the future payments by the bond. The company can therefore only default on any of its coupon payment dates and the default will only occur if the company does not own sufficient capital to pay off the bondholder. To determine whether the company defaults at a specific coupon payment, the assets of the company need to be compared to the specific coupon payment the company has to pay. Therefore, consider a process $\{A(t)\}_{t>0}$ governing the value of the company. The company will default if coupon payment at time $\tau_j \in \mathcal{T}^{\text{coup}}$ exceeds the value of the company $A(\tau_j)$ at the coupon payment date τ_j, i.e., the company will default if,

$$\begin{cases} A(\tau_j) < F\alpha & \text{for } j \in \{1, \dots, n-2\}, \\ A(\tau_j) < F(1+\alpha) & \text{for } j = n-1. \end{cases}$$

Contrary to shareholders, bondholders are among the first who may have a claim right on the company value when a default occurs. Which part of the remaining capital of the company may be recovered by a bondholder is referred to as the seniority of the bond. A bondholder, therefore, has a more senior claim on the remaining capital than the shareholder, as the shareholder is only entitled to recover any of the remaining capital when all other claim rights have been exercised. By the limited liability condition, the bondholder can only recover the maximum of the assets available after default. It is assumed that, when the company defaults, the company will be liquidated and the remaining value of the company is totally obtained. Hence, the payoff of a corporate bond, in case of default at time $\tau_j \in \mathcal{T}^{\text{coup}}$, is the minimum of the maximum amount of assets that can be recovered and her claim under the bond, i.e. $\min(A(\tau_j), F(1+\alpha))$. The recovery of the company value upon a default may be extended with a recovery rate. However, within this thesis, a full recovery will be assumed. Denote the payoff function of the corporate bond at time $t \in [t_0, T]$ by $H(t, A(t))$. At non-coupon payment dates $t \notin \mathcal{T}^{\text{coup}}$, the corporate bond will have no payoff. On the other hand, if $t \in \mathcal{T}^{\text{coup}}$ is a coupon payment date, the bond will pay

$$H(t, A(t)) = \begin{cases} \min(\alpha F, A(t)) & \text{for } j \in \{1, \dots, n-2\}, \\ \min((1+\alpha)F, A(t)) & \text{for } j = n-1. \end{cases}$$

A plot of the payoff function H of a corporate bond at a coupon date $\tau_j \in \mathcal{T}^{\text{coup}}$ and at maturity T is given in Figure 1.3.

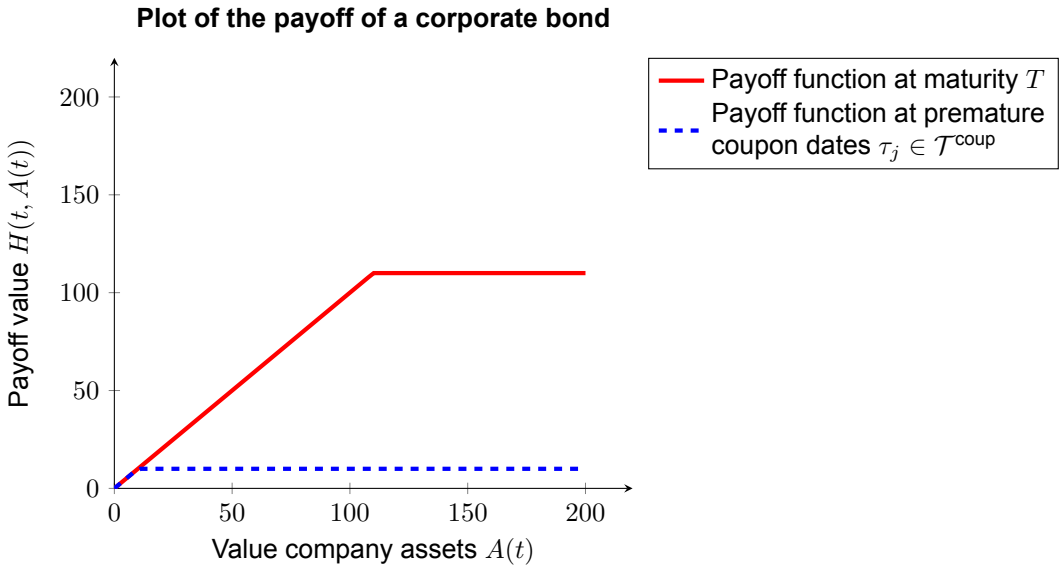


Figure 1.3: A plot of the payoff function corresponding to a corporate bond. The bond has a face value $F = 100$ with regular coupon payments $c_j = \alpha F = 10$ for $\alpha = 0.1$ at coupon payment dates $\tau_j \in \mathcal{T}^{\text{coup}}$. The red line corresponds with the payoff at maturity T, i.e. $H(T, A(T)) = \min((1+\alpha)F, A(T))$. The blue line shows the payoff at any premature coupon payment date $\tau_j \in \mathcal{T}^{\text{coup}}$, i.e. $H(\tau_j, A(\tau_j)) = \min(\alpha F, A(\tau_j))$ for $\tau_j < T$.

The value of any contingent claim under the equivalent martingale measure $\mathbb{Q}$ is equal to the discounted expected value of its future cash flows, where the discounting is performed with respect to the money-market account [32].

Definition 1.2. The money-market account yields the possibility to invest in a risk-free asset. The value of the money-market account at time $t \in [t_0, T]$ is denoted by $M(t)$. It is assumed that the initial value of the money-market account at time t_0 is $M(t_0) = 1$. Furthermore, it is assumed that the money-market account evolves following the differential equation,

$$\mathrm{d}M(t) = r(t)M(t)\mathrm{d}t.$$

The process $\{r(t)\}_{t\geq 0}$, which governs the growth of the money-market account, is referred to as the instantaneous short rate (or abbreviated as short rate).

When an investor invests in a money-market account, the investment will have a (locally) risk-free return [8]. Although the return is locally risk-free, the amount of the return may vary and can even be negative which will reverse the direction of the interest payments. Given the initial value of the money-market account and the differential equation governing its value, the value of the money-market account at time T, with an initial investment of $M(t)$ at time $0 \leq t < T$, can be given in closed-form [32],

$$M(T) = M(t)\mathrm{e}^{\int_t^T r(s)\mathrm{d}s},$$

where the $\{r(t)\}_{t\geq 0}$ is the short rate process. Note that based on the available information of the process $\{r(t)\}_{t\geq 0}$, the value of $M(T)$ for given $M(t)$ may be a stochastic quantity.

The value of a corporate bond can be expressed in a martingale formulation under the equivalent martingale measure $\mathbb{Q}$ [32]. In particular, the discounted value of the corporate bond needs to be a martingale under the measure $\mathbb{Q}$. Denote the value at time $t \in [t_0, T]$ of a corporate bond as $V(t, A(t))$, then the martingale property

$$\frac{V(t, A(t))}{M(t)} = \mathbb{E}\left[\left.\frac{V(T, A(T))}{M(T)}\right| \mathcal{F}_t\right], \tag{1.1}$$

needs to hold under the measure $\mathbb{Q}$, where $\mathcal{F}_t$ is the filtration generated by the processes $\{r(t)\}_{t\geq 0}$ and $\{A(t)\}_{t\geq 0}$. Recall the zero-coupon bond with maturity T defined in Definition 1.1. As the zero-coupon

bond is a special case within the class of bonds with a payoff $P(T,T) = 1$ at maturity, the value $P(t,T)$ at time t can be derived from the value given in Equation (1.1) and is equal to

$$P(t,T) = \mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\mathrm{d}s} \;\middle|\; \mathcal{F}_t \right],$$

where $\mathcal{F}_t$ is the filtration corresponding to the process $\{r(t)\}_{t\geq 0}$ at time t.

1.2.2. Convertible option

A corporate bond can be embedded with a convertible option, to make the bond more attractive for investors. By exercising the option, also called converting the bond, the corporate bond is terminated and the bondholder receives a predetermined amount of shares in return. The fraction of total shares that the bondholder receives after converting the convertible bond is called the conversion factor and is denoted by $\gamma \in [0,1]$. Exercising the conversion right can only be done once at $m-1$ predetermined dates $\mathcal{T}^{\text{conv}}$. If the bondholder decides to convert the bond, she is no longer entitled to any future payments under the bond. If the convertible option is not exercised during any of the possible exercise dates $\mathcal{T}^{\text{conv}}$, the issuer of the bond is obliged to continue paying any payments that are agreed on at the issuance of the bond. These payments include any further coupon payments and the payment of the face value of the bond at maturity.

When the bond is converted, the bondholder receives a predetermined amount of shares and thus the value of the option is dependent on the share price (equity value). However, the valuation problem of the corporate bond is described in terms of the value of the company itself. Fortunately, these two concepts can be linked through the capital structure of the company [21]. By assuming that the Modigliani-Miller theorem holds, the company value can be written in terms of an equity and debt component [31, 21]. Let $\{S(t)\}_{t\geq 0}$ be a stochastic process, which governs the value of the stock of the company. Furthermore, assume that the company only issues one convertible bond. Denote the value of the convertible bond at time $t \in [t_0, T]$ by $V_{\text{conv}}(t, A(t))$. Due to the capital structure of the company, the relation between the company value, the debt, and the equity of the company is given by

$$S(t) = A(t) - V_{\text{conv}}(t, A(t)). \tag{1.2}$$

As the debt will be terminated upon conversion, the equation

$$S(t) = A(t),$$

holds in case the convertible bond is converted. Hence, if the bondholder decides to exercise her option, she will receive the value $\gamma A(t)$ in shares.

It is assumed that the main objective of a rational investor will be to maximize her wealth. At maturity T, the bond will therefore be converted if and only if the conversion value exceeds the face value F of the bond increased with the final coupon payment αF. This yields that the payoff of a convertible bond $H_{\text{conv}}(T, A(T))$ at maturity T is given by

$$H_{\text{conv}}(T, A(T)) = \begin{cases} H(T, A(T)) & \text{for } F(1+\alpha) \geq \gamma A(T), \\ \gamma A(T) & \text{for } F(1+\alpha) < \gamma A(T). \end{cases}$$

The convertible bond may be exercised before maturity T. However, Ingersoll Jr showed that the optimal exercising strategy for a convertible bond will be exercising at maturity. For completeness, the optimal conversion moment for convertible bonds will be proven [21].

Theorem 1.1. *Assume a convertible bond $V_{\text{conv}}(t, A(t))$ with maturity T issued by a company with value process $\{A(t)\}_{t\geq 0}$. Let the convertible bond have a constant conversion factor γ, and assume that the stock concerning the company does not pay any dividends. Then a rational investor would not convert the convertible bond prior to maturity.*

Proof. Denote the value of the convertible bond at time t by $V_{\text{conv}}(t, A(t))$. Suppose two portfolios P_A and P_B, where P_A contains only the convertible bond with value $V_{\text{conv}}(t, A(t))$ at time t and P_B contains

the fraction γ of the convertible bond together with the fraction γ of shares of the company with value $\gamma S(t)$. By Equation (1.2), the value of portfolio P_B will become:

$$\gamma(V_{\text{conv}}(t, A(t)) + S(t)) = \gamma A(t),$$

and thus the portfolio P_B always has the value $\gamma A(t)$. If the holder of portfolio P_A decides to convert his convertible, then the value of portfolio P_A and portfolio P_B will be the same for the remaining time within the time horizon $[t_0, T]$. If, on the other hand, the holder of portfolio P_A decides to not convert her convertible bond, then the payout at maturity T within portfolio P_A will be

$$\begin{cases} A(T) & \text{for } \gamma A(T) \leq A(T) \leq F, \\ F & \text{for } \gamma A(T) \leq F \leq A(T), \\ \gamma A(T) & \text{for } F \leq \gamma A(T) \leq A(T). \end{cases} \tag{1.3}$$

Hence, the value of portfolio P_A at maturity T will always be greater or equal to the value of portfolio P_B. Now suppose that there exists a $t \in [t_0, T]$ such that $V_{\text{conv}}(t, A(t)) \leq \gamma A(t)$. Furthermore, take a short position in portfolio P_B and a long position in portfolio P_A. Decide not to convert the position in portfolio P_A until maturity T. As by Equation (1.3), the value of portfolio P_A at maturity T is always greater or equal than the value of portfolio P_B at maturity. Hence, there exists an arbitrage. Since arbitrages are not allowed, it must hold that $V_{\text{conv}}(t, A(t)) > \gamma A(t)$. However, since $V_{\text{conv}}(t, A(t)) > \gamma A(t)$ it is not optimal to convert before maturity. Hence a rational investor will not convert his claim before maturity T, which concludes the proof. □

A plot of the payoff function of the convertible bond at maturity is given in Figure 1.4.

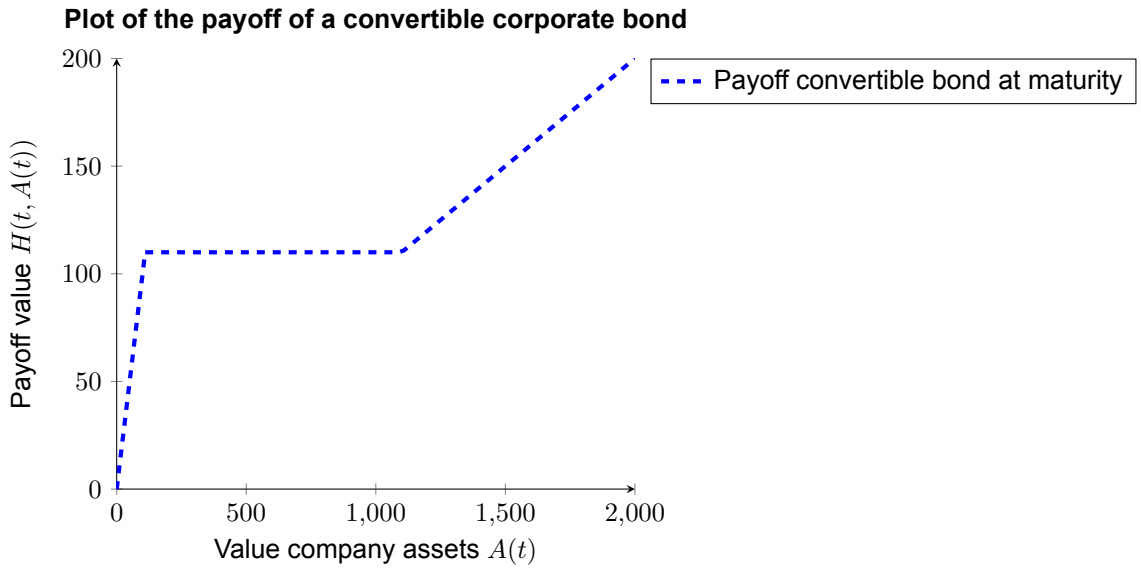


Figure 1.4: A plot of the payoff function corresponding to a convertible corporate bond for different times within the time horizon $[t_0, T]$. The bond has a face value $F = 100$ with regular coupon payments $\alpha = 0.1$ at coupon payment dates $\tau_j \in \mathcal{T}^{\text{coup}}$. Furthermore, the conversion factor $\gamma = 0.1$. The blue line corresponds with the payoff at maturity T, i.e. $H_{\text{conv}}(T, A(T))$. The first part of the curve corresponds to the payoff at default. The middle part represents the constant payoff in the case of the final payment at maturity. The last part of the curve represents the payoff when conversion is the optimal strategy.

As the optimal exercise moment for the convertible bond is at maturity T, a convertible bond without any coupon payments is equivalent to a European-style option. At maturity T, the value of the convertible bond will be equal to the payoff $H_{\text{conv}}(T, A(T))$, i.e.

$$V_{\text{conv}}(T, A(T)) = H_{\text{conv}}(T, A(T)).$$

Furthermore, the current value of the convertible bond is given as the discounted payoff at maturity T under the measure $\mathbb{Q}$,

$$V_{\text{conv}}(t_0, A(t_0)) = \mathbb{E}\left[\mathrm{e}^{-\int_{t_0}^{T} r(s)\mathrm{d}s} H_{\text{conv}}(T, A(T)) \;\middle|\; \mathcal{F}_{t_0} \right],$$

where $\mathcal{F}_{t_0}$ is the filtration corresponding to the process $\{A(t)\}_{t\geq 0}$ and $\{r(t)\}_{t\geq 0}$.

Though the assumption of no dividends can easily be relaxed, in this thesis, the assumption of no dividends is made to simplify the model. Relaxing the no divided condition may imply that Theorem 1.1 does no longer hold. In the case Theorem 1.1 does not hold, the optimal exercising policy can still be determined. Under the assumption of a rational investor, the investor will always try to maximize her wealth. As a consequence, the investor will exercise her option if and only if the expected future discounted value of the bond is less than the value at immediate exercise. To make this more precise, the continuation value of the convertible bond is defined.

Definition 1.3. Assume a convertible bond with exercise dates $t_l \in \mathcal{T}^{\text{conv}}$. The continuation value $C(t_l, A(t_l))$ of the convertible bond at time t_l is defined as

$$C(t_l, A(t_l)) = \mathbb{E}\left[\mathrm{e}^{-\int_{t_j}^{t_{j+1}} r(s)\mathrm{d}s} V_{\text{conv}}(t_{j+1}, A(t_{j+1})) \;\middle|\; \mathcal{F}_{t_j} \right],$$

where $V_{\text{conv}}(t_{j+1}, A(t_{j+1}))$ is the convertible bond value at time $t_{j+1} \in \mathcal{T}^{\text{conv}}$ and $\mathcal{F}_{t_j}$ is the filtration at time t_j corresponding to the processes $\{r(t)\}_{t\geq 0}$ and $\{A(t)\}_{t\geq 0}$ governing the value of the interest rates and the company value respectively.

Given the continuation value $C(t_l, A(t_l))$, the optimal policy for the investor to exercise her option is to exercise if

$$C(t_l, A(t_l)) \leq \gamma A(t_l),$$

where γ is the conversion factor.

1.2.3. Callable option

A callable option gives the holder of the option, the issuer of the bond, the right to buy back the bond for a predetermined price. The value for which the bond can be bought back is called the call price, $K(t)$, and may be dependent on the time of exercise. It will be assumed that the call price is a penalty function $\beta(t) : \mathbb{R}_+ \to \mathbb{R}_+$ to the face value F. In practice, most call price functions are step functions over time. For each predetermined time interval in which the call option can be exercised corresponding call prices are agreed upon before issuing the bond. These step functions can be written in the form,

$$K(t) = F\beta(t),$$

for a corresponding step function β. Exercising the option is often referred to as calling the bond, and can only be done once for a predefined set of $p-1$ exercise dates, $\mathcal{T}^{\text{call}} = \{\tau_1, \ldots, \tau_{p-1}\}$. At any date that is not included in the set of predetermined exercising moments, the issuer of the bond is not allowed to call the bond. When the bond is called, the bondholder is no longer entitled to any of the future payments by the bond. However, if the bond is not called during any of the exercise moments, the issuer of the bond is obliged to continue paying any future coupon payments and the face value of the bond at maturity. The callable option embedded within the bond can only be called before the maturity T of the bond. The set of call dates is often limited to a certain period. The period in which the call option cannot be called is referred to as the call protection period. When the call option is called, the bondholder is also entitled to the accrued interest over the period between the last and the upcoming coupon payment. Let the accrued interest $I_j(t)$ at time t between coupon payment dates t_j and t_{j+1} be given as

$$I_j(t) = \frac{t - t_j}{t_{j+1} - t_j} F\alpha,$$

for $j \in \{1, \ldots, n-2\}$. In general, the accrued coupon payment $I(t)$ for an arbitrary time $t \in [t_0, T]$ is given by

$$I(t) = \sum_{j=1}^{n-2} I_j(t) \mathbb{1}_{\{t\in[t_j, t_{j+1}]\}}(t).$$

Denote the value of the callable bond at time $t \in [t_0, T]$ by $V_{\text{call}}(t, A(t))$. As the bond can only be called before maturity, the payoff at maturity is equal to the payoff of the underlying bond,

$$H_{\text{call}}(T, A(T)) = H(T, A(T)),$$

in case of a corporate bond. At any exercise date $\tau_j \in \mathcal{T}^{\text{call}}$ before maturity T, the bond can be called and the payoff will be the call price $K(\tau_j)$ increased with the accrued interest $I(\tau_j)$, i.e.

$$H_{\text{call}}(\tau_j, A(\tau_j)) = F\beta(\tau_j) + I(\tau_j).$$

Note that, if the call date coincides with a coupon payment date, the accrued interest will be equal to the coupon payment, i.e. $I(\tau_{j+1}) = \alpha F$ for $\tau_{j+1} \in \mathcal{T}^{\text{coup}}$. An example of the payoff of a call option over time is given in Figure 1.5.

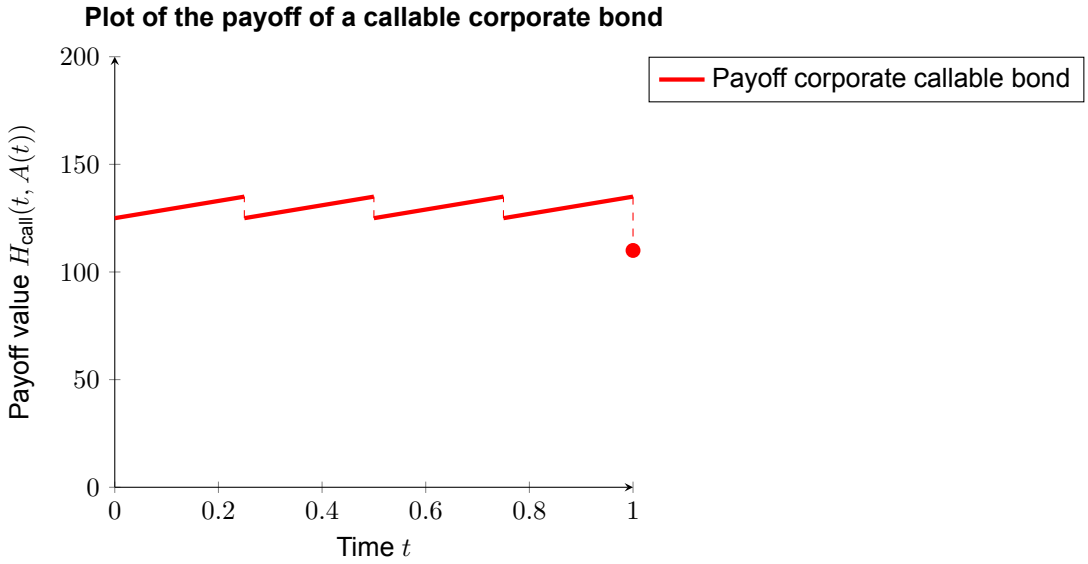


Figure 1.5: A plot of the payoff function corresponding to a callable corporate bond for different times t within the time horizon $[t_0, T]$. The bond has a face value $F = 100$ with regular coupon payments $\alpha = 0.1$ at coupon payment dates $\tau_j \in \mathcal{T}^{\text{coup}}$. Furthermore, the call price is constant and equal to $K(t) \equiv K = 125$. Given a set of call dates $\mathcal{T}^{\text{call}}$, the payoff, $H_{\text{call}}(\tau_j, A(\tau_j))$ for $\tau_j \in \mathcal{T}^{\text{call}}$, of the callable bond can be derived from the given plot.

At maturity T, the value of the callable bond is known and given by,

$$V_{\text{call}}(T, A(T)) = H_{\text{call}}(T, A(T)).$$

By assumption (see Appendix A.1), the main objective of the management of the company is to maximize the wealth of the shareholders. By the capital structure assumption,

$$A(t) = S(t) + V_{\text{call}}(t, A(t)),$$

and therefore maximizing the wealth of the shareholders is equivalent to minimizing the value of the callable bond. The exercising moment τ_{call} of the call option will therefore be that moment $\tau_{\text{call}} \in \mathcal{T}^{\text{call}}$ such that the value of the bond is minimized, i.e.

$$\tau_{\text{call}} = \underset{\tau \in \mathcal{T}^{\text{call}}}{\arg\min}\, \mathbb{E}\left[\mathrm{e}^{-\int_{t_0}^{\tau} r(s)\mathrm{d}s} V_{\text{call}}(\tau, A(\tau)) \,\middle|\, \mathcal{F}_0 \right].$$

1.2.4. Callable convertible bond

A corporate bond may also be embedded with a callable and convertible option. The bondholder then owns the convertible option and the issuer of the bond has the right to exercise the call option. It is assumed that the exercising dates for the call option coincide with the exercising dates of the convertible option, i.e. $\mathcal{T}^{\text{call}} \subseteq \mathcal{T}^{\text{conv}} \subset [t_0, T]$. Furthermore, if a convertible bond is embedded with a call option, a call notice period may be included within the callable option. The call notice period gives the bondholder time to decide on converting the bond. When the remaining time until maturity is smaller than the call notice period, the call notice period will be equal to the remaining time until maturity T. At the end of the call notice period, the bondholder receives the call price $K(\tau_j)$ raised with the accrued interest $I(\tau_j)$

over the call notice period t_c when accepting the call, or the conversion value $\gamma A(t+t_c)$ when deciding to convert. A conversion during a call notice is often referred to as a forced conversion [3]. Note that it is assumed the bondholder is not entitled to any accrued interest during the call notice period. The assumption of not adding the accrued interest obtained during the call notice period is non-restrictive and can straightforwardly be relaxed for cases where the bond does pay the accrued interest over the call notion period.

Still, the call option can only be exercised before maturity T. Hence, the payoff at maturity will be,

$$H_{\mathsf{cc}}(T, A(T)) = H_{\mathsf{conv}}(T, A(T)).$$

Moreover, the issuing company can only call the bond before maturity T, if it has sufficient capital to pay the call price increased with the accrued interest. When the bond is called at any of the call dates, the bondholder has to choose between conversion and collecting the call price increased with the accrued interest. First, assume that the call date coincides with a coupon payment date and assume that the issuing company has sufficient capital to exercise the call. If the call price combined with the accrued coupon payment exceeds the conversion value, the bondholder will accept the call. Otherwise, the bondholder will exercise her option and convert the bond. On the other hand, if the company does not have sufficient capital for calling the bond, the company may also default on the regular coupon payment. Next assume that the call date does not coincide with a coupon payment date. The company can only call the bond if the company has sufficient capital to pay the call price increased with the accrued interest. If the bond is called, the bondholder will exercise the conversion option if the conversion value exceeds the call prices increased with the accrued interest. Otherwise, she will accept the call price together with the accrued interest. At the end of the call notice period, the bondholder will communicate her choice and accept the payments. The payoff at call $\tau_j \in \mathcal{T}^{\mathsf{call}}$, adjusted for the call notice period t_c, will therefore become,

$$H_{\mathsf{cc}}(\tau_j) = \begin{cases} K(\tau_j) + I(\tau_j) & \text{for } \gamma A(\tau_j) \leq K(\tau_j) + I(\tau_j) \leq A(\tau_j), \\ H(\tau_j, A(\tau_j)) & \text{for } A(\tau_j) \leq K(\tau_j) + I(\tau_j), \\ \gamma A(\tau_j) & \text{for } K(\tau_j) + I(\tau_j) \leq \gamma A(\tau_j), \end{cases}$$

where the form $K(t) = F\beta(t)$, for the call penalty function β and face value F, is assumed for the call price, $I(t)$ is the accrued interest at time t and $A(t)$ is the value of the company at time t. A graphical representation of the payoff function for different moments in time is given in Figure 1.6.

Let the value of the callable convertible bond at time t be denoted as $V_{\mathsf{cc}}(t, A(t))$. As the callable convertible bond will have a payoff directly after the call notice period, the discounting period should also be increased with the call notice period. Therefore, the management of the company will call the bond at the date $\tau_{\mathsf{call}} \in \mathcal{T}^{\mathsf{call}}$ if the call date satisfies

$$\tau_{\mathsf{call}} = \underset{\tau \in \mathcal{T}^{\mathsf{call}}}{\arg\min}\, \mathbb{E}\left[\mathrm{e}^{-\int_{t_0}^{\tau+t_c} r(s)\mathrm{d}s} V_{\mathsf{cc}}(\tau + t_c, A(\tau + t_c)) \;\middle|\; \mathcal{F}_0 \right],$$

when a call notice period is included. The optimal conversion time may be different due to the presence of the call option. Ingersoll Jr shows that, even if a call option is embedded within the convertible bond, the optimal exercising moments for the convertible bond will still be at the moment the bond is called or at maturity T [21]. However, to generalize the optimal moment for converting the bond, assume that the objective of the holder of the option is to maximize her wealth. Therefore, the bond will be converted at time $\tau_{\mathsf{conv}} \in \mathcal{T}^{\mathsf{conv}}$, if at that time the expected discounted value of the bond $V_{\mathsf{cc}}(\tau_{\mathsf{conv}}, A(\tau_{\mathsf{conv}}))$ is maximized, i.e.

$$\tau_{\mathsf{conv}} = \underset{\tau \in \mathcal{T}^{\mathsf{cc}}}{\arg\max}\, \mathbb{E}\left[\mathrm{e}^{-\int_{t_0}^{\tau} r(s)\mathrm{d}s} V_{\mathsf{cc}}(\tau, A(\tau)) \;\middle|\; \mathcal{F}_0 \right].$$

Denote the termination date of the callable convertible bond by τ, which is given by the first moment that the bond is terminated, i.e. $\tau = \tau_{\mathcal{D}} \wedge \tau_{\mathsf{conv}} \wedge \tau_{\mathsf{call}} \wedge T$, where $\tau_{\mathcal{D}}$ is the moment at which the company will default, i.e. $\tau_{\mathcal{D}} = \inf\{t : A(t) < \alpha F\}$ for face value F and coupon rate α. The value of the callable convertible bond at time t_0 can therefore be given in terms of its termination date τ,

$$V_{\mathsf{cc}}(t_0, A(t_0)) = \mathbb{E}\left[\mathrm{e}^{-\int_{t_0}^{\tau^c} r(s)\mathrm{d}s} V_{\mathsf{cc}}(\tau_c, A(\tau_c)) \;\middle|\; \mathcal{F}_{t_0} \right],$$

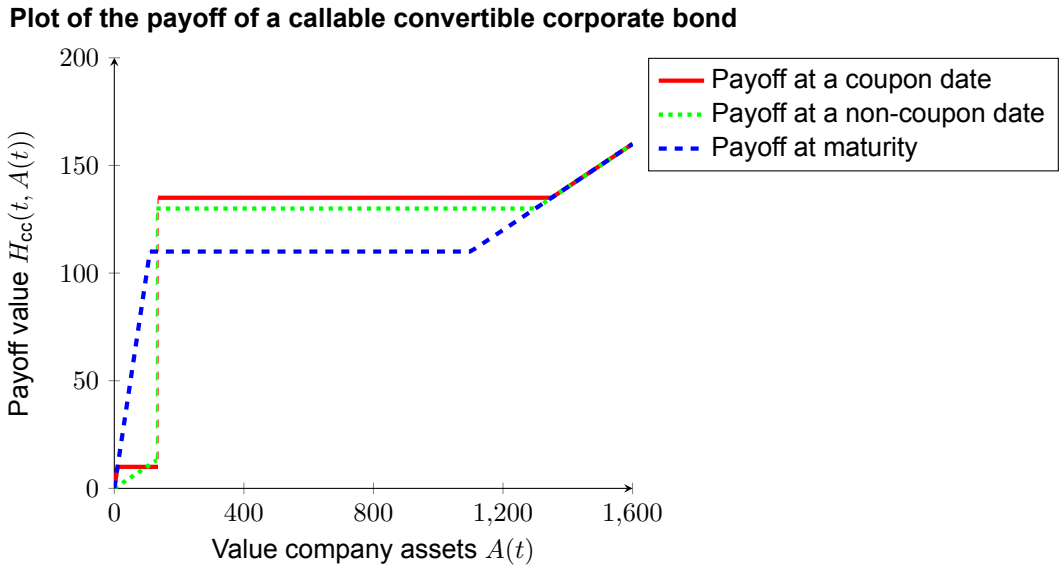


Figure 1.6: A plot of the payoff function of a callable convertible for different moments in time; a coupon payment date, a non-coupon payment date, and maturity. The call price is equal to $K(t) = 125$ for all moments considered. At maturity, no call price is considered and only the convertible option is given. At the coupon payment date, a coupon payment rate $\alpha = 0.1$ is considered. For the non-coupon payment date, it is assumed that the accrued interest $I(t + t_c) = 5$. The face value $F = 100$ of the bond is given. The red line shows the payoff of a callable convertible bond at a coupon date before maturity, i.e. $H_{cc}(\tau_j, A(\tau_j)), \tau_j \in \mathcal{T}^{coup}, \tau_j < T$. The green line shows the payoff of a callable convertible bond at a non-coupon date, i.e. $H_{cc}(\tau_j, A(\tau_j)), \tau_j \notin \mathcal{T}^{coup}, \tau_j < T$. The blue line represents the payoff of the callable convertible bond at maturity T, i.e. $H_{cc}(T, A(T))$.

where $\tau^c = \tau + t_c \mathbb{1}_{\{\tau = \tau_{call}\}}$, for the call notice period t_c and filtration $\mathcal{F}_{t_0}$ corresponding to the interest rate process $\{r(t)\}_{t \geq 0}$ and company value process $\{A(t)\}_{t \geq 0}$ at time t_0.

For both the management of the company and the investor the optimal exercise policy can be determined. Based on the interests of the management and the interests of the investor, the call and convertible option holder will exercise the option when the future discounted value of the bond exceeds the immediate payoff or when the immediate exercise value of the bond exceeds the discounted future value respectively. To make the policy for both parties within the contract more precise, the following definition for the continuation value of the bond will be given.

Definition 1.4. The continuation value of the callable convertible bond $C(t_j, A(t_j))$ at time t_j within the set of intermediate exercise and coupon dates $\mathcal{T}^{coup} \cup \mathcal{T}^{call} \cup \mathcal{T}^{conv}$, is equal to the expected discounted value of the bond under the measure $\mathbb{Q}$ at the subsequent date $t_{j+1} \in \mathcal{T}^{coup} \cup \mathcal{T}^{call} \cup \mathcal{T}^{conv}$, i.e.

$$C(t_j, A(t_j)) = \mathbb{E}\left[e^{-\int_{t_j}^{t_{j+1}} r(s) ds} V_{cc}(t_{j+1}, A(t_{j+1})) \,\middle|\, \mathcal{F}_{t_j} \right],$$

where $V_{cc}(t_{j+1}, A(t_{j+1}))$ is the bond value at time t_{j+1} and $\mathcal{F}_{t_j}$ is the filtration corresponding to the processes $\{r(t)\}_{t \geq 0}$ and $\{A(t)\}_{t \geq 0}$ at time t_j.

The convertible option holder will exercise her option when the immediate exercise of the option will exceed the continuation value $C(t_j, A(t_j))$. Hence, the exercise policy will be to exercise the convertible option if

$$C(t_j, A(t_j)) \leq \gamma A(t_j).$$

The management of the company, on the other hand, needs to take the call notion period into account, as the investor will receive her payoff at the end of the call notion period t_c. When the call notice period t_c is taken into account, the continuation value needs to be evaluated at time $t_{j+1} + t_c$ for $t_{j+1} \in \mathcal{T}^{coup} \cup \mathcal{T}^{call} \cup \mathcal{T}^{conv}$.

Definition 1.5. The continuation value $C_{call}(t_j, A(t_j))$ of a callable convertible bond taken with respect

to the call notion period t_c will be defined as

$$C_{\text{call}}(t_j, A(t_j)) = \mathbb{E}\left[\mathrm{e}^{-\int_{t_j}^{t_{j+1}+t_c} r(s)\mathrm{d}s} V_{\text{cc}}(t_{j+1}+t_c, A(t_{j+1}+t_c)) \;\middle|\; \mathcal{F}_{t_j} \right],$$

where $V_{\text{cc}}(t_{j+1}, A(t_{j+1}))$ is the bond value at time t_{j+1} and $\mathcal{F}_{t_j}$ is the filtration corresponding to the processes $\{r(t)\}_{t\geq 0}$ and $\{A(t)\}_{t\geq 0}$ at time t_j.

As mentioned, the main objective of the management of the company is to maximize the wealth of the shareholders. As a consequence, they will try to minimize the value of the bond. Therefore, the policy followed by the management of the company will be to exercise the call option as soon as the value at immediate exercise exceeds the continuation value of the bond, when considering the call notice period, i.e.

$$C_{\text{call}}(t_j, A(t_j)) \geq K(t_j) + I(t_j), \tag{1.4}$$

for the call price function $K(t_j)$ and the accrued interest $I(t_j)$ at time t_j. Note that, from the capital structure of the company, if the condition in Equation (1.4) is satisfied, the company will also be able to pay the call value $K(t_j)$ increased with the accrued interest rate $I(t_j)$ at the moment the call option is exercised.

2

Model description

2.1. Introduction

In this thesis, a structural default model is proposed for modeling the value of different sorts of bonds. Structural models assume that the default of a company is intrinsically driven by the ability of the company, issuing the bond, to fulfill the obliged claims by the bond. Therefore, these models assume that the value of a company can be modeled by a stochastic process. If the value of the company is sufficient, it has the resources to fulfill the claims. On the other hand, if the value is lower than the claim, the company will default and be liquidated.

In this section, a stochastic model for the company value process $\{A(t)\}_{t\geq 0}$ is proposed. It is assumed that the company value follows a jump-diffusion process. Furthermore, an extension of the model the short rate process $\{r(t)\}_{t\geq 0}$ is also assumed to follow a stochastic process. To be able to fit the current term structure, the short rate is modeled under the Heath-Jarrow-Morton framework. Both models combined, are included within the so-called class of affine (jump-)diffusion models. After defining the dynamics of the risk factors, this chapter will continue with a brief explanation of the affine class. Finally, using the properties of processes in the affine class the characteristic function for the model can be derived. The characteristic function of the model will be used in later chapters to value different sorts of bonds.

2.2. Dynamics of the risk factors

The process $\{A(t)\}_{t\geq 0}$ governs the company value and will be modeled by a jump-diffusion process. A jump-diffusion process can be divided into two sub-processes; a continuous diffusion process and a discontinuous jump process. Consider a constant volatility $\sigma \in \mathbb{R}_+$ corresponding to the volatility of the process $\{A(t)\}_{t\geq 0}$. The continuous part of the process $\left\{A^C(t)\right\}_{t\geq 0}$ will be defined as

$$\mathsf{d}A^C(t) = r(t)A(t)\mathsf{d}t + \sigma A(t)\mathsf{d}W_A(t),$$

for $\{W_A(t)\}_{t\geq 0}$, a Brownian motion under the measure $\mathbb{Q}$. The jump arrivals will be assumed to follow a Poisson process $\{\mathcal{P}(t)\}_{t\geq 0}$ under the measure $\mathbb{Q}$ with intensity parameter $\lambda \in \mathbb{R}_+$. Furthermore, let J be a stochastic variable that models the size of the jumps. Assume that the jump size J, the Poisson process $\{\mathcal{P}(t)\}_{t\geq 0}$, and the Brownian motion $\{W_A(t)\}_{t\geq 0}$ are mutually independent. The process governing the value of the assets of the company, $\{A(t)\}_{t\geq 0}$, follows the jump-diffusion process given by

$$\frac{\mathsf{d}A(t)}{A(t)} = r(t)\mathsf{d}t + \sigma \mathsf{d}W_A(t) + \left(\mathrm{e}^J - 1\right)\mathsf{d}\mathcal{P}(t). \tag{2.1}$$

As the process $\{A(t)\}_{t\geq 0}$ follows a jump-diffusion model, the corresponding market will no longer be complete [32]. As a consequence, the risk-neutral measure is no longer unique. To ensure the no-arbitrage condition, the equivalent martingale measure $\mathbb{Q}$ is introduced. Under the measure $\mathbb{Q}$, the discounted company value will be a martingale. Furthermore, the measure $\mathbb{Q}$ will be equivalent to the

real-world measure $\mathbb{P}$. A more detailed discussion of the equivalent martingale measure lies outside the scope of this thesis but can for instance be found in the work of Oosterlee and Grzelak [32].

In the literature, two generally accepted models for the jump-diffusion processes are known and were given by Kou and Merton [22, 30]. Both models assume a different distribution for the jump sizes J. The model of Merton assumes that the jump sizes J are normally distributed with mean μ_J and variance σ_J^2. On the other hand, the model of Kou assumes that the jump sizes follow a non-symmetric double exponentially distribution. The density function corresponding to the model proposed by Kou is given by [32]

$$f_J(x) = p_1\alpha_1 \mathrm{e}^{-\alpha_1 x}\mathbb{1}_{\{x\geq 0\}} + p_2\alpha_2\mathrm{e}^{\alpha_2 x}\mathbb{1}_{\{x<0\}}.$$

The parameter $p_1 \in [0,1]$ yields the probability of a upward jump, $p_2 = 1 - p_1$ that of a downward jump and the parameters α_1 and α_2 yield the means of the individual exponential distributions. To ensure the existence of the first moment, the means are set as $\alpha_1 > 1$ and $\alpha_2 > 0$ [32]. A derivation of the density property is given in Appendix B.1.

To be able to derive the characteristic function, the model described in Equation (2.1) will be transformed. Define the log-transform of the process $\{A(t)\}_{t\geq 0}$ by

$$X(t) = \log\left(A(t)\right),$$

such that the dynamics of the process $\{X(t)\}_{t\geq 0}$ are given by

$$\mathsf{d}X(t) = \left(r(t) - \lambda\mathbb{E}\left[\mathrm{e}^J - 1\right] - \frac{1}{2}\sigma^2\right)\mathsf{d}t + \sigma(t)\mathsf{d}W_A(t) + J\mathsf{d}\mathcal{P}(t).$$

The dynamics of the process $\{X(t)\}_{t\geq 0}$ are a direct application of Ito's lemma to the transformation $X(t)$. In addition, the diffusion part of the process $\{X(t)\}_{t\geq 0}$ will be defined as $X^C(t)$ and is given by

$$\mathsf{d}X^C(t) = \left(r(t) - \frac{1}{2}\sigma^2\right)\mathsf{d}t + \sigma\mathsf{d}W_A(t).$$

If the diffusion part of the process $\{X(t)\}_{t\geq 0}$ is removed, the part of the process that governs the jumps remains. Denote $\left\{X^J(t)\right\}_{t\geq 0}$ for the jump part of the process, the dynamics of this process are given by

$$\mathsf{d}X^J(t) = -\lambda\mathbb{E}\left[\mathrm{e}^J - 1\right]\mathsf{d}t + J\mathsf{d}\mathcal{P}(t).$$

2.2.1. The Hull-White dynamics under the Heath-Jarrow-Morton framework

The short rates will be modeled by the one-factor Hull-White dynamics under the Heath-Jarrow-Morton (HJM) framework. The HJM framework models the short rate process directly from the yield curve, which can be done by fitting the current forward rates to the current yield curve. To make this more precise, denote the instantaneous forward rate, the interest rate return at time t on a forward contract starting at time T and maturing at $\mathsf{d}t$ (an infinitesimal moment in time) later, as $f_r(t,T)$. Under the HJM framework, the forward rates can be determined based on the market yield. Given the current market prices of zero-coupon bonds $P_{\mathsf{mkt}}(0,t)$ with maturity t, the forward rate f_r under the HJM framework is given by [32]

$$f_r(0,t) = -\frac{\partial \log P_{\mathsf{mkt}}(0,t)}{\partial t}.$$

Furthermore, it is assumed that the instantaneous forward rates follow a stochastic process [32],

$$\mathsf{d}f_r(t,T) = \alpha^{\mathbb{Q}}(t,T)\mathsf{d}t + \bar{\eta}(t,T)\mathsf{d}W_f(t),$$

where

$$\alpha^{\mathbb{Q}}(t,T) = \bar{\eta}(t,T)\int_t^T \bar{\eta}(t,z)\mathsf{d}z,$$

for a function $\bar{\eta}$ and $\{W_f\}_{t\geq 0}(t)$ a Brownian motion under the equivalent martingale measure $\mathbb{Q}$. As a consequence, the arbitrage-free representation of the instantaneous forward model is determined by its diffusion function $\bar{\eta}$.

The Hull-White model can be retrieved by setting [32]

$$\bar{\eta}(t,T) = \eta \mathrm{e}^{-\kappa(T-t)},$$

where $\kappa, \eta \in \mathbb{R},\ \eta > 0,\ \kappa \neq 0$ are the speed of mean reversion and diffusion parameters respectively. By setting the function θ as [32]

$$\theta(t) = \frac{1}{\kappa}\frac{\partial}{\partial t} f_r(0,t) + f_r(0,t) + \frac{\eta^2}{2\kappa^2}(1 - \mathrm{e}^{-2\kappa t}), \tag{2.2}$$

the Hull-White dynamics can be written as

$$\mathrm{d}r(t) = \kappa(\theta(t) - r(t))\mathrm{d}t + \eta \mathrm{d}W_r(t), \tag{2.3}$$

where $\{W_r(t)\}_{t\geq 0}$ is a Brownian motion under the equivalent martingale measure $\mathbb{Q}$. The specific derivations are given by Oosterlee and Grzelak [32]. The Brownian motion $\{W_r(t)\}_{t\geq 0}$ defined for the short rate dynamics and the Brownian motion $\{W_A(t)\}_{t\geq 0}$ corresponding to the process $\{X(t)\}_{t\geq 0}$ may be correlated. This correlation factor will be denoted by $\rho \in [-1,1]$ and is given such that

$$\mathrm{d}W_r(t)\mathrm{d}W_A(t) = \rho \mathrm{d}t.$$

2.2.2. Considerations concerning the calibration of the model

As discussed in Section 1.1.1, structural default models are difficult to calibrate due to the absence of historical data for the company value. Therefore, a short note on the calibration of the proposed model will be given. When the interest rate is assumed to be stochastic, the parameters of the interest rate process will be calibrated on market data independently of the company value part [17]. Then the company value parameters and the correlation parameter ρ need to be determined. When the correlation parameter ρ is known, the parameters of the company value process can be calibrated based on CDS market quotes and book values as discussed in Section 1.1.1. For instance, the method proposed by Fang et al. is referred to as a suitable calibration method [15, 3].

The parameter ρ cannot be inferred from the analysis between the company value and the interest rate, as no historical data is available. Therefore, ρ needs to be determined based on expert judgment and will be an imposed parameter. An alternative would be to only consider the equity part of the company value and use historical share prices to calibrate the correlation parameter ρ. However, note that this approach will contain an approximation error due to the capital structure of the company. A third option is to use an implied correlation parameter obtained via adding the correlation parameter to the set of company value parameters that will be calibrated using the CDS market quotes.

2.3. Derivation of the characteristic functions

The analysis of affine models was first proposed by Duffie, Pan, and Singleton and later extensively studied within the literature (see for example [32]). Affine models are commonly used models as they have useful properties. Let $\{\boldsymbol{X}(t)\}_{t\geq 0}$ be a multidimensional stochastic process, such that at each time step $t \in [t_0, T]$, the state of the process is given by the multidimensional random variable,

$$\boldsymbol{X}(t) = [r(t)\ X_0(t)\ X_1(t)\ \ldots\ X_{k-2}(t)]^{\mathsf{T}}, \tag{2.4}$$

for $k \in \mathbb{N}, k > 1$. Note that in the case $k = 1$, the Equation (2.4) reduces to the one-dimensional case and $k = 0$ is not defined. Suppose that the dynamics of the multidimensional stochastic process $\{\boldsymbol{X}(t)\}_{t\geq 0}$ are given by a system of stochastic differential equations,

$$\mathrm{d}\boldsymbol{X}(t) = \boldsymbol{\mu}(t, \boldsymbol{X}(t))\mathrm{d}t + \boldsymbol{\sigma}(t, \boldsymbol{X}(t))\mathrm{d}\bar{\boldsymbol{W}}(t), \tag{2.5}$$

for functions $\boldsymbol{\mu}(t, \boldsymbol{X}(t))$, $\boldsymbol{\sigma}(t, \boldsymbol{X}(t))$ and a vector of independent Brownian motions $\bar{\boldsymbol{W}}(t)$. Affine diffusion theory links the structure of the system of stochastic differential equations to the discounted characteristic function.

Definition 2.1. Let $f_{\boldsymbol{X}} : \mathbb{R}^k \to [0,\infty)$ be the transition density function for the process $\{\boldsymbol{X}(t)\}_{t\geq 0}$. The discounted characteristic function $\phi_{\boldsymbol{X}} : \mathbb{R}^k \to \mathbb{C}$ with initial date $t \in [t_0, T)$ corresponding to the system defined in Equation (2.5), is defined as

$$\phi_{\boldsymbol{X}}(u,t,T) = \mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\mathrm{d}s + \mathrm{i}u^{\mathsf{T}}\boldsymbol{X}(T)} \;\middle|\; \mathcal{F}_t\right] = \int_{\mathbb{R}^k} \mathrm{e}^{-\int_t^T r(s)\mathrm{d}s}\mathrm{e}^{\mathrm{i}u^{\mathsf{T}}x} f_{\boldsymbol{X}}(x)\mathrm{d}x,$$

where $u \in \mathbb{R}^k$ and $\boldsymbol{X}$ is defined as in Equation (2.4) [32].

In particular, the discounted characteristic function of affine models can be derived in terms of (complex) ordinary differential equations. For commonly used models, these differential equations are often solvable, which yields a closed-form solution for the discounted characteristic function. Hence, when using Fourier techniques, affine models may be used to approximate the value of a financial derivative through its closed-form characteristic function.

The system given in Equation (2.5) is in the class of affine diffusion models, if the functions $\boldsymbol{\mu}(t,\boldsymbol{X}(t))$ and $\boldsymbol{\sigma}(t,\boldsymbol{X}(t))$ satisfy the regularity conditions (see [32]) and in particular satisfy the affinity conditions which are given by [32]

$$\begin{aligned} \boldsymbol{\mu}(t,\boldsymbol{X}(t)) &= a_0 + a_1\boldsymbol{X}(t), \\ \boldsymbol{r}(t,\boldsymbol{X}(t)) &= r_0 + r_1^{\mathsf{T}}\boldsymbol{X}(t) \\ \left(\boldsymbol{\sigma}(t,\boldsymbol{X}(t))\boldsymbol{\sigma}(t,\boldsymbol{X}(t))^{\mathsf{T}}\right)_{ij} &= (c_0)_{ij} + (c_1)_{ij}^{\mathsf{T}}\boldsymbol{X}_j(t), \end{aligned} \tag{2.6}$$

for $r_0 \in \mathbb{R}$, $r_1, a_0 \in \mathbb{R}^k$, $a_1, c_0 \in \mathbb{R}^{k\times k}$ and $c_1 \in \mathbb{R}^{k\times k\times k}$. Furthermore, $\boldsymbol{r}$ is the interest rate process with respect to the discounting within the system. Consider a time horizon $[t_0, T]$ for $T \in \mathbb{R}_+$ and initial date $0 = t_0 < T$. In particular, if the system $\boldsymbol{X}$ is included in the class of affine diffusion models, the discounted characteristic function is known in closed form [32].

Theorem 2.1. *Assume a stochastic process $\{\boldsymbol{X}(t)\}_{t\geq 0}$ and corresponding system of stochastic differential equations as given in Equation* (2.5) *satisfying the regularity conditions and the affinity conditions given in Equation* (2.6)*. The discounted characteristic function corresponding to the process $\{\boldsymbol{X}(t)\}_{t\geq 0}$ is known in closed form and given by*

$$\phi_{\boldsymbol{X}}(u,t,T) = \mathrm{e}^{A(u,\tau)+\boldsymbol{B}^{\mathsf{T}}(u,\tau)\boldsymbol{X}(t)},$$

for $u \in \mathbb{R}^k,\ k \in \mathbb{N},\ k > 0$ and where $\tau = T - t$ is the time to maturity, $A : \mathbb{R}^{k+1} \to \mathbb{C}$ and $\boldsymbol{B}$ a vector of functions $B_j : \mathbb{R}^{k+1} \to \mathbb{C}$ for $j \in \{0,\ldots,m-1\}$, such that $\boldsymbol{B}(u,\tau) = [B_0(u,\tau)\ \ldots\ B_{m-1}(u,\tau)]^{\mathsf{T}}$. Moreover, the functions $A \equiv A(u,\tau)$ and $\boldsymbol{B} \equiv \boldsymbol{B}(u,\tau)$ satisfy the complex ordinary stochastic differential equations

$$\begin{aligned} \frac{dA}{d\tau} &= -r_0 + \boldsymbol{B}^{\mathsf{T}}a_0 + \frac{1}{2}\boldsymbol{B}^{\mathsf{T}}c_0B, \\ \frac{d\boldsymbol{B}}{d\tau} &= -r_1 + a_1^{\mathsf{T}}\boldsymbol{B} + \frac{1}{2}\boldsymbol{B}^{\mathsf{T}}c_1\boldsymbol{B}. \end{aligned}$$

Furthermore, the initial conditions of the system of complex ordinary differential equations satisfy $A(u,0) = 0$, $\boldsymbol{B}(u,0) = \mathrm{i}u$ for $u \in \mathbb{R}^k$.

A proof of Theorem 2.1 is given by Duffie, Pan, and Singleton [11]. The model, as described in Section 2.2, falls within the affine class. As a consequence, the (discounted) characteristic function of the model can be written in terms of ordinary (complex) differential equations.

2.3.1. The zero-coupon bond

The value $P(t,T)$ at time t of a zero coupon bond maturing at time T can be written in terms of the instantaneous short rate $\{r(t)\}_{t\geq 0}$, and is given by [32]

$$P(t,T) = \mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\mathrm{d}s} \;\middle|\; \mathcal{F}_t\right].$$

In particular, the value $P(t,T)$ of a zero-coupon bond can be written in closed form. This closed-form will be obtained by using a decomposition of the interest rate process $\{r(t)\}_{t\geq 0}$. The Hull-White

interest rate process can be decomposed in a deterministic function ψ and a stochastic process $\tilde{r}$. The decomposition into a deterministic and a simpler stochastic mean-reversion process will simplify the computations.

Lemma 2.1. *The interest rate process $\{r(t)\}_{t\geq 0}$ can be decomposed in a deterministic function ψ and a mean-reversion process $\{\tilde{r}(t)\}_{t\geq 0}$, which satisfy*

$$\begin{cases} r(t) & = \tilde{r}(t) + \psi(t), \\ \psi(t) & = r_0 \mathrm{e}^{-\kappa t} + \kappa \int_0^t \theta(z)\mathrm{e}^{-\kappa(t-z)}\mathsf{d}z, \\ \mathsf{d}\tilde{r}(t) & = -\kappa\tilde{r}(t)\mathsf{d}t + \eta \mathsf{d}W_r^{\mathbb{Q}}(t), \end{cases}$$

where η and κ are the diffusion and speed of mean reversion parameters in the Hull-White process defined in Equation (2.3)*. For the function ψ it holds that $\psi(0) = r_0$, the initial value of the Hull-White process at t_0. Furthermore, the initial condition of the mean-reversion process $\{\tilde{r}(t)\}_{t\geq 0}$ is given by $\tilde{r}(0) = 0$.*

Lemma 2.1 can be verified by Ito's lemma and a proof is given by Grzelak, Oosterlee, and Van Weeren [18]. As the decomposed interest rate process $\{\tilde{r}(t)\}_{t\geq 0}$ also satisfies the affinity conditions, Theorem 2.1 can be used to compute the discounted characteristic function in closed form.

Theorem 2.2. *The mean-reversion process $\{\tilde{r}(t)\}_{t\geq 0}$ is affine and therefore Theorem 2.1 applies. As a consequence, the discounted characteristic function corresponding to the process $\{\tilde{r}(t)\}_{t\geq 0}$ can be written in closed form, i.e.*

$$\tilde{\phi}_{\tilde{r}}(u,t,T) = \mathrm{e}^{A(u,\tau)+B(u,\tau)\tilde{r}(t)},$$

for $\tau = T - t$, where the complex functions $A : \mathbb{R}^2 \to \mathbb{C}$ and $B : \mathbb{R}^2 \to \mathbb{C}$ are given by

$$\begin{aligned} A(u,\tau) &= \frac{\eta^2}{2\kappa^3}\left(\kappa\tau - 2(1-\mathrm{e}^{-\kappa\tau}) + \frac{1}{2}(1-\mathrm{e}^{-2\kappa\tau})\right) - \mathrm{i}u\frac{\eta^2}{2\kappa^2}(1-\mathrm{e}^{-\kappa\tau})^2 \\ &\quad - u^2\frac{\eta^2}{4\kappa}(1-\mathrm{e}^{-2\kappa\tau}) \\ B(u,\tau) &= \mathrm{i}u\mathrm{e}^{-\kappa\tau} - \frac{1}{\kappa}(1-\mathrm{e}^{-\kappa\tau}), \end{aligned}$$

for $u \in \mathbb{R}$.

A proof of Theorem 2.2 is given by Grzelak, Oosterlee, and Van Weeren [18].

Using the result of Theorem 2.2 yields the tool for obtaining the value of the zero-coupon bond in closed form [32]. Firstly, the discounted characteristic function of the process $\{r(t)\}_{t\geq 0}$ is obtained by applying the decomposition given in Lemma 2.1 to the discounted characteristic function of $\{r(t)\}_{t\geq 0}$,

$$\begin{aligned} \tilde{\phi}_r(u,t,T) &= \mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\mathsf{d}s + \mathrm{i}ur(T)} \;\middle|\; \mathcal{F}_t\right] \\ &= \mathrm{e}^{-\int_t^T \psi(s)\mathsf{d}s + \mathrm{i}u\psi(T)}\mathbb{E}\left[\mathrm{e}^{-\int_t^T \tilde{r}(s)\mathsf{d}s + \mathrm{i}u\tilde{r}(T)} \;\middle|\; \mathcal{F}_t\right] \\ &= \mathrm{e}^{-\int_t^T \psi(s)\mathsf{d}s + \mathrm{i}u\psi(T) + A(u,\tau) + B(u,\tau)\tilde{r}(t)} \\ &= \mathrm{e}^{-\int_t^T \psi(s)\mathsf{d}s + \mathrm{i}u\psi(T)}\tilde{\phi}_{\tilde{r}}(u,t,T), \end{aligned}$$

under the equivalent martingale measure $\mathbb{Q}$, where $\tau = T - t$ and $\tilde{\phi}_{\tilde{r}}$ is the discounted characteristic function of the process $\{\tilde{r}(t)\}_{t\geq 0}$ as given in Theorem 2.2. Hence, for $u = 0$, the discounted characteristic function of the instantaneous forward rate is equal to the value of a zero-coupon bond at time t, i.e. $\tilde{\phi}_r(0,t,T) = P(t,T)$. As a direct result of Theorem 2.2, the value of a zero-coupon bond maturing at time T can be expressed in a closed-form function,

$$P(t,T) = \tilde{\phi}_r(0,t,T) = \mathrm{e}^{-\int_t^T \psi(s)\mathsf{d}s + A(\tau) - B(\tau)\tilde{r}(t)}, \tag{2.7}$$

where

$$\begin{aligned} A(\tau) &= \frac{\eta^2}{2\kappa^3}\left(\kappa\tau - 2(1-\mathrm{e}^{-\kappa\tau}) + \frac{1}{2}(1-\mathrm{e}^{-2\kappa\tau})\right) \\ B(\tau) &= \frac{1}{\kappa}(1-\mathrm{e}^{-\kappa\tau}), \end{aligned} \tag{2.8}$$

such that $A(\tau) \equiv A(0,\tau)$ and $B(\tau) = -B(0,\tau)$.

2.3.2. Characteristic function of the Black-Scholes Hull-White model

The discounted characteristic function of the system, as given in Definition 2.1, is defined as

$$\phi_{r,X}(u,w,t,T) = \mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\mathrm{d}s + \mathrm{i}(ur(T)+wX(T))} \;\middle|\; \mathcal{F}_t\right], \tag{2.9}$$

As the jump part is independent of the continuous part, the discounted characteristic function in Equation (2.9) can be written as the multiplication between the discounted characteristic function corresponding to the diffusion part and the characteristic function corresponding to the jump process, i.e.

$$\begin{aligned} \phi_{r,X}(u,w,t,T) &= \mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\mathrm{d}s + \mathrm{i}(ur(T)+wX^C(T))} \;\middle|\; \mathcal{F}_t\right]\mathbb{E}\left[\mathrm{e}^{\mathrm{i}wX^J(T)} \;\middle|\; \mathcal{F}_t\right] \\ &= \phi_{r,X^C}(u,w,t,T)\phi_{X^J}(u,w,t,T), \end{aligned} \tag{2.10}$$

where $X^C(T)$ denotes the value of the diffusion process $\left\{X^C(t)\right\}_{t\geq 0}$ at time T and $X^J(T)$ is used to refer to the value of the jump process $\left\{X^J(t)\right\}_{t\geq 0}$ at time T as defined in Section 2.2. In this section, the characteristic function of the diffusion part will be derived. To obtain the full discounted characteristic function of the process, including the jumps, only a multiplication is required with the independent jump component, as described in Equation (2.10). The characteristic function of the jump component will be discussed in Section 2.3.4.

Using only the diffusion part, the model can be written in matrix form, where the correlation matrix is decomposed using the Cholesky decomposition to correlate the independent Brownian motions $\left\{\bar{W}_r(t)\right\}_{t\geq 0}$ and $\left\{\bar{W}_A(t)\right\}_{t\geq 0}$ under the equivalent martingale measure,

$$\begin{bmatrix} \mathrm{d}r(t) \\ \mathrm{d}X^C(t) \end{bmatrix} = \begin{bmatrix} \kappa(\theta(t)-r(t)) \\ r(t) - \frac{1}{2}\sigma^2 \end{bmatrix}\mathrm{d}t + \begin{bmatrix} \eta & 0 \\ \sigma\rho & \sigma\sqrt{1-\rho^2} \end{bmatrix}\begin{bmatrix} \mathrm{d}\bar{W}_r(t) \\ \mathrm{d}\bar{W}_A(t) \end{bmatrix}, \tag{2.11}$$

where $r(t)$ is the value of the interest rate process at time t, κ and η are the speed of mean reversion and the diffusion parameters in the Hull-White process defined in Equation (2.3) respectively, θ is the mean reversion function defined in Equation (2.2) and ρ and σ are the correlation and volatility parameter respectively. As the diffusion part of the model is included in the affine class, Theorem 2.1 can be applied.

Theorem 2.3. *The system given in Equation* (2.11) *is included in the affine class of diffusion models. As a consequence, the discounted characteristic function corresponding to the system is given by*

$$\phi_{r,X^C}(u,w,t,T) = \mathrm{e}^{A_1(u,w,\tau)+A_2(u,w,\tau)r(t)+\mathrm{i}wX^C(t)}. \tag{2.12}$$

where $\tau = T-t$ is the time until maturity and $u, w \in \mathbb{R}$. Furthermore, $A_1 : \mathbb{R}^3 \to \mathbb{C}$, $A_2 : \mathbb{R}^3 \to \mathbb{C}$ and $A_3 = \mathrm{i}w$ are functions which satisfy the complex ordinary differential equations given in Equation (2.5).

Solving the corresponding complex ordinary differential equations, yields

$$\begin{aligned}
A_1(u,w,\tau) &= (\kappa \mathrm{i}u - (\mathrm{i}w-1)) \int_0^\tau \theta(T-z)\mathrm{e}^{-\kappa z}\mathsf{d}z + (\mathrm{i}w-1)\int_t^T \theta(s)\mathsf{d}s \\
&\quad + \frac{1}{2}\sigma^2 \mathrm{i}w(\mathrm{i}w-1)\tau \\
&\quad + \frac{\eta^2}{2}\left[-\frac{u^2}{2\kappa}(1-\mathrm{e}^{-2\kappa\tau}) + \frac{\mathrm{i}u(\mathrm{i}w-1)}{\kappa^2}\left(1+\mathrm{e}^{-2\kappa\tau}-2\mathrm{e}^{-\kappa\tau}\right)\right. \\
&\quad \left. + \frac{(\mathrm{i}w-1)^2}{2\kappa^3}\left(2\kappa\tau + 4\mathrm{e}^{-\kappa\tau} - \mathrm{e}^{-2\kappa\tau} - 3\right)\right] \\
&\quad + \frac{\eta\sigma\rho\mathrm{i}w}{\kappa^2}\left[(\mathrm{i}w-1)\left(\kappa\tau + \mathrm{e}^{-\kappa\tau} - 1\right) - \kappa\mathrm{i}u(1-\mathrm{e}^{-\kappa\tau})\right], \\
A_2(u,w,\tau) &= \mathrm{i}u\mathrm{e}^{-\kappa\tau} + \frac{(\mathrm{i}w-1)}{\kappa}(1-\mathrm{e}^{-\kappa\tau}).
\end{aligned}$$

A proof of Theorem 2.3 is given in Appendix D.1. Note that the function A_2 can be written in terms of the function B, given in Equation (2.8), i.e.

$$A_2(u,w,t) = \mathrm{i}u\mathrm{e}^{-\kappa\tau} + (w\mathrm{i}-1)B(\tau). \tag{2.13}$$

To be able to apply a zero-centered grid, it is more convenient to use a process that is more centered around zero. Therefore, also the characteristic function under the decomposed interest rate process $\{\tilde{r}(t)\}_{t\geq 0}$ will be derived. The corresponding system of stochastic differential equations based on this decomposed short rate process $\{\tilde{r}(t)\}_{t\geq 0}$, yields,

$$\begin{bmatrix} \mathsf{d}\tilde{r}(t) \\ \mathsf{d}X^C(t) \end{bmatrix} = \begin{bmatrix} \kappa\tilde{r}(t) \\ \tilde{r}(t) + \psi(t) - \frac{1}{2}\sigma^2 \end{bmatrix} \mathsf{d}t + \begin{bmatrix} \eta & 0 \\ \sigma\rho & \sigma\sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} \mathsf{d}\bar{W}_r(t) \\ \mathsf{d}\bar{W}_A(t) \end{bmatrix}, \tag{2.14}$$

for the process $\{\tilde{r}(t)\}_{t\geq 0}$ and the function ψ defined in Lemma 2.1. Also, the decomposed system as given in Equation (2.14) is included in the affine class of diffusion processes.

Theorem 2.4. *The system in Equation* (2.14) *is included in the affine class of diffusion processes. As a consequence, the discounted characteristic function of the system is given by*

$$\phi_{\tilde{r}(t),X^C}(u,w,t,T) = \mathrm{e}^{\bar{A}(u,w,\tau)+A_2(u,w,\tau)\tilde{r}(t)+\mathrm{i}wX^C(t)},$$

where A_2 is given in Equation (2.13) *and $\bar{A}$ is given by*

$$\begin{aligned}
\bar{A}(u,w,\tau) &= \mathrm{i}w\int_0^\tau \psi(T-z)\mathsf{d}z + \frac{1}{2}\sigma^2\mathrm{i}w(\mathrm{i}w-1)\tau \\
&\quad + \frac{\eta^2}{2}\left[\frac{(\mathrm{i}w-1)^2}{2\kappa^3}\left(2\tau\kappa + 4\mathrm{e}^{-\kappa\tau} - \mathrm{e}^{-2\kappa\tau} - 3\right) + \frac{(\mathrm{i}w-1)\mathrm{i}u}{2\kappa^2}\left[\mathrm{e}^{-2\kappa\tau} - 2\mathrm{e}^{-\kappa\tau} + 1\right]\right. \\
&\quad \left. - \frac{(\mathrm{i}u)^2}{2\kappa}(\mathrm{e}^{-2\kappa\tau}-1)\right] + \eta\sigma\rho\mathrm{i}w\left[\frac{\mathrm{i}w-1}{\kappa^2}(\kappa\tau + \mathrm{e}^{-\kappa\tau} - 1) - \frac{\mathrm{i}u}{\kappa}(\mathrm{e}^{-\kappa\tau}-1)\right],
\end{aligned} \tag{2.15}$$

where $u, w \in \mathbb{R}$ and the integral $\int_0^\tau \psi(T-z)\mathsf{d}z$ can be computed numerically.

A proof of Theorem 2.4 is given in Appendix D.2. Applying the decomposition to the discounted characteristic function under the mean reversion process $\{r(t)\}_{t\geq 0}$ yields

$$\begin{aligned}
\mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\mathsf{d}s + \mathrm{i}(ur(T)+wX^C(T))} \;\middle|\; \mathcal{F}_t\right] &= \mathrm{e}^{-\int_t^T \psi(s)\mathsf{d}s + \mathrm{i}u\psi(T)}\mathbb{E}\left[\mathrm{e}^{-\int_t^T \tilde{r}(s)\mathsf{d}s + \mathrm{i}(u\tilde{r}(T)+wX^C(T))} \;\middle|\; \mathcal{F}_t\right] \\
&= \mathrm{e}^{-\int_t^T \psi(s)\mathsf{d}s + \mathrm{i}u\psi(T)}\phi_{\tilde{r}(t),X^C}(u,w,t,T),
\end{aligned}$$

which gives a zero-centered approach for computing the characteristic function under the system given in Equation (2.11).

2.3.3. Characteristic function under a variable transformation

One disadvantage of the form presented in Equation (2.12), is the requirement of the state variables $r(t)$ (or $\tilde{r}(t)$) and $X^C(t)$ to compute the discounted characteristic function. At time $t_0 = 0$, the values of $r(t_0)$ and $X^C(t_0)$ are known, however values $r(t)$ and $X^C(t)$ for $t_0 < t \leq T$ remain stochastic quantities. Within the valuation of early exercise products, exercise decisions need to be made for exercising moments between the initial date of valuation t_0 and maturity T. Hence, the form presented in Equation (2.12) cannot be used, as there is no information about the state variables within the interval $(t_0, T]$. Therefore, through a transformation, a stateless form will be derived. Ballotta and Kyriakou show that it is beneficial to first change the measure under which the characteristic function $\phi_{\tilde{r},X^C}$ is computed. The resulting expression contains a conditional expectation under the T-forward measure, multiplied with a factor representing the value of a zero-coupon bond. This zero-coupon bond value can be used to define a variable transformation under which the characteristic function is free of the state variables.

First, an expression for the characteristic function under the T-forward measure $\mathbb{Q}^T$, of the system of stochastic differential equations given in Equation (2.14) will be given.

Lemma 2.2. *The characteristic function of the system given in Equation* (2.14) *with respect to the filtration $\mathcal{F}_t$ under the $T-$forward measure, denoted as $\mathbb{E}^T[\cdot \mid \mathcal{F}_t]$, is given by*

$$\mathbb{E}^T\left[\mathrm{e}^{\mathrm{i}\left(u\tilde{r}(T)+wX^C(T)\right)} \;\middle|\; \mathcal{F}_t\right] = \frac{\mathrm{e}^{-\int_t^T \psi(s)\mathrm{d}s}\mathbb{E}\left[\mathrm{e}^{-\int_t^T \tilde{r}(s)\mathrm{d}s+\mathrm{i}\left(u\tilde{r}(T)+wX^C(T)\right)} \;\middle|\; \mathcal{F}_t\right]}{P(t,T)},$$

where $P(t,T)$ is the value of zero-coupon bond at time t maturing at time T and ψ is the deterministic function given in Lemma 2.1.

The proof of Lemma 2.2 is given in Appendix D.3. Lemma 2.2 yields a direct relation between the measure $\mathbb{Q}$ and the $T-$forward measure $\mathbb{Q}^T$, which can be exploited to obtain an expression for the characteristic function under the T-forward measure. The results can be split into a deterministic part and a stochastic part, this result is stated in Lemma 2.3.

Lemma 2.3. *The characteristic function of the system given in Equation* (2.14) *under the $T-$forward measure is known in closed form and given by*

$$\mathbb{E}^T\left[\mathrm{e}^{\mathrm{i}\left(u\tilde{r}(T)+wX^C(T)\right)} \;\middle|\; \mathcal{F}_t\right] = \mathrm{e}^{-A(\tau)+\bar{A}(u,w,\tau)}\mathrm{e}^{\mathrm{i}ue^{-\kappa\tau}\tilde{r}(t)+\mathrm{i}w(X^C(t)+B(\tau)\tilde{r}(t))},$$

where the function $\tilde{A}$ is given in Theorem 2.4 and,

$$A(\tau) = \frac{\eta^2}{2\kappa^3}\left(\kappa\tau - 2(1-\mathrm{e}^{-\kappa\tau}) + \frac{1}{2}(1-\mathrm{e}^{-2\kappa\tau})\right),$$

as derived in Section 2.3.1.

The proof of Lemma 2.3 is given in Appendix D.4. Using both the stochastic and the deterministic part, a transformation can be defined under which the closed-form expression of its characteristic function is free of the state variables $\tilde{r}(t)$ and $X^C(t)$. The transformation and the resulting characteristic function under the T-forward measure are summarized in Theorem 2.5.

Theorem 2.5. *Let the functions g_r and g be given as*

$$g_r(x) = x\mathrm{e}^{-\kappa\tau},\; g(x,y) = y + B(\tau)x,$$

where $B(\tau) = \frac{1}{\kappa}(1-\mathrm{e}^{-\kappa\tau})$ and $\tau = T - t$. Furthermore, define the variable transformation (Z_r, Z) as

$$\begin{cases} Z_r(T) &= \tilde{r}(T) - g_r(\tilde{r}(t)) \\ Z(T) &= X^C(T) - g(\tilde{r}(t), X^C(t)) \end{cases}$$

Then the characteristic function of the stochastic variables (Z_r, Z) under the $T-$forward measure,

$$\phi_{Z_r,Z}(u,w,t,T) = \mathbb{E}^T\left[\mathrm{e}^{\mathrm{i}(uZ_r(T)+wZ(T))} \;\middle|\; \mathcal{F}_t\right],$$

is known in closed form and free of any state variables. In particular, the characteristic function $\phi_{Z_r,Z}(u,w,t,T)$ *is given by*

$$\phi_{Z_r,Z}(u,w,t,T) = \mathrm{e}^{\tilde{A}(u,w,\tau)-A(\tau)}, \tag{2.16}$$

where $u,w \in \mathbb{R}$ *and* $\tau = T-t$.

The proof of Theorem 2.5 is given in Appendix D.5. Equation (2.16) shows that the transformed characteristic function under the T-forward measure is known and can be given free of the state variables $\tilde{r}(T)$ and $X^C(T)$. The resulting expression can be exploited in pricing problems where the state variables for $\tilde{r}(t)$ and $X^C(t)$, for $t_0 < t < T$, are required but are not known.

2.3.4. Characteristic function of the jump-diffusion model

As discussed, by assumption, the jump and the continuous parts of the process are mutually independent. Due to this independence assumption, the characteristic function could be written in the form,

$$\phi_{r,X}(u,w,t,T) = \phi_{r,X^C}(u,w,t,T)\mathbb{E}\left[\mathrm{e}^{\mathrm{i}wX^J(T)} \;\middle|\; \mathcal{F}_t\right] = \phi_{r,X^C}(u,w,t,T)\phi_{r,X^J}(w,t,T),$$

where the continuous discounted characteristic function can be written under a variable transformation such that it no longer depends on any of the state variables. It remains therefore to find the characteristic function of the transformation under the jump-diffusion process. Note that the characteristic function for the jumps can be written as [32]

$$\phi_{r,X^J}(w,t,T) = \mathbb{E}\left[\mathrm{e}^{\mathrm{i}w\left(-\lambda t\mathbb{E}\left[\mathrm{e}^J-1\right]+\sum_{k=1}^{\mathcal{P}(t)}J_k\right)} \;\middle|\; \mathcal{F}_t\right] = \mathrm{e}^{-\lambda t\left(\mathbb{E}\left[\mathrm{e}^J-1\right]-\mathbb{E}\left[\mathrm{e}^{\mathrm{i}wJ}-1\right]\right)}.$$

Adding jumps to the model is therefore restricted on the choice of $\mathbb{E}\left[\mathrm{e}^J\right]$ and $\mathbb{E}\left[\mathrm{e}^{\mathrm{i}wJ}\right]$.

For both the Merton model as the Kou and Wang model, the characteristic function under the jump-diffusion model can be derived. The specific functions for the expectations $\mathbb{E}\left[\mathrm{e}^J\right]$ and $\mathbb{E}\left[\mathrm{e}^{\mathrm{i}wJ}\right]$ are included in Table 2.1. Using the expectations given in Table 2.1 for the jump part of the characteristic function ϕ_{r,X^J} multiplied with the continuous part of the transformed characteristic function $\phi_{Z_r,Z}$, the transformed characteristic function for the model defined in Section 2.2 under the transformation can be obtained.

Model name	$\mathbb{E}\left[\mathrm{e}^J\right]$	$\mathbb{E}\left[\mathrm{e}^{\mathrm{i}wJ}\right]$
Merton	$\mathrm{e}^{\mu_J+\frac{1}{2}\sigma_J^2}$	$\mathrm{e}^{\mathrm{i}w\mu_J-\frac{1}{2}w^2\sigma_J^2}$
Kou and Wang	$\mathrm{e}^{p_1\frac{\alpha_1}{\alpha_1-1}+p_2\frac{\alpha_2}{\alpha_2+1}}$	$\mathrm{e}^{p_1\frac{\alpha_1}{\alpha_1-\mathrm{i}w}+p_2\frac{\alpha_2}{\mathrm{i}w+\alpha_2}}$

Table 2.1: This table shows the model specifications of the two main used models for jump-diffusion processes within the literature. The model proposed by Merton, assumes normal distributed jump sizes with mean μ_J and variance σ_J^2. The model proposed by Kou and Wang assumes jump sizes that are non-symmetric double exponentially distributed, with upward probability $p_1 \in [0,1]$ and downward probability $p_2 = 1-p_1$ [22]. The α_1 and α_2 parameters represent the intensities of the individual exponential distributions [23].

3

Monte Carlo methods for callable convertible bond approximations

3.1. Introduction

The thesis will present two ways of valuing the callable convertible bond. In this chapter, Monte Carlo (MC) techniques will be discussed. the bond value at time t_0 can be approximated through a dynamic programming approach using simulated paths for all of the risk factors. First, the payoff at maturity T will be computed. Subsequently, the bond value at each premature time step can be obtained using the already obtained future values of the bond. By the characteristics of the callable convertible bond, at each premature time step, the exercising decisions can be made using a least-squares regression technique [26].

The chapter will start with a discussion on the simulation of sample paths of the underlying risk factors $\{r(t)\}_{t\geq 0}$ and $\{A(t)\}_{t\geq 0}$. The MC techniques will be explained following the complexity of the bond products. First, MC techniques will be applied to the less complex convertible bonds without a call option and then to the more complex callable convertible bonds that include a call option. The valuation of the convertible bond will illustrate how MC techniques may be applied. The basic notion of the convertible bond will give rise to the applicability of MC techniques to the more complex callable convertible bond product.

3.2. Simulation of the model

Recall the diffusion part of the system described in Chapter 2 and given in Equation (2.11). The system given in Equation (2.11) can be simulated in two steps. First, the interest rate process is simulated, using independent realizations of the Brownian motion. Consider an equally spaced time grid $t_l = l\Delta_t$ of M nodes over the time horizon given by $[t_0, T]$, for $l \in \{0, \ldots, \mathrm{M}-1\}$ with $\mathrm{M} \in \mathbb{N}$, $\Delta_t = \frac{T}{\mathrm{M}-1}$ and such that $t_0 = 0$ and $t_{\mathrm{M}-1} = T$. The stochastic process for the short rate is given by

$$\mathsf{d}r(t) = \kappa(\theta(t) - r(t))\mathsf{d}t + \eta \mathsf{d}\bar{W}_r(t), \tag{3.1}$$

where κ is the speed of mean reversion, η is the volatility of the interest rate process, θ is the mean reversion function as given in Equation (2.2), and for a Brownian motion $\left\{\bar{W}_r(t)\right\}_{t\geq 0}$ under the equivalent martingale measure $\mathbb{Q}$. Assume that the initial value of the process is given by $r(t_0) = r_0$. Denote the value $r_l \equiv r(t_l)$ for t_l on the time grid. Simulations will be based on numerical integration. An increment $t_l + \Delta_t$ of the short rate, following the stochastic differential equation given in Equation (3.1) for known value r_l is given by

$$r_{l+1} \approx r_l + \kappa \int_{t_l}^{t_{l+1}} \theta(s) - r(s)\mathsf{d}s + \eta(\bar{W}_r(t_{l+1}) - \bar{W}_r(t_l)). \tag{3.2}$$

Increments of a Brownian motion $\{W(t)\}_{t\geq 0}$ over a time interval $[t_l, t_{l+1}]$ are independently normally distributed with mean zero and variance $t_{l+1} - t_l$, i.e. [32]

$$W(t_{l+1}) - W(t_l) \sim \mathcal{N}\left(0, t_{l+1} - t_l\right).$$

Using an Euler integration scheme for the integral presented in Equation (3.2) and using the normal distribution for the simulation of the Brownian increment, the following recursive simulation scheme can be derived for the simulation of the short rate;

$$r_{l+1} \overset{\mathsf{d}}{\approx} r_l + \kappa(\theta(t_l) - r_l)\Delta_t + \eta\sqrt{\Delta_t} Z_l^r, \tag{3.3}$$

where Z_l^r is a realization of the standard normal distribution.

The second step in simulating the system given in Equation (2.11), is simulating the log-transform of the company value. Assume that the initial company value is given as $A(t_0) = A_0$, such that its log-transform has initial value, $X(t_0) = \log(A_0)$. Recall the stochastic differential equation for the transformed company value, which includes stochastic jumps,

$$\mathsf{d}X(t) = \left(r(t) - \lambda\mathbb{E}\left[\mathrm{e}^J - 1\right] - \frac{1}{2}\sigma^2\right)\mathsf{d}t + \sigma\left(\rho\mathsf{d}\bar{W}_r(t) + \sqrt{1-\rho^2}\mathsf{d}\bar{W}_A(t)\right) + J\mathsf{d}\mathcal{P}(t), \tag{3.4}$$

where σ is the volatility, ρ is the correlation, λ is the intensity of the jump process, independent Brownian motions $\left\{\bar{W}_r(t)\right\}_{t\geq 0}$ and $\left\{\bar{W}_A(t)\right\}_{t\geq 0}$, an independent stochastic jump-size J and the Poisson process $\{\mathcal{P}(t)\}_{t\geq 0}$ independent from the jump sizes J and Brownian motion $\left\{\bar{W}_A(t)\right\}_{t\geq 0}$. Note that the short rate process can be simulated following the scheme as given in Equation (3.3). Denote the value of $X(t_l) = X_l$ for the time step t_l on the time grid. Assume that the values r_l and X_l are known. The integral formulation of the stochastic differential equation given in Equation (3.4) can be written as

$$\begin{aligned} X_{l+1} \approx X_l + \int_{t_l}^{t_{l+1}} & r(s) - \lambda\mathbb{E}\left[\mathrm{e}^J - 1\right] - \frac{1}{2}\sigma^2\mathsf{d}s \\ & + \sigma\left(\rho(\bar{W}_r(t_{l+1}) - \bar{W}_r(t_l)) + \sqrt{1-\rho^2}(\bar{W}_A(t_{l+1}) - \bar{W}_A(t_l))\right) \\ & + J(\mathcal{P}(t_{l+1}) - \mathcal{P}(t_l)). \end{aligned}$$

Using an Euler integration scheme, combined with the standard normal distributed Brownian increments, a simulation scheme can be derived. Let Z_l^A be a realization of a standard normal distribution, which is independent of Z_l^r. The Poisson process has independent Poisson distributed increments with intensity $\lambda\mathsf{d}t$, i.e.

$$\mathcal{P}(t_{l+1}) - \mathcal{P}(t_l) \sim \text{Poisson}\left(\lambda\mathsf{d}t\right).$$

Hence, let P_l be a realization of a Poisson distribution with intensity parameter $\lambda\mathsf{d}t$. Furthermore, given the distribution for the jump sizes J, let $\bar{J}$ be a realization of the jump size. The simulation scheme for the company value under the log dynamics, adjusted with jumps and using an Euler integration scheme is then given by

$$X_{l+1} \overset{\mathsf{d}}{\approx} X_l + \left(r_l - \lambda\mathbb{E}\left[\mathrm{e}^J - 1\right] - \frac{1}{2}\sigma^2\right) + \sigma\left(\rho Z_l^r + \sqrt{1-\rho^2}Z_l^A\right) + \bar{J}P_l. \tag{3.5}$$

The value for the expectation $\mathbb{E}\left[\mathrm{e}^J\right]$ depends on the chosen model. For the popular models proposed by Merton and Kou and Wang, the function corresponding to the expected value is included in Table 2.1 in Chapter 2. In many software packages, no simulation algorithm for the double exponential model presented by Kou and Wang is given. Hence, a sampling algorithm based on the uniform distribution is given in Appendix B.1.2. This sampling algorithm can be used to sample the company value under the jump-diffusion model proposed by Kou and Wang. Henceforth, the simulation schemes in Equation (3.5) and in Equation (3.3) can be used to simulate sample paths for the company value and the interest rate respectively.

3.3. Valuation of convertible bonds

The valuation of a convertible bond can be done very straightforwardly using MC techniques. Consider the time grid, defined in Section 3.2, i.e. $t_l = l\Delta_t$ for $\Delta_t = \frac{T}{\mathrm{M}-1}$ with maturity T. Make sure the coupon payment dates $\tau_j \in \mathcal{T}^{\text{coup}}$ coincide with the time grid such that $t_0 < \tau_1$ and $\tau_{n-1} = T$. Simulate N paths for the process $\{A(t)\}_{t\geq 0}$ and $\{r(t)\}_{t\geq 0}$ of $\mathrm{M}-1$ nodes respectively. In particular, the values for the process $\{A(t)\}_{t\geq 0}$ can be simulated by simulating paths for the log-transform $\{X(t)\}_{t\geq 0}$, $\{X_{l,k} \mid l \in \{0,\ldots,\mathrm{M}-1\}, k \in \{0,\ldots,\mathrm{N}-1\}\}$, and then use the transformation $A_{l,k} = \mathrm{e}^{X_{l,k}}$ to obtain realizations of the process $\{A(t)\}_{t\geq 0}$. Denote the realization of the company value and the short rate process at time t_l, along path $k \in \{0,\ldots,\mathrm{N}-1\}$ by $A_{l,k}$ and $r_{l,k}$ respectively. Let the fixed coupons be given as

$$c_j = \alpha F, \ \forall j,$$

for face value F and fixed coupon rate α. Note that the default can be determined at each time step. Let $\gamma \in (0,1]$ be the predetermined conversion rate. As discussed in Section 1.2.2, the optimal exercise moment for a convertible bond is at maturity T.

The value of the bond $V_{\text{conv}}(t_{\mathrm{M}-1}, A_{\mathrm{M}-1,k})$ at maturity $T = t_{\mathrm{M}-1}$ along path k is equal to the payoff of the convertible bond at maturity T, which is given by

$$H(t_{\mathrm{M}-1}, A_{\mathrm{M}-1,k}) = \min(A_{\mathrm{M}-1,k}, \max((1+\alpha)F, \gamma A_{\mathrm{M}-1,k})), \ \forall k.$$

Each premature value of the convertible bond along each path k, $V(t_l, A_{l,k})$ for $l \in \{1,\ldots,\mathrm{M}-2\}$, will be obtained by looping backwards in time. For each time step t_l for $l < \mathrm{M}-1$ and for each path $k \in \{0,\ldots,\mathrm{N}-1\}$, the payoff of the bond is,

$$H(t_l, A_{l,k}) = \min(\alpha F, A_{l,k}).$$

The value of the bond, however, also will take future payments into consideration. Hence, the value of the bond $V_{\text{conv}}(t_l, A_{l,k})$ for time step $t_0 < t_l < T$ will be

$$V_{\text{conv}}(t_l, A_{l,k}) = \begin{cases} \alpha F + \mathrm{e}^{-r_{l,k}\Delta_t} V_{\text{conv}}(t_{l+1}, A_{l+1,k}) & \text{for } \alpha F \geq A_{l,k}, \\ A_{l,k} & \text{for } \alpha F < A_{l,k}, \end{cases}$$

in case $t_l \in \mathcal{T}^{\text{coup}}$ and $V(t_l, A_{l,k}) = \mathrm{e}^{-r_{l,k}\Delta_t} V(t_{l+1}, A_{l+1,k})$ when $t_l \notin \mathcal{T}^{\text{coup}}$. At the initial time step t_0, the value of the convertible bond $V(t_0, A(t_0))$ can be approximated by taking the discounted mean over all the obtained convertible bond values along each path following the relation,

$$V_{\text{conv}}(t_0, A(t_0)) = \mathbb{E}\left[\mathrm{e}^{-\int_{t_0}^{t_1} r(s)\mathrm{d}s} V_{\text{conv}}(t_1, A(t_1)) \;\middle|\; \mathcal{F}_{t_0} \right].$$

3.4. Valuation of callable convertible bonds

Consider a callable convertible option embedded within a corporate bond. Assume that the call price is given by $K(t)$, which is a deterministic function of time and can be decomposed as $K(t) = \beta(t)F$ for a penalty function β and face value F. Consider the time grid, defined in Section 3.2, i.e. $t_l = l\Delta_t$ for $\Delta_t = \frac{T}{\mathrm{M}-1}$ with maturity T. Furthermore, let both the coupon dates $\mathcal{T}^{\text{coup}}$ and the exercise dates of the call option $\mathcal{T}^{\text{call}}$ coincide with the defined time nodes on the grid. It will be assumed that $\mathcal{T}^{\text{coup}} = \mathcal{T}^{\text{call}} = \{\tau_1,\ldots,\tau_{n-1}\}$ and that each date coincides with the time grid. However, to possibly account for a call protection period, also other configurations are possible. Consider the approach for simulating values of the underlying risk factors $\{A(t)\}_{t\geq 0}$ and $\{r(t)\}_{t\geq 0}$ as discussed in Section 3.3 and denote $A_{l,k}$ and $r_{l,k}$ for the realization of $A(t_l)$ and $r(t_l)$ along path k respectively. In addition, let α be the fixed coupon rate and assume a conversion rate $\gamma \in (0,1]$.

Write $V_{\text{cc}}(t, A(t))$ for the value of the callable convertible bond at time t. By definition, the callable convertible bond cannot be called at maturity T, hence its value at maturity $t_{\mathrm{M}} - 1 = T$ is equal to the payoff of a convertible bond at maturity, i.e.

$$V_{\text{cc}}(t_{\mathrm{M}-1}, A_{\mathrm{M}-1,k}) = H_{\text{conv}}(t_{\mathrm{M}-1}, A_{\mathrm{M}-1,k}) = \min(A_{\mathrm{M}-1,k}, \max((1+\alpha)F, \gamma A_{\mathrm{M}-1,k})),$$

along al paths $k \in \{0, \dots, \mathrm{N}-1\}$. The value $V_{\mathsf{cc}}(t_l, A_{l,k})$ of the callable convertible bond for each time $t_0 < t_l < T$ along path k can be determined by looping backwards in time. If the time step t_l does not coincide with a coupon payment, the value of the bond will be set to the continuation value of the bond, i.e.

$$V_{\mathsf{cc}}(t_l, A_{l,k}) = C(t_l, A(l,k)),$$

for C as defined in Definition 1.3. Subsequently assume that the time step $t_l \in \mathcal{T}^{\mathsf{coup}}$. At each premature time step t_l, each party within the contract needs to decide whether or not to exercise their option. First, the management of the company needs to decide whether or not to call the bond. Secondly, based on the decision of the management of the company, the bondholder needs to decide whether or not to convert the bond.

Recall the continuation values $C(t_l, A(t_l))$ and $C_{\mathsf{call}}(t_l, A(t_l))$ of the bond as defined in Definition 1.4 and Definition 1.5 respectively. These continuation values define the optimal exercise policies for both parties within the contract. As the coupon payment dates and the call dates coincide, the accrued interest on the coupon payments equals the coupon payment αF. As discussed in Section 1.2.4, the optimal exercise policy for the management of the company will therefore be,

$$C_{\mathsf{call}}(t_l, A(t_l)) > (\beta(t_l) + \alpha)F.$$

The bondholder will have the objective to maximize her wealth. Therefore, if the management of the company calls the bond, the bondholder will accept the call price increased with the accrued interest only if it exceeds the conversion value. On the other hand, if the management of the company decides not to call the bond, the bondholder will compare the continuation value of the bond to the payoff at immediate exercise. The bondholder will then follow the optimal conversion policy as discussed in Section 1.2.4 and exercise the conversion option if and only if

$$C(t_l, A(t_l)) \leq \gamma A(t_l),$$

where γ is the conversion factor and $C(t_l, A(t_l))$ is the continuation value of the bond given at time t_l and defined in Definition 1.4.

Using the continuation values for both parties within the contract, the value of the bond $V_{\mathsf{cc}}(t_l, A_{l,k})$ at time t_l along path k can be determined,

$$V_{\mathsf{cc}}(t_l, A_{l,k}) = \begin{cases} A_{l,k} & \text{for } A_{l,k} < \alpha F, \\ \max((\beta(t_l)+\alpha)F, \gamma A_{l,k}) & \text{for } C_{\mathsf{call}}(t_l, A_{l,k}) > (\beta(t_l)+\alpha)F \\ & \qquad \text{and } A_{l,k} \geq \alpha F, \\ \gamma A_{l,k} & \text{for } C(t_l, A_{l,k}) < \gamma A_{l,k} \text{ and } A_{l,k} \geq \alpha F, \\ \alpha F + \mathrm{e}^{-r_{l,k}\Delta_t} V_{\mathsf{cc}}(t_{l+1}, A_{l,k+1}) & \quad \text{for } (C_{\mathsf{call}}(t_l, A_{l,k}) \leq (\beta(t_l)+\alpha)F \\ & \text{or } C(t_l, A_{l,k}) \geq \gamma A_{l,k}) \text{ and } A_{l,k} \geq \alpha. \end{cases}$$

At the initial time t_0, no coupon payments are paid and there will not be any exercise opportunities. Hence, the value $V(t_0, A_{0,k})$ of the bond at time t_0 along path k will be the discounted value of the bond at time t_1. An approximation for the convertible bond value can therefore be obtained, by taking the discounted mean using all M paths of the bond values $V(t_1, A_{1,k})$, which follows from the relation,

$$V_{\mathsf{cc}}(t_0, A(t_0)) \approx \mathbb{E}\left[\mathrm{e}^{-\int_{t_0}^{t_1} r(s)\mathrm{d}s} V_{\mathsf{cc}}(t_1, A(t_1)) \;\middle|\; \mathcal{F}_{t_0} \right].$$

It now remains to find an approximation for the continuation values $C_{\mathsf{call}}(t_l, A_{l,k})$ and $C(t_l, A_{l,k})$. The continuation values will be computed using least-squares regression [26]. Both continuation values are obtained in a similar way. Assume that the continuation value $C(t_l, A_{l,k})$ can be approximated by a linear combination of basis functions, i.e.

$$C(t_l, A_{l,k}) \approx \hat{C}_{l,k} = \sum_{\nu=0}^{m-1} a_\nu^{\mathsf{conv}} \xi_\nu(A_{l,k}),$$

for a set $\xi_\nu : \mathbb{R} \to \mathbb{R}$ basis functions and a_ν^{conv} coefficients, where $\nu \in \{0, \dots, m-1\}$, $m \in \mathbb{N}$. Then the parameter vector $a^{\text{conv}} = [a_0^{\text{conv}}, \dots, a_{m-1}^{\text{conv}}]$ can be obtained using least-squares regression, i.e.

$$a^{\text{conv}} = \underset{b \in \mathbb{R}^{m-1}}{\text{argmin}} \| \mathrm{e}^{-r_l \Delta_t} V_{\text{cc}}(t_{l+1}, A_{l+1}) - b\mathrm{X}(A_l) \|^2,$$

where A_l and r_l are the vectors containing all the realizations of $A(t_l)$ and $r(t_l)$ along all $\mathrm{N}-1$ paths respectively, and $V_{\text{cc}}(t_{l+1}, A_{l+1})$ are the approximated bond values at time t_{l+1} along all $\mathrm{N}-1$ paths. Moreover, the matrix $\mathrm{X}(A_l)$ is the explanatory matrix with entries $(\mathrm{X}(A_l))_{l,\nu} = \xi_\nu(A_l)$. Analogously, it is assumed that the continuation value for the management of the company, $C^{\text{call}}(t_l, A_{l,k})$, can also be approximated by the same set of basis functions but different coefficients,

$$C_{\text{call}}(t_l, A_{l,k}) \approx \hat{C}_{\text{call}} = \sum_{\nu=0}^{m-1} a_\nu^{\text{conv}} \xi_\nu(A_{l,k}). \tag{3.6}$$

The coefficients in Equation (3.6) can be obtained by using least-squares regression with respect to the call notice period, i.e.

$$a^{\text{call}} = \underset{b \in \mathbb{R}^{m-1}}{\text{argmin}} \| \mathrm{e}^{-r_l(\Delta_t + t_c)} V_{\text{cc}}(t_{l+c}, A_{l+c}) - b\mathrm{X}(A_l) \|^2,$$

where $V_{\text{cc}}(t_{l+c}, A_{l+c})$ are the approximated callable convertible bond value at time $t_{l+c} = t_l + t_c$ along all $\mathrm{N}-1$ paths for given call notion period t_c and $a^{\text{call}} = [a_0^{\text{call}}, \dots, a_{m-1}^{\text{call}}]$ is the parameter vector.

4

Fourier-based methods for the valuation of callable convertible bonds

4.1. Introduction

The second class of methods that will be discussed in this thesis for valuing callable convertible bonds are Fourier-based techniques. Fourier techniques are based on the approximation of the Fourier transform. In general, there are two types of Fourier transforms; continuous Fourier transforms and discrete Fourier transforms. The different definitions differ up to their conjugates and scaling factors. As in continuous time, the Fourier transform can be used to transfer the pricing problem to Fourier space. Theoretically, the pricing problem can be solved in Fourier space and the result can be recovered using the inverse Fourier transform. However, computing the Fourier transform and its inverse may be difficult. Therefore, the transform and its inverse will be numerically approximated. Using the numerical approximation, a pricing problem can be defined based on the convolution between the option value and the density of the underlying risk factors.

In this chapter both the CONV method used by Ballotta and Kyriakou and the COS method for valuing callable convertible bonds will be discussed [3]. In addition, the valuation methods for the less complex convertible and more complex callable convertible bonds will be discussed. In the next section, a numerical integration scheme based on the fast Fourier transform will be given, from which the continuous Fourier transform and its inverse can be approximated. Using these results, a numerical approximation for a convertible bond value will be derived. This Fourier-based approach will, in the subsequent section, be extended by approximating the value of a callable and convertible bond. Subsequently, a valuation method using the COS method will be derived. First, a general expression for the COS method in one and two dimensions will be given. Subsequently, these expressions will be applied to an algorithm that can be used to obtain convertible bond prices. This algorithm will be extended to the Bermudan case that can be used to value callable convertible issues. Under stochastic interest rates, only a valuation algorithm for non-coupon-paying convertible bonds will be given.

4.2. Approximation of the continuous Fourier transform

The Fourier transform will be defined based on probability density functions. In probability theory, the Fourier transform of the probability density function is also known as the characteristic function. In general, the definition of the Fourier transform is not fixed and may vary up to a scaling constant and conjugate imaginary exponent. Within this thesis, the conventional form used in probability theory will be used to define the continuous Fourier transform.

4.2.1. One-dimensional Fourier transform

Definition 4.1. Consider a probability density function $f_X : \mathbb{R} \to [0, \infty)$ corresponding to a one-dimensional stochastic variable X. The Fourier transform of the density function f_X is called the

characteristic function and is defined as

$$\phi_X(u) = \mathrm{F}[f](u) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}ux} f_X(x) \mathrm{d}x,$$

for $x, u \in \mathbb{R}$. Corresponding to the Fourier transform, the inverse Fourier transform is defined as

$$f_X(x) = \mathrm{F}^{-1}[\phi_X](x) = \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}ux} \phi_X(u) \mathrm{d}u.$$

The notion of the continuous Fourier transform can also be defined in discrete space. When the summation in discrete space is finite, the corresponding sum is called the discrete Fourier transform. A definition is given in Definition 4.2.

Definition 4.2. Let $N \in \mathbb{N}$ and consider sequences $\{x_j\}_{j=0}^{j=N-1}$ and $\{u_n\}_{n=0}^{n=N-1}$. The one-dimensional discrete Fourier transform and its inverse transform are defined as the finite summations,

$$D_n\left(x_j\right) = \sum_{j=0}^{N-1} \mathrm{e}^{\mathrm{i}2\pi \frac{nj}{N}} x_j \text{ and } D_j^{-1}\left(u_n\right) = \sum_{n=0}^{N-1} \mathrm{e}^{-\mathrm{i}2\pi \frac{jn}{N}} u_n,$$

respectively.

A numerical approximation of the Fourier transform in one dimension, can be derived using the trapezoidal integration scheme. First, the integral is truncated to an interval on which the function f_X is almost zero. Next, the truncated integral is approximated using the Newton-Cotes integral rule with trapezoidal weights. At last, by defining a specific grid scheme, the discrete Fourier transform as defined in Definition 4.2 can be used to obtain a fast numerical approximation of the Fourier integral. In particular, recovering values under the transform defined by the discrete Fourier transform can be done very efficiently. The algorithm that enables this fast and efficient computation is often referred to as the Fast Fourier transform (FFT) [27].

Theorem 4.1 (Trapezoidal Fourier approximation)**.** *The continuous Fourier transform defined in Definition 4.1 can be approximated using the trapezoidal integration, i.e.* $\bar{\mathrm{F}}[f](u_n) \approx \mathrm{F}[f](u_n)$. *Assume* $N \in \mathbb{N}$ *and* $j, n \in \{0, \ldots, N-1\}$. *Let* $x_j = \left(j - \frac{N}{2}\right) \Delta_x$ *be an equally spaced grid for a chosen* $\Delta_x \in \mathbb{R}_+$. *Furthermore, let* $u_n = \left(n - \frac{N}{2}\right) \Delta_u$ *and impose the restriction,*

$$\Delta_u \Delta_x = \frac{2\pi}{N}.$$

The resulting numerical integration scheme $\bar{\mathrm{F}}$ *under the trapezoidal weights is then given by*

$$\bar{\mathrm{F}}[f](u_n) = \Delta_x \left[(-1)^{n-\frac{N}{2}} D_n\left(f(x_j)(-1)^j\right) - \frac{1}{2}\left(e^{iu_n x_0} f(x_0) + e^{iu_n x_{N-1}} f(x_{N-1})\right) \right],$$

where the discrete Fourier transform $D_n(\cdot)$ *is defined as in Definition 4.2. Similarly, the inverse continuous Fourier transform can be approximated using trapezoidal integration, i.e.* $\bar{\mathrm{F}}^{-1}[\phi](x_j) \approx \mathrm{F}^{-1}[\phi](x_j)$. *The numerical integration scheme* $\bar{\mathrm{F}}^{-1}$ *for the inverse Fourier transform with trapezoidal weights is given by*

$$2\pi \bar{\mathrm{F}}^{-1}[\phi](x_j) \approx \Delta_u \left[(-1)^{j-\frac{N}{2}} D_j^{-1}\left(\phi(u_n)(-1)^n\right) - \frac{1}{2}\left(e^{iu_0 x_j} \phi(u_0) + e^{iu_{N-1} x_j} \phi(u_{N-1})\right) \right],$$

where the inverse discrete Fourier transform $D_j^{-1}(\cdot)$ *is defined as in Definition 4.2.*

Both the derivation of the Fourier transform as its inverse are given in Appendix D.6. As the approximations can be written in terms of the discrete Fourier transform, the numerical approximations given in Theorem 4.1 enable the use of FFT algorithms to obtain an approximation for the continuous Fourier transform efficiently. For the approximation $\bar{\mathrm{F}}[f]$ (and $\bar{\mathrm{F}}^{-1}[\phi]$) to be effective, the grid x_j should be chosen in such a way that the Fourier integral approaches zero outside the grid.

4.2.2. Two-dimensional Fourier transform

Both the notion of the continuous and discrete Fourier transform can be extended to higher dimensions. Similar to the one-dimensional case, the definition of the Fourier transform will be based on the characteristic function corresponding to a multidimensional random variable.

Definition 4.3. Let $\boldsymbol{X}$ be a multidimensional stochastic variable, where $\boldsymbol{X} = [X_0 \ \dots \ X_{k-1}]^\mathsf{T}$ for stochastic variables $X_0, \dots, X_{k-1}$ and $k \in \mathbb{N},\ k > 0$. Consider a probability density function $f_{\boldsymbol{X}} : \mathbb{R}^k \to [0, \infty)$ corresponding to a stochastic variable $\boldsymbol{X}$. The Fourier transform of the density function $f_{\boldsymbol{X}}$ is called the characteristic function and is defined as

$$\phi_{\boldsymbol{X}}(\boldsymbol{u}) = \mathrm{F}_k[f](\boldsymbol{u}) = \int_{\mathbb{R}^k} \mathrm{e}^{\mathrm{i}\boldsymbol{u}^\mathsf{T}\boldsymbol{x}} f_{\boldsymbol{X}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x},$$

for $\boldsymbol{x}, \boldsymbol{u} \in \mathbb{R}^k$. Corresponding to the Fourier transform, the inverse Fourier transform is defined as

$$f_{\boldsymbol{X}}(x) = \mathrm{F}_k^{-1}[\phi](\boldsymbol{x}) = \frac{1}{(2\pi)^k} \int_{\mathbb{R}^k} \mathrm{e}^{-\mathrm{i}\boldsymbol{u}^\mathsf{T}\boldsymbol{x}} \phi_{\boldsymbol{X}}(\boldsymbol{u}) \mathrm{d}\boldsymbol{u}.$$

Extending Definition 4.2 to higher dimensions will be done by adding new summations over the higher dimensional variables. To illustrate the resulting definitions for the discrete Fourier transform and its inverse transform, the definition for the two-dimensional case will be given in Defintion 4.4.

Definition 4.4. Let $N_0, N_1 \in \mathbb{N}$ and consider four sequences $\{x_j\}_{j=0}^{j=N_0-1}$, $\{y_k\}_{k=0}^{k=N_1-1}$, $\{u_n\}_{n=0}^{n=N_0-1}$ and $\{w_m\}_{m=0}^{m=N_1-1}$. Then the discrete Fourier transform of a function f in two dimensions will be defined as,

$$D_{n,m}\left(f(x_j, y_k)\right) = \sum_{j=0}^{N_0-1} \sum_{k=0}^{N_1-1} \mathrm{e}^{\mathrm{i}2\pi\left(\frac{nj}{N_0} + \frac{mk}{N_1}\right)} f(x_j, y_k).$$

Moreover, the inverse discrete Fourier transform of a function ϕ in two dimensions will be defined as,

$$D_{j,k}^{-1}\left(\phi(u_n, w_m)\right) = \sum_{n=0}^{N_0-1} \sum_{m=0}^{N_1-1} \mathrm{e}^{-\mathrm{i}2\pi\left(\frac{nj}{N_0} + \frac{mk}{N_1}\right)} \phi(u_n, w_m).$$

The discrete Fourier transform in higher dimensions can be retrieved by following the given recursion for each new transformed variable.

The multidimensional Fourier transform extends the one-dimensional case and therefore its approximation is similar to the approximation of the one-dimensional case. Moreover, higher dimensions follow the same pattern of derivation and are therefore also similar to the presented cases. A general formula, however, may become complex and does not fit the purpose of this thesis. Therefore, only an approximation for the two-dimensional case will be given.

Theorem 4.2 (2d trapezoidal Fourier approximation)**.** *The two-dimensional Fourier transform as defined in Definition 4.3 for $k = 2$, can be approximated using the multidimensional quadrature rule. Let the two grids $x_j = \left(j - \frac{N_0}{2}\right)\Delta_x,\ \Delta_x \in \mathbb{R}_+$ and $y_k = \left(k - \frac{N_1}{2}\right)\Delta_y,\ \Delta_y \in \mathbb{R}_+$ be equally spaced grids and define two Fourier grids u_n and w_m such that $u_n = \left(n - \frac{N_0}{2}\right)\Delta_u$ and $w_m = \left(m - \frac{N_1}{2}\right)\Delta_w$ for Δ_u and Δ_w subject to the condition*

$$\Delta_x \Delta_u = \frac{2\pi}{N_0}, \text{ and } \Delta_y \Delta_w = \frac{2\pi}{N_1}.$$

Then the two-dimensional Fourier integral can be approximated by $\bar{\mathrm{F}}_2[f](u_n, w_m) \approx \mathrm{F}_2[f](u_n, w_m)$, for

the integration scheme

$$\begin{aligned}\frac{\bar{\mathrm{F}}_2[f](u_n,w_m)}{\Delta_x\Delta_y} &= (-1)^{n+m+\frac{N_1+N_0}{2}}\bar{D}_{n,m}\left(f(x_j,y_k)\right)\\ &-\frac{1}{2}\Big((-1)^{m+\frac{N_1}{2}}\mathrm{e}^{\mathrm{i}ux_0}\bar{D}_m\left(f(x_0,y_k)\right)+(-1)^{m+\frac{N_1}{2}}\mathrm{e}^{\mathrm{i}ux_{N_0-1}}\bar{D}_m\left(f(x_{N_0-1},y_k)\right)\\ &+(-1)^{n+\frac{N_0}{2}}\mathrm{e}^{\mathrm{i}w_my_0}\bar{D}_n\left(f(x_j,y_0)\right)+(-1)^{n+\frac{N_0}{2}}\mathrm{e}^{\mathrm{i}w_my_{N_1-1}}\bar{D}_n\left(f(x_j,y_{N_1-1})\right)\Big)\\ &+\frac{1}{4}\left(\mathrm{e}^{\mathrm{i}ux_0}\mathrm{e}^{\mathrm{i}wy_0}f(x_0,y_0)+\mathrm{e}^{\mathrm{i}ux_0}\mathrm{e}^{\mathrm{i}wy_{N_1-1}}f(x_0,y_{N_1-1})\right.\\ &\left.+\mathrm{e}^{\mathrm{i}ux_{N_0-1}}\mathrm{e}^{\mathrm{i}wy_0}f(x_{N_0-1},y_0)+\mathrm{e}^{\mathrm{i}ux_{N_0-1}}\mathrm{e}^{\mathrm{i}wy_{N_1-1}}f(x_{N_0-1},y_{N_1-1})\right),\end{aligned}$$

where $\bar{D}_{n,m}\left(f(x_j,y_k)\right)=D_{n,m}\left(f(x_j,y_k)(-1)^{k+j}\right)$ *for the two-dimensional DFT* $D_{n,m}\left(\cdot\right)$ *given in Definition 4.4 and* $\bar{D}_n\left(f(x_j)\right)=D_n\left(f(x_j)(-1)^j\right)$ *for the one-dimensional DFT* $D_n\left(\cdot\right)$ *given in Definition 4.2. In addition, the two-dimensional inverse Fourier transform can be approximated using the trapezoidal integration rule. Moreover, it holds that* $\bar{\mathrm{F}}_2^{-1}[\phi](x_j,y_k)\approx\mathrm{F}_2^{-1}[\phi](x_j,y_k)$*, where it holds that*

$$\begin{aligned}\frac{4\pi^2}{\Delta_u\Delta_v}\bar{\mathrm{F}}_2^{-1}[\phi](x_j,y_k) &= (-1)^{k+j+\frac{N_0+N_1}{2}}\bar{D}_{j,k}^{-1}\left(\phi(u_n,w_m)\right)\\ &-\frac{1}{2}\Big((-1)^{j+\frac{N_0}{2}}\bar{D}_j^{-1}\left(\phi(u_n,w_0)\right)+(-1)^{j+\frac{N_0}{2}}\bar{D}_j^{-1}\left(\phi(u_n,w_{N_1-1})\right)\\ &+(-1)^{k+\frac{N_1}{2}}\bar{D}_k^{-1}\left(\phi(u_0,w_m)\right)+(-1)^{k+\frac{N_1}{2}}\bar{D}_k^{-1}\left(\phi(u_{N_0-1},w_m)\right)\Big)\\ &+\frac{1}{4}\Big(\mathrm{e}^{-\mathrm{i}(u_0x+w_0y)}\phi(u_0,w_0)+\mathrm{e}^{-\mathrm{i}(u_0x+w_{N_1-1}y)}\phi(u_0,w_{N_1-1})\\ &+\mathrm{e}^{-\mathrm{i}(u_{N_0-1}x+w_0y)}\phi(u_{N_0-1},w_0)+\mathrm{e}^{-\mathrm{i}(u_{N_0-1}x+w_{N_1-1}y)}\phi(u_{N_0-1},w_{N_1-1})\Big),\end{aligned}$$

where $\bar{D}_{j,k}^{-1}\left(\phi(u_n,w_m)\right)=D_{j,k}^{-1}\left(\phi(u_n,w_m)(-1)^{n+m}\right)$ *for the two-dimensional inverse DFT* $D_{n,m}^{-1}\left(\cdot\right)$ *given in Definition 4.4 and* $\bar{D}_j^{-1}\left(\phi(u_n)\right)=D_j^{-1}\left(\phi(u_n)(-1)^n\right)$ *for the one-dimensional inverse DFT* $D_n^{-1}\left(\cdot\right)$ *given in Definition 4.2.*

The derivations of the integration schemes are given in Appendix D.7. Similar to the one-dimensional case, also higher dimensional DFT's can be computed efficiently by use of a higher dimensional Fast Fourier algorithm. As the integration schemes are written in terms of the DFT, the two-dimensional Fourier transform and its inverse can be efficiently computed by use of the two-dimensional FFT algorithm. Similarly to the one-dimensional case, for Theorem 4.2 to be effective, the grids x_j and y_k cannot be chosen as too small or too big. A bad choice of Δ_x and Δ_y will result in a bad accuracy when performing approximating the Fourier transform.

4.3. Convolution-based approximation methods for valuing bonds

Assume that the time horizon is given by $[t_0,T]$ and define the time grid $t_l=l\Delta_l$ for $\Delta_t=\frac{T}{\mathrm{M}-1}$ such that $l\in\{0,\ldots,\mathrm{M}-1\}$. Let the coupon dates $\mathcal{T}^{\mathsf{coup}}$ coincide with the defined time grid t_l. Let $\gamma\in(0,1]$ be the conversion factor. Furthermore, recall the coupon rates α, such that $c_j=\alpha F,\ \forall j$, for the face value of the bond F as defined in Section 1.2. The corporate bonds will be valued using a Fourier technique which is often referred to as the convolution method [27, 3]. To illustrate the algorithm, first a constant interest rate $r\equiv r(t)$ will be assumed under which the proposed algorithm will be used to approximate the value of a convertible bond under the structural default model given in Section 2.2. In the next subsection, the algorithm will be extended to incorporate stochastic interest rates.

The valuation method is based on the continuation value of the corporate bond. At maturity T, the bond value is known and is equal to its payoff value,

$$V_{\mathsf{conv}}(T,A(T))=H_{\mathsf{conv}}(T,A(T))=\min(A(T),\max((1+\alpha)F,\gamma A(T))),$$

For each intermediate time $t_0<t_l<T$ before maturity, the bond value is equal to the liquidation of the company $A(t_l)$ in case of a default or to the future discounted payoff increased with the coupon

payment at time $t_l \in \mathcal{T}^{\text{coup}}$. The value of a convertible bond $V_{\text{conv}}(t_l, A(t_l))$ at time $t_l \in \mathcal{T}^{\text{coup}}$ can be described as

$$V_{\text{conv}}(t_l, A(t_l)) = \begin{cases} A(t_l) & \text{for } A(t_l) < \alpha F, \\ \alpha F + \mathrm{e}^{-r\Delta_t} V_{\text{conv}}(t_{l+1}, A(t_{l+1})) & \text{for } A(t_l) \geq \alpha F. \end{cases}$$

Note that at time $t_l \notin \mathcal{T}^{\text{coup}}$, the value of the bond will be equal to the continuation value of the bond, i.e. $V_{\text{conv}}(t_l, A(t_l)) = \mathrm{e}^{-r\Delta_t} V_{\text{conv}}(t_{l+1}, A(t_{l+1}))$.

To optimally use the efficient characteristics of the convolution method, the corporate bond value will be computed under a transformation. Define the log-face transform under the final payment to be

$$X_F(t) = \log\left(\frac{A(t)}{F}\right) \iff A(t) = F\mathrm{e}^{X_F(t)}, \tag{4.1}$$

where $X_F(t_0) = \log(\frac{A_0}{F})$ is the initial value, for initial value $A(t_0) = A_0$ of the process $\{A(t)\}_{t\geq 0}$. An important observation is that the dynamics of the process $\{X_F(t)\}_{t\geq 0}$ are equal to the dynamics of $\{X(t)\}_{t\geq 0}$ defined in Chapter 2, as the transformation is done for a constant value F. The payoff under the log-face transformation defined in Equation (4.1) is then given by

$$V_{\text{conv}}(T, X_F(T)) = H_{\text{conv}}(T, X_F(T)) = F\min(\mathrm{e}^{X_F(T)}, \max((1+\alpha), \gamma\mathrm{e}^{X_F(T)})).$$

Let the scaled payoff function at maturity, therefore, be given by

$$\bar{H}_{\text{conv}}(T, X_F(T)) = \min(\mathrm{e}^{X_F(T)}, \max((1+\alpha), \gamma\mathrm{e}^{X_F(T)})),$$

such that $H_{\text{conv}}(T, X_F(T)) = F\bar{H}_{\text{conv}}(T, X_F(T))$. As a consequence, the scaled value of the convertible bond will therefore be given as,

$$\bar{V}_{\text{conv}}(t_l, X_F(t_l)) = \begin{cases} \mathrm{e}^{X_F(t_l)} & \text{for } \mathrm{e}^{X_F(t_l)} < \alpha, \\ \alpha + \mathrm{e}^{-r\Delta_t} \bar{V}_{\text{conv}}(t_{l+1}, X_F(t_{l+1})) & \text{for } \mathrm{e}^{X_F(t_l)} \geq \alpha. \end{cases}$$

The approach using scaled functions will enable the method to compute bond prices for vectors of face values.

4.3.1. Convertible bonds under constant interest rates

Let $\bar{C}_l$ be the scaled continuation value of the convertible bond, $\bar{C}_l \equiv FC(t_l, X_F(t_l))$ with $C(t_l, X_F(t_l))$ as defined in Definition 1.3, that under the assumption of constant interest rates is given as,

$$\bar{C}_l = \mathrm{e}^{-r\Delta_t}\mathbb{E}\left[\bar{V}_{\text{conv}}(t_{l+1}, X_F(t_{l+1})) \;\middle|\; \mathcal{F}_{t_l}\right],$$

and can be computed using Fourier techniques. Indeed, the general form of the scaled continuation value under the Fourier transform is given as,

$$\mathrm{e}^{r\Delta_t}\mathrm{F}[\bar{C}_l](u) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}x_l u}\bar{C}_l(t_l, x_l)\mathrm{d}x_l.$$

Assume that the condition [27]

$$f_{X_F}(x|x_l) = f_{X_F}(x - x_l), \tag{4.2}$$

holds for the discounted transition density $f_{X_F}(x|y)$ corresponding to the process $\{X_F(t)\}_{t\geq 0}$ going from state y to state x. In particular, the condition in Equation (4.10) holds for processes with independent increments [27]. Under the assumption stated in Equation (4.10), the continuation value $\bar{C}$ can be viewed as a convolution between the scaled option price $\bar{V}_{\text{conv}}(t_{l+1}, x_{l+1}) \equiv \bar{V}_{l+1}$ at time t_{l+1} and the transition density function f_{X_F}. The resulting convolution can be computed as a multiplication between the option value $\bar{V}_{l+1}$ and the density f_{X_F} in Fourier space. The following theorem summarizes this observation.

Theorem 4.3. *Assume that Equation* (4.2) *holds for the transition density f_{X_F} corresponding to the process $\{X_F\}_{t\geq 0}$. The Fourier transform of $\bar{C}_l$ is can be computed as*

$$\mathrm{F}[\bar{C}_l](u) = \mathrm{e}^{r\Delta_t}\mathrm{F}[\bar{V}_{l+1}](u)\phi_{X_F}(-u),$$

where $F[\bar{V}_{l+1}]$ is the Fourier transform of the scaled value function $\bar{V}_{\text{conv}}(t_{l+1}, X_F(t_l))$ of the bond at time t_{l+1} and ϕ_{X_F} the characteristic function corresponding to the discounted density function f_{X_F} over the time period Δ_t.

The proof of the theorem is given in Appendix D.9. Fourier techniques take place in two spaces, the asset space, and the Fourier space. Therefore, define a grid $x_j = \left(j - \frac{N}{2}\right)\Delta_x$ in asset-face space, for $N \in \mathbb{N}$, $j \in \{0, \ldots, N-1\}$ and $\Delta_x \in \mathbb{R}$. Moreover, define a grid in Fourier space given by $u_n = \left(n - \frac{N}{2}\right)\Delta_u$, for $n \in \{0, \ldots, N-1\}$ and impose the restriction,

$$\Delta_u \Delta_x = \frac{2\pi}{N}.$$

The convertible bond value will be obtained by computing the intermediate bond values backward in time. At time step T, the bond value is known and given by

$$\bar{V}_{\text{conv}}(T, x_j) = \bar{H}_{\text{conv}}(T, x_j).$$

For each time step $t_l < T$, the continuation value of the corporate bond can be computed using the approximation of the Fourier transform. Assume that the scaled value of the bond at time t_{l+1} is known and given by $\bar{V}_{t_{l+1}} \equiv \bar{V}_{\text{conv}}(t_{l+1}, x)$. For time steps t_l, $l \in \{1, \ldots, \mathrm{M}-2\}$ it holds that

$$\bar{C}(t_l, x_j) \approx \mathrm{e}^{-r\Delta_t} \bar{\mathrm{F}}^{-1}\left[\bar{\mathrm{F}}[\bar{V}_{l+1}](u)\phi_{X_F}(-u)\right](x_j),$$

where the numerical approximation of the Fourier transform is given in Theorem 4.1. To obtain the value of the option at time t_l the relation,

$$\bar{V}_{\text{conv}}(t_l, x_j) = \begin{cases} \mathrm{e}^{x_j} & \text{for } \mathrm{e}^{x_j} < \alpha, \\ \alpha + \bar{C}(t_l, x_j) & \text{for } \mathrm{e}^{x_j} \geq \alpha, \end{cases}$$

can be used if $t_l \in \mathcal{T}^{\text{coup}}$ and $\bar{V}_{\text{conv}}(t_l, x_j) = \bar{C}(t_l, x_j)$ otherwise. At time step t_0, there will be no coupon payments. Hence, it holds that the value of the option will be equal to its continuation value, i.e.

$$\bar{V}_{\text{conv}}(t_0, X_F(t_0)) = \bar{C}(t_0, X_F(t_0)).$$

As a consequence, the values $\bar{V}_{\text{conv}}(t_0, X_F(t_0))$ for different values x_j on the grid can be obtained by performing the Fourier scheme,

$$\bar{V}_{\text{conv}}(t_0, x_j) \approx \mathrm{e}^{-r\Delta_t} \bar{\mathrm{F}}^{-1}\left[\bar{\mathrm{F}}[\bar{V}_1](u)\phi_{X_F}(-u)\right](x_j).$$

An approximation of the corporate bond value $V_{\text{conv}}(t_0, X_F(t_0))$ at the initial time t_0 can be computed using the relation;

$$V_{\text{conv}}(t_0, X_F(t_0)) = F\bar{V}_{\text{conv}}(t_0, X_F(t_0)),$$

where the values for $\bar{V}_{\text{conv}}(t_0, X_F(t_0))$ corresponding to different values of F can be obtained by using interpolation with respect to the grid x_j.

4.3.2. Convertible bonds under stochastic interest rates

This section will continue under the assumption of stochastic interest rates, where the dynamics of the short rate process $\{r(t)\}_{t\geq 0}$ follow the Hull-White dynamics, as described in Chapter 2. Similar to the approach in one dimension, the continuation value of the bond will be used to obtain the bond value at time t_0. However, to ease computations, the continuation value of the bond under the $t_{l+1}-$forward measure will be used.

Lemma 4.1. *Let*

$$C(t_l, X_F(t_l)) = \mathbb{E}\left[\mathrm{e}^{-\int_{t_l}^{t_{l+1}} r(s)ds} V_{\text{conv}}(t_{l+1}, X_F(t_{l+1})) \;\middle|\; \mathcal{F}_{t_l}\right],$$

be the continuation value of the convertible bond under the equivalent martingale measure $\mathbb{Q}$, where $\mathcal{F}_{t_l}$ is the filtration at time t_l with respect to the processes $\{X_F(t)\}_{t\geq 0}$ and $\{r(t)\}_{t\geq 0}$ and where $V_{\text{conv}}(t_l, X_F(t_l))$ is the value of the convertible bond at time t_l for $l \in \{0, \ldots, \mathrm{M}-2\}$. The continuation value $C(t_l, X_F(t_l))$ of the bond under the $t_{l+1}-$forward measure is given by

$$C(t_l, X_F(t_l)) = P(t_l, t_{l+1})\mathbb{E}^{t_{l+1}}\left[V_{\text{conv}}(t_{l+1}, X_F(t_{l+1})) \mid \mathcal{F}_{t_l}\right].$$

The proof of Lemma 4.1 is given in Appensdix D.8.

To be able to compute the continuation value of the bond under the CONV method, the approach of Ballotta and Kyriakou is considered [3]. As in Section 4.3.1, the Fourier approximations will be done based on the scaled payoff and continuation value. As the decomposed interest rate process $\{\tilde{r}(t)\}_{t\geq 0}$ is stochastic, the value of the zero-coupon bond $P(t_l, t_{l+1})$ for $l > 0$ will also be stochastic. To remove the dependence on the stochastic state variable $\tilde{r}(t_l)$ for $l > 0$, the transformation proposed in Section 2.3 will be applied. Recall the transformation under which the characteristic function will be free of state variables as

$$\begin{cases} Z_r(t_{l+1}) &= \tilde{r}(t_{l+1}) - g_r(\tilde{r}(t_l)), \\ Z(t_{l+1}) &= X_F(t_{l+1}) - g(\tilde{r}(t_l), X_F(t_l)), \end{cases} \tag{4.3}$$

where $B(\tau) = \frac{1}{\kappa}(1-\mathrm{e}^{-\kappa\tau})$, $g_r(x) = x\mathrm{e}^{-\kappa\Delta_t}$ and $g(x,y) = y + B(\Delta_t)x$. Under the variable transformation given in Equation (4.3), the continuation value can be written as a convolution with respect to the transformed variables $g_r(\tilde{r}(t_l))$ and $g(\tilde{r}(t_l), X_F(t_l))$.

Theorem 4.4. *Let Z_r and Z be two random variables defined as in Equation* (4.3)*. The continuation value $C(t_l, \tilde{r}(t_l), X_F(t_l)) \equiv C(t_l, X_F(t_l))$ as defined in Lemma 4.1 is related to the continuation value $C(t_l, g_r(\tilde{r}(t_l)), g(\tilde{r}(t_l), X_F(t_l)))$ under the transformation $g_r(x) = x\mathrm{e}^{-\kappa\Delta_t}$ and $g(x,y) = y + B(\Delta_t)x$ via,*

$$C(t_l, \tilde{r}(t_l), X_F(t_l)) = C(t_l, g_r(\tilde{r}(t_l)), g(\tilde{r}(t_l), X_F(t_l))),$$

where the convolution under the transformation is given by

$$\frac{C(t_l, g_r(\tilde{r}(t_l)), g(\tilde{r}(t_l), X_F(t_l)))}{P(t_l, t_{l+1})} = \int_{\mathbb{R}^2} V_{\mathsf{conv}}(t_{l+1}, z_r + g_r(\tilde{r}(t_l)), z + g(\tilde{r}(t_l), X_F(t_l))) f_{Z_r,Z}(z, z_r) \mathsf{d}(z_r, z).$$

A proof of Theorem 4.4 is given in Appendix D.10. To apply the convolution method, define two grids in asset space, $x_j = \left(j - \frac{N}{2}\right)\Delta_x$ and $y_k = \left(k - \frac{N}{2}\right)\Delta_y$, for $N = N_0 = N_1 \in \mathbb{N}$, $j, k \in \{0, \ldots, N-1\}$ and $\Delta_x, \Delta_y \in \mathbb{R}_+$. For ease of computation, it is assumed that both grids have the same number of steps. Within the different steps of the algorithm, the interpretation of the grids might be different. Both the x_j and y_k grid will serve two interpretations. The grid x_j will correspond to both the values of the interest rate process $\tilde{r}(t)$ and the transformed interest rate process $g_r(\tilde{r}(t))$. Similarly, the grid y_k will correspond to the transformed asset $X_F(t)$ and its transform $g(\tilde{r}(t), X_F(t))$. Moreover, define grids in Fourier space given by $u_n = \left(n - \frac{N}{2}\right)\Delta_u$ and $w_m = \left(m - \frac{N}{2}\right)\Delta_w$, for $n, m \in \{0, \ldots, N-1\}$ and impose the restrictions,

$$\Delta_u\Delta_x = \frac{2\pi}{N},\ \Delta_w\Delta_y = \frac{2\pi}{N}.$$

The corporate bond value will be obtained by computing the intermediate bond values backward in time. At time step T, the bond value is known and given by

$$\bar{V}_{\mathsf{conv}}(T, y_k) = \bar{H}_{\mathsf{conv}}(T, y_k).$$

Given the scaled corporate bond value at time t_{l+1}, $\bar{V}_{l+1} \equiv \bar{V}_{\mathsf{conv}}(t_{l+1}, y_k)$, the scaled continuation value under the transformation $(g_r(x), g(x,y))$, $\hat{C}(t_l, x, y) \equiv \frac{FC(t_l, g_r(x), g(x,y))}{P(t_l, t_{l+1})}$, can be obtained using the approximation of the Fourier transform and its inverse.

Theorem 4.5. *The continuation value under the transformation given in Theorem 2.5 can be computed using the two-dimensional Fourier transform,*

$$\hat{C}(t_l, x_j, y_k) = \mathrm{F}^{-1}[\mathrm{F}[\bar{V}_{l+1}](u_n, w_m)\phi_{Z,Z_r}(-u_n, -w_m)](x_j, y_k), \tag{4.4}$$

where $\phi_{Z_r,Z}$ is the characteristic function corresponding to the transformed variables Z_r and Z over the time interval $\tau = \Delta_t$.

The proof of Theorem 4.5 is given in Appendix D.11. The continuation value $\hat{C}(t_l, x, y)$ is computed with respect to the transformation $g(x_j)$ and $g(x_j, y_k)$. To obtain the continuation value $\bar{C}(t_l, x_j, y_k) \equiv FC(t_l, x_j, y_k)$, interpolation and the relation for the zero-coupon bond is used,

$$\bar{C}(t_l, x_j, y_k) = \mathrm{e}^{A(\tau)-B(\tau)x_j}\hat{C}(t_l, g_r(x_j), g(x_j, y_k)). \tag{4.5}$$

As values of $g(x_j, y_k)$ may lie outside the grid defined by y_k, linear extrapolation is used to obtain the values for $g(x_j, y_k)$ outside the grid y_k. Using the obtained values for $\bar{C}(t_l, y_k) \equiv \bar{C}(t_l, x_j, y_k)$, the approximated continuation value under the decomposed interest $\tilde{r}(t_l)$ and transformed log-face value $X_F(t_l)$, the value of the corporate bond can be computed using,

$$\bar{V}_{\mathsf{conv}}(t_l, y_k) = \begin{cases} \mathrm{e}^{y_k} & \text{for } \mathrm{e}^{y_k} < \alpha, \\ \alpha + \bar{C}(t_l, y_k) & \text{for } \mathrm{e}^{y_k} \geq \alpha, \end{cases}$$

when $t_l \in \mathcal{T}^{\mathsf{coup}}$ and $\bar{V}_{\mathsf{conv}}(t_l, y_k) = \bar{C}(t_l, y_k)$ otherwise. At time step t_0, there will be no coupon payments. As in the one-dimensional case, it holds that the value of the option will be equal to its continuation value, i.e.

$$\bar{V}_{\mathsf{conv}}(t_0, X_F(t_0)) = \bar{C}(t_0, X_F(t_0)).$$

As a consequence, the values $\bar{V}_{\mathsf{conv}}(t_0, X_F(t_0))$ for different values x_j and y_k on the grids, the continuation value $\hat{C}(x_j, y_k)$ under the transform $g_r(x_j)$ and $g(x_j, y_k)$ can be obtained by performing the Fourier scheme as given in Equation (4.4), and using interpolation to obtain the continuation value $\bar{C}(x_j, y_k)$ at time t_0. An approximation of the corporate bond value $V_{\mathsf{conv}}(t_0, X_F(t_0))$ at the initial time t_0 can be computed using the relation;

$$V_{\mathsf{conv}}(t_0, X_F(t_0)) = F\bar{C}(t_0, X_F(t_0)).$$

4.3.3. Callable convertible bonds under stochastic interest rates

Given the Fourier approach discussed in Section 4.3.1 and Section 4.3.2, the valuation algorithm can easily be extended to obtain values for the callable convertible bond. This section will continue under the assumption of stochastic interest rates. The derivations of the callable convertible bond under constant interest rates are analogously derived by combining the results of this section and Section 4.3.1. Consider a callable convertible bond with face value F and conversion factor γ. Note that the payoff at maturity T of the callable convertible bond is only dependent on the convertible feature, as the bond can not be called at maturity T. Hence, the payoff of the callable convertible bond at maturity T is equal to the payoff of a convertible bond and defined by,

$$H_{\mathsf{conv}}(T, A(T)) = \min(A(T), \max((1+\alpha)F, \gamma A(T))). \tag{4.6}$$

The corresponding payoff scaled to the face value F is given by,

$$\bar{H}_{\mathsf{conv}}(T, X_F(T)) = \min(e^{X_F(T)}, \max(1+\alpha, \gamma \mathrm{e}^{X_F(T)})).$$

For every pre-mature date $t_0 < t_l < T$, the bond may be called or be converted. Assume that the call price $K(t)$ corresponding to the callable feature of the bond can be written as a time-dependent penalty function β with respect to the face value F of the bond, i.e.

$$K(t) = \beta(t)F.$$

Furthermore, assume that the coupon, callable and convertible dates coincide, $\mathcal{T}^{\mathsf{coup}} = \mathcal{T}^{\mathsf{call}} = \mathcal{T}^{\mathsf{conv}}$. The call notion period, corresponding to the bond is given in the indenture of the bond and is denoted as $t_c \in \mathbb{R}_+$.

Both the bondholder and the management of the company will try to find their optimal exercise moment to exercise their option. For the management of the company, the optimal exercise moment is chosen in such a way that the value of the bond is minimized. The strategy of minimizing the bond value will optimize the wealth of the shareholders, which is a direct result of the capital structure of the company. To apply the CONV method, the continuation value under the t_{l+1} measure need to be computed.

Lemma 4.2. *Let $V_{\mathsf{cc}}(t, X_F(t))$ be the callable convertible bond value at time $t \in [t_0, T]$ for a given value of the transformed process $X_F(t)$. Let the continuation value of the callable convertible bond under the call notice period t_c be defined as in Definition 1.5. Then the continuation value $C^{\mathsf{call}}(t_l, X_F(t_l))$ under the t_{l+1}-forward measure is given by,*

$$\begin{aligned} C^{\mathsf{call}}(t_l, X_F(t_l)) &= \mathbb{E}\left[\mathrm{e}^{-\int_{t_l}^{t_{l+1}+t_c} r(s)\mathsf{d}s} V_{\mathsf{cc}}(t_{l+1}+t_c, X_F(t_{l+1}+t_c)) \;\middle|\; \mathcal{F}_{t_l} \right] \\ &= P(t_l, t_{l+1}+t_c)\mathbb{E}^{t_{l+1}+t_c}\left[V_{\mathsf{cc}}(t_{l+1}+t_c, X_F(t_{l+1}+t_c)) \mid \mathcal{F}_{t_l} \right]. \end{aligned} \tag{4.7}$$

The proof of Lemma 4.2 is analogous to the proof of Lemma 4.1 given in Appensdix D.8 and will therefore not be restated. The bond will be called as soon as the discounted future bond value is expected to exceed the call price of the bond increased with the accrued coupon payments. As the call dates and the coupon dates coincide, the accrued interest is equal to the coupon payment. As a consequence, the call option will be exercised if

$$C^{\mathsf{call}}(t_l, X_F(t_l)) > (\beta(t_l) + \alpha)F. \tag{4.8}$$

When the bond is called, the bondholder has to choose between the call prices increased with the coupon payment and forced conversion. As the bondholder will try to maximize her wealth, she will choose the maximum of the call prices increased with the coupon payment and the value upon conversion.

Contrary to the management of the company, the bondholder will exercise her option to maximize her wealth. The continuation value of the bond with respect to the bondholder does not take into account the call notice period. Hence, the convertible continuation value is given by $C(t_l, A(t_l))$,

$$\begin{aligned} C(t_l, X_F(t_l)) &= \mathbb{E}\left[\mathrm{e}^{-\int_{t_l}^{t_{l+1}} r(s)\mathsf{d}s} V_{\mathsf{cc}}(t_{l+1}, X_F(t_{l+1})) \;\middle|\; \mathcal{F}_{t_l} \right] \\ &= P(t_l, t_{l+1})\mathbb{E}^{t_{l+1}}\left[V_{\mathsf{cc}}(t_{l+1}, X_F(t_{l+1})) \mid \mathcal{F}_{t_l} \right], \end{aligned} \tag{4.9}$$

under the $t_{l+1}-$forward measure and similar to the continuation value of the convertible bond as discussed in Lemma 4.1. To maximize her wealth, the bondholder will thus exercise if the immediate payoff of the convertible option exceeds the continuation value, i.e.

$$C(t_l, A(t_l)) < \gamma F \mathrm{e}^{X_F(t_l)}. \tag{4.10}$$

As discussed in Chapter 3, at the moment of conversion the company value and the market value of the equity of the company will coincide due to the capital structure of the company. As a consequence, if the bond is converted at time t_l, its corresponding payoff at conversion will be $\gamma F \mathrm{e}^{X_F(t_l)}$.

The callable convertible bond value can be approximated following the pricing algorithm discussed in Section 4.3.2. Note that, even when the call notice period is considered, the convolution argument under the transformed variables Z_r and Z still holds. Recall the defined grids in asset space, $x_j = \left(j - \frac{N}{2}\right)\Delta_x$ and $y_k = \left(k - \frac{N}{2}\right)\Delta_y$, for $N = N_0 = N_1 \in \mathbb{N}$, $j, k \in \{0, \ldots, N-1\}$ and $\Delta_x, \Delta_y \in \mathbb{R}_+$. Moreover, the grids in Fourier space are given by $u_n = \left(n - \frac{N}{2}\right)\Delta_u$ and $w_m = \left(m - \frac{N}{2}\right)\Delta_w$, for $n, m \in \{0, \ldots, N-1\}$. The restrictions,

$$\Delta_u \Delta_x = \frac{2\pi}{N},\ \Delta_w \Delta_y = \frac{2\pi}{N},$$

will be imposed to enable the fast Fourier algorithm. As discussed in the previously described approaches (Section 4.3.1 and Section 4.3.2), a scaled problem will be solved to be able to efficiently calculate the callable convertible bond values for a vector of face values. At maturity T, the value of the callable convertible bond is known and given by $\bar{H}_{\mathsf{conv}}(T, y_k)$.

The algorithm will continue backward in time. Assume that the values $V_{l+1} \equiv V_{\mathsf{cc}}(t_{l+1}, y_k)$ and $V_{l+c} \equiv V_{\mathsf{cc}}(t_l + t_c, y_k)$ are known for all $k \in \{0, \ldots, N-1\}$ and recall their scaled variants are given by $\bar{V}_{l+1} = FV_{l+1}$ and $\bar{V}_{l+c} = FV_{l+c}$ respectively. To emphasize the dependence of the algorithm on the values of the decomposed interest rate processes $\{\tilde{r}(t)\}_{t \geq 0}$, denote the continuation value of the callable

convertible bond $C(t_l, \tilde{r}(t_l), X_F(t_l)) \equiv C(t_l, X_F(t_l))$ and $C_{\text{call}}(t_l, \tilde{r}(t_l), X_F(t_l)) \equiv C_{\text{call}}(t_l, X_F(t_l))$ as given in Equation (4.9) and Equation (4.7) respectively. At each time step $t_0 < t_l < T$, two continuation values need to be estimated, $\bar{C}_{\text{call}}(t_l, x_j, y_k) = FC_{\text{call}}(t_l, x_j, y_k)$ and $\bar{C}(t_l, x_j, y_k) = FC(t_l, x_j, y_k)$. Both continuation values can be approximated following the convolution method. Using the results of Theorem 4.5, the continuation value $\hat{C}(t_l, x_j, y_k) \equiv FC(t_l, g_r(x_j), g(x_j, y_k))$ under the transformation $(Z_r, Z) \equiv (Z_r(t_{l+1}), Z(t_{l+1}))$ can be obtained by computing

$$\hat{C}(t_l, x_j, y_k) \approx \bar{\mathrm{F}}[\bar{\mathrm{F}}^{-1}[\bar{V}_{l+1}](u_n, w_m)\phi_{Z,Z_r}(-u_n, -w_m; \Delta_t)](x_j, y_k), \tag{4.11}$$

where $\phi_{Z,Z_r}(u_n, w_m; \Delta_t)$ is the characteristic function with respect to (Z_r, Z) for $\tau = \Delta_t$. Using the same approach as described in Section 4.3.2, the value of $\bar{C}(t_l, x_j, y_k)$ can be obtained using interpolation, with linear extrapolation for the values that lie outside the grid. The bond value $P(t_l, t_{l+1})$ corresponding to the value x_j can then be used to address for the discounting. Similarly, the continuation value $\hat{C}_{\text{call}}(t_l, x, y) \equiv FC_{\text{call}}(t_l, g_r(x), g(x, y))$ corresponding to the call feature, can be derived using a similar approximation as discussed in Theorem 4.5,

$$\hat{C}_{\text{call}}(t_l, x_j, y_k) \approx \bar{\mathrm{F}}[\bar{\mathrm{F}}^{-1}[\bar{V}_{l+c}](u_n, w_m)\phi_{Z,Z_r}(-u_n, -w_m; \Delta_t + t_c)](x_j, y_k),$$

however, using $\phi_{Z,Z_r}(u_n, w_m; \Delta_t + t_c)$ to correct for the call notice period. The continuation values $\bar{C}_{\text{call}}(t_l, x_j, y_k)$ can be obtained using interpolation with respect to the transformations $g_r(x_j)$ and $g(x_j, y_k)$,

$$\bar{C}_{\text{call}}(t_l, x_j, y_k) = \mathrm{e}^{A(\tau) - B(\tau)x_j}\hat{C}_{\text{call}}(t_l, g_r(x_j), g(x_j, y_k)).$$

Based on the computed continuation values, the bond value at time $t_0 < t_l < T$ can be approximated. If the company value does not provide sufficient resources to fulfill the obligations, the company will default. Furthermore, if one of the conditions in Equation (4.8) or Equation (4.10) apply, the corresponding option will be exercised respectively and the payoff of the option is known. Otherwise, the option will not be exercised and the decision process will continue. Combining all possibilities, yields that the scaled option value $\bar{V}_{\text{cc}}(t_l, x_j, y_k)$ at time $t_l \in \mathcal{T}^{\text{coup}}$ is given by

$$\bar{V}_{\text{cc}}(t_l, x_j, y_k) = \begin{cases} \mathrm{e}^{y_k} & \text{for } \mathrm{e}^{y_k} < \alpha, \\ \max(\beta(t_l) + \alpha, \gamma\mathrm{e}^{y_k}) & \text{for } \bar{C}_{\text{call}}(x_j, y_k) \geq \beta(t_l) + \alpha \text{ and } \mathrm{e}^{y_k} \geq \alpha, \\ \gamma\mathrm{e}^{y_k} & \text{for } \bar{C}(x_j, y_k) \geq \gamma\mathrm{e}^{y_k},\ \mathrm{e}^{y_k} \geq \alpha \\ & \text{ and } \bar{C}_{\text{call}}(x_j, y_k) < \beta(t_l) + \alpha, \\ \alpha + \bar{C}(x_j, y_k) & \text{otherwise.} \end{cases}$$

When $t_l \notin \mathcal{T}^{\text{coup}}$, the scaled value of the bond $\bar{V}_{\text{cc}}(t_l, x_j, y_k)$ at time t_l reduces to $\bar{C}(x_j, y_k)$.

At the initial time t_0, the bond will neither be exercisable nor payout any coupon payments. Hence, the value of the bond will be equal to its continuation value at time t_0. The scaled continuation value $\bar{C}(x_j, y_k)$ can be obtained using Equation (4.11) and performing interpolation and linear extrapolation on the plain $(g_r(x_j), g(x_j, y_k), \hat{C}(x_j, y_k))$. Using the value for $P(t_0, t_1)$, the bond value can therefore be approximated by,

$$V_{\text{cc}}(t_0, X_F(t_0)) \approx FP(t_0, t_1)\bar{C}(r_0, X_F(t_0)).$$

4.4. Cosine based approximation methods for valuing bonds

Besides the convolution approximation method, also a cosine-based method exists which can be used to approximate the value of a financial derivative. This section will discuss the cosine method, which is often referred to as the COS method. The one-dimensional COS method is already discussed in detail by for example Oosterlee and Grzelak [32]. Hence for completeness, only the results of the one-dimensional COS method will be stated.

Theorem 4.6. *Let $\phi_X : \mathbb{R} \to \mathbb{C}$ be the characteristic function corresponding to the one-dimensional process $\{X(t)\}_{t \geq 0}$. Let $N \in \mathbb{N}$ and $k \in \{0, \ldots, N-1\}$. Furthermore, assume that the interval $[a, b]$*

sufficiently covers the transition density function corresponding to ϕ_X, such that it is almost zero outside of the interval. Then, the one-dimensional COS formula is given by

$$\bar{G}[H_k](x) = \sum_{k=0}^{N-1}{}' \,\mathsf{Re}\left[\phi_X\left(\frac{k\pi}{b-a}\right)\mathrm{e}^{-\mathrm{i}k\pi\frac{a}{b-a}}\right]H_k,$$

where $\sum'$ is the summation with the first term multiplied with one-half, $\mathsf{Re}[z]$ *corresponds to the real part of the complex number $z \in \mathbb{C}$ and H_k are the cosine series coefficients of a function g, defined as*

$$H_k = \frac{2}{b-a}\int_a^b g(x)\cos\left(k\pi\frac{y-a}{b-a}\right)\mathrm{d}y.$$

A full derivation is given by Oosterlee and Grzelak [32].

The notion of the COS method given in Theorem 4.6 is extended to two dimensions by Ruijter and Oosterlee [33]. Following the results presented by Ruijter and Oosterlee, also a two-dimensional version of the COS method can be given.

Theorem 4.7. *Consider a two-dimensional characteristic function $\phi_{\boldsymbol{X}} : \mathbb{R}^2 \to \mathbb{C}$. Let $N_1, N_2 \in \mathbb{N}$ and $(k_1, k_2) \in \{0, \ldots, N_1\} \times \{0, \ldots, N_2\}$. Let $[a_1, b_1] \times [a_2, b_2]$ be an interval, sufficiently capturing the transition density corresponding $\phi_{\boldsymbol{X}}$, that is, the density function approaches zero outside the defined interval. The two-dimensional COS formula is then given by*

$$\begin{aligned}\bar{G}_2[H_{k_1,k_2}](x,y) = \frac{1}{2}\sum_{k_1=0}^{N_1-1}{}'\sum_{k_2=0}^{N_2-1}{}' \,\mathsf{Re}\Bigg[&\phi_{\boldsymbol{X}}\left(\frac{k_1\pi}{b_1-a_1},\frac{k_2\pi}{b_2-a_2}\right)\mathrm{e}^{-\mathrm{i}k_2\pi\frac{a_2}{b_2-a_2}} \\ &+\phi_{\boldsymbol{X}}\left(\frac{k_1\pi}{b_1-a_1},-\frac{k_2\pi}{b_2-a_2}\right)\mathrm{e}^{\mathrm{i}k_2\pi\frac{a_2}{b_2-a_2}}\Bigg]H_{k_1,k_2},\end{aligned}$$

where the two-dimensional cosine series coefficients of the function $g : \mathbb{R}^2 \to \mathbb{R}$ are defined as

$$H_{k_1,k_2} = \int_{a_1}^{b_1}\int_{a_1}^{b_2} g(x,y)\cos\left(k_1\pi\frac{x-a_1}{b_1-a_1}\right)\cos\left(k_2\pi\frac{y-a_2}{b_2-a_2}\right)\mathrm{d}x\mathrm{d}y.$$

The derivation of the two-dimensional COS formula is given by Ruijter and Oosterlee [33]. The one-dimensional and two-dimensional COS methods will be used to give an alternative Fourier approach for valuing the corporate callable convertible bond.

4.4.1. Convertible bonds under constant interest rates

As with the CONV method (Section 4.3), first, the valuation method in a one-dimensional setting will be discussed. Recall the same setting as defined in Section 4.3, where the time grid is defined as $t_l = l\Delta_t$ for $\Delta_t = \frac{T}{\mathrm{M}-1}$ such that $l \in \{0, \ldots, \mathrm{M}-1\}$. Furthermore, recall the coupon payment dates $\mathcal{T}^{\mathsf{coup}}$ and convertible dates $\mathcal{T}^{\mathsf{conv}}$, which all coincide with the time steps. Let γ be the conversion factor and let the coupon payments equal αF for a fixed coupon rate α and face value F. Assume a constant interest rate $r(t) \equiv r \in \mathbb{R}$ and assume the company value process is defined by $\{A(t)\}_{t\geq 0}$ as described in Section 2.2. Similar to the CONV method, define the transformation

$$X_F(t) = \log\left(\frac{A(t)}{F}\right) \iff A(t) = F\mathrm{e}^{X_F(t)}, \tag{4.12}$$

where $t \in [t_0, T]$, $X_F(t_0) = \log(\frac{A_0}{F})$ is the initial value, for initial value $A(t_0) = A_0$ of the process $\{A(t)\}_{t\geq 0}$.

The scaled payoff of the bond at maturity T under the transformation in Equation (4.12) is given by

$$\bar{H}_{\mathsf{conv}}(T, X_F(T)) = \min\left(e^{X_F(T)}, \max(1+\alpha, \gamma\mathrm{e}^{X_F(T)})\right), \tag{4.13}$$

such that $H_{\text{conv}}(T, X_F(T)) = F\bar{H}_{\text{conv}}(T, X_F(T))$. Denote the scaled value of the convertible bond at time $t_0 < t_l < T$ as $V_{\text{conv}}(t_l, X_F(t_l))$. As the optimal exercise policy is to only consider the option to convert at maturity T, the intermediate values of the convertible bond are given by

$$\bar{V}_{\text{conv}}(t_l, y_k) = \begin{cases} \mathrm{e}^{X_F(t_l)} & \text{for } \mathrm{e}^{X_F(t_l)} < \alpha, \\ \alpha + \bar{C}(t_l, y_k) & \text{for } \mathrm{e}^{X_F(t_l)} \geq \alpha, \end{cases}$$

where $\bar{C}(t_l, y_k)$ is the scaled continuation value of the convertible bond, such that $C(t_l, y_k) = F\bar{C}(t_l, y_k)$ for $C(t_l, y_k)$ defined as in Definition 1.3.

To be able to apply the COS method, the cosine coefficients $V_k(t_1)$ need to be determined, such that the bond value can be approximated via,

$$V_{\text{conv}}(t_0, X(t_0)) \approx F\mathrm{e}^{-r\Delta_t}\bar{G}[V_k(t_1)](x), \tag{4.14}$$

using the cosine approximation method $\bar{G}[\cdot](x)$ as defined in Theorem 4.6. To obtain the cosine coefficients $V_k(t_1)$, a recursive scheme is proposed which will determine all intermediate coefficients $V_k(t_l)$ for $l \in \{1, \ldots, \mathrm{M}-1\}$. At the last step of the recursion, the coefficients $V_k(t_1)$ are approximated and the bond value can be determined via Equation (4.14). The recursion will be performed backward in time, starting at $t_{\mathrm{M}-2}$. Given the cosine coefficients, $V_k(t_{l+1})$ for $l \in \{1, \ldots, \mathrm{M}-2\}$, the continuation value of the convertible bond can be approximated by applying the COS method,

$$\bar{C}(t_l, X_F(t_l)) \approx \mathrm{e}^{-r\Delta_t}\bar{G}[V_k(t_{l+1})](x).$$

From Theorem 1.1, it is known that the convertible bond will not be exercised at any of the intermediate steps. Hence, the intermediate cosine coefficients for $t_l \in \mathcal{T}^{\text{coup}}$ are only dependent on the default and the continuation interval, i.e.

$$V_k(t_l) = D_k(t_l, \mathcal{D}) + P_k(t_l, \mathcal{C}),$$

where the default and the continuation interval $\mathcal{D} \cup \mathcal{C}$ are spanning the whole interval $[a, b]$ and

$$\begin{cases} D_k(t_l, x_1, x_2) & = \dfrac{2}{b-a}\displaystyle\int_{x_1}^{x_2} \mathrm{e}^y \cos\left(k\pi\frac{y-a}{b-a}\right)\mathrm{d}y, \\ P_k(t_l, x_1, x_2) & = \dfrac{2}{b-a}\displaystyle\int_{x_1}^{x_2} \alpha\cos\left(k\pi\frac{y-a}{b-a}\right)\mathrm{d}y + C_k(t_l, x_1, x_2), \\ C_k(t_l, x_1, x_2) & = \dfrac{2}{b-a}\displaystyle\int_{x_1}^{x_2} C(t_l, x)\cos\left(k\pi\frac{y-a}{b-a}\right)\mathrm{d}y. \end{cases} \tag{4.15}$$

When $t_l \notin \mathcal{T}^{\text{coup}}$, the value coefficients equal the continuation coefficients C_k over the whole interval $[a, b]$, i.e.

$$V_k(t_l) = C_k(t_l, a, b). \tag{4.16}$$

The continuation interval $\mathcal{C}$ and default interval $\mathcal{D}$ can be determined based on the expression for $V(t_l, X_F(t_l))$ at the intermediate times t_l. As the company will default in case $\mathrm{e}^{X_F(t_l)} < \alpha$, the default interval and continuation interval are given by,

$$\mathcal{D} = [a, \log(\alpha)),\ \mathcal{C} = [\log(\alpha), b].$$

The coefficients $D_k(t_l, \mathcal{D})$ over the default interval $\mathcal{D}$ can be computed in closed form. The result is given in Lemma 4.3

Lemma 4.3. *Let $D_k(t_l, x_1, x_2)$ be given as in Equation* (4.15)*, then for all $l \in \{1, \ldots, \mathrm{M}-2\}$, the coefficients $D_k(t_l, x_1, x_2)$ are known in closed form and given by*

$$D_k(t_l, x_1, x_2) = \frac{2}{b-a}\chi_k(x_1, x_2),$$

where the function χ_k is given in Lemma C.1.

Lemma 4.3 follows directly from Lemma C.1. Substitution of the COS approximation of $\bar{C}(t_l, X_F(t_l))$ in Equation (4.15) and performing some basic analytical derivations, yield the following result.

Lemma 4.4. *The cosine coefficients for the continuation value $C_k(t_l, x)$ at time t_l can be approximated based on the cosine value coefficients $V_j(t_{l+1})$ at time step t_{l+1} via,*

$$C_k(t_l, x_1, x_2) \approx \mathrm{e}^{-r\Delta_t} {\sum_{j=0}^{N-1}}' \mathsf{Re}\left[\phi_X\left(\frac{j\pi}{b-a}\right) V_j(t_{l+1}) \mathcal{A}_{j,k}(t_l, x_1, x_2)\right].$$

Here the coefficients $\mathcal{A}_{j,k}(t_l, x)$ are known in closed form and given by

$$\mathcal{A}_{j,k}(t_l, x_1, x_2) = -\frac{\mathrm{i}}{\pi}(\tilde{\mathcal{A}}^{+}_{j,k}(t_l, x_1, x_2) + \tilde{\mathcal{A}}^{-}_{j,k}(t_l, x_1, x_2)).$$

The coefficients $\tilde{\mathcal{A}}^{+}_{j,k}(t_l, x_1, x_2)$ and $\tilde{\mathcal{A}}^{-}_{j,k}(t_l, x_1, x_2)$ can be written in matrix form,

$$\tilde{\mathcal{A}}^{-}_{j}(t_{t_l}, x_1, x_2) = \begin{bmatrix} \mathrm{A}_0 & \mathrm{A}_1 & \cdots & \mathrm{A}_{N-1} \\ \mathrm{A}_{-1} & \mathrm{A}_0 & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathrm{A}_{1-N} & \cdots & & \mathrm{A}_0 \end{bmatrix}, \; \tilde{\mathcal{A}}^{+}_{j}(t_l, x_1, x_2) = \begin{bmatrix} \mathrm{A}_0 & \mathrm{A}_1 & \cdots & \mathrm{A}_{N-1} \\ \mathrm{A}_1 & \mathrm{A}_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathrm{A}_{N-1} & \cdots & & \mathrm{A}_{2N+2} \end{bmatrix},$$

where,

$$\mathrm{A}_j = \begin{cases} \dfrac{x_2 - x_1}{b-a}\mathrm{i}\pi & j = 0, \\ \dfrac{\mathrm{e}^{\mathrm{i}j\pi\frac{x_2-a_1}{b-a}} - \mathrm{e}^{\mathrm{i}j\pi\frac{x_1-a_1}{b-a}}}{j} & j \neq 0. \end{cases}$$

The proof of Lemma 4.4 is given in Appendix D.12. The matrices $\tilde{\mathcal{A}}^{+}_{j,k}(t_l, x_1, x_2)$ and $\tilde{\mathcal{A}}^{-}_{j,k}(t_l, x_1, x_2)$ are Hankel and Toeplitz matrices which enables the possibility to compute them by use of the DFT defined in Definition 4.2. The coefficients for $P_k(t_l, x_1, x_2)$ can now be computed straightforwardly using the results of Equation (4.15) and Theorem 4.4.

Lemma 4.5. *Given the continuation coefficients given in Lemma 4.4, the cosine coefficients $P_k(t_l, x_1, x_2)$ are given by*

$$P_k(t_l, x_1, x_2) = \alpha\frac{2}{b-a}\psi_k(x_1, x_2) + C_k(t_l, x_1, x_2),$$

where ψ_k is given in Lemma C.1.

The proof of Lemma 4.5 follows directly from the definitions of the functions $\psi_k(x_1, x_2)$ and $C_k(t_l, x_1, x_2)$.

To enable the recursion, only the coefficients for $V_k(t_{\mathrm{M}-1})$ are left for computation. Fortunately, the coefficients $V_k(t_{\mathrm{M}-1})$ are known in closed form. The result is given in Theorem 4.8

Theorem 4.8. *Let the scaled payoff function at maturity T be given as in Equation* (4.13)*. The cosine coefficients $V_k(t_T)$ at maturity T are given by the integral,*

$$V_k(T) = \frac{2}{b-a}\int_a^b \bar{H}_{\mathsf{conv}}(T, x)\cos\left(k\pi\frac{y-a}{b-a}\right)\mathsf{d}y,$$

for $a, b \in \mathbb{R}$ chosen as described in Theorem 4.6 and where $\bar{H}$ is given in Equation (4.13) *is known in closed form and given by*

$$V_k(T) = \frac{2}{b-a}\left[\chi_k(a, \log(1+\alpha)) + (1+\alpha)\psi_k\left(\log(1+\alpha), \log\left(\frac{1+\alpha}{\gamma}\right)\right) + \gamma\chi_k\left(\log\left(\frac{1+\alpha}{\gamma}\right), b\right)\right].$$

Furthermore, the functions χ_k and ψ_k are known in closed form and given in Lemma C.1.

The proof of Theorem 4.8 is given in Appendix D.13.

4.4.2. Callable convertible bonds under constant interest rates

Denote the scaled value of the callable convertible bond at time t_l by $\bar{V}_{\text{cc}}(t_l, X_F(t_l))$, such that $V_{\text{cc}}(t_l, X_F(t_l)) = F\bar{V}_{\text{cc}}(t_l, X_F(t_l))$ is the value of the callable convertible bond. At maturity T, the value of the callable convertible bond equals the value of a convertible bond as the callable option can only be exercised prematurely. The scaled payoff of the callable convertible bond at maturity T is therefore given by $\bar{H}_{\text{conv}}$ as given in Equation (4.13). The value $\bar{V}_{\text{cc}}(t_l, X_F(t_l))$ of the bond at each intermediate time step $t_0 > t_l > T$ is defined by

$$\bar{V}_{\text{cc}}(t_l, X_F(t_l)) = \begin{cases} \mathrm{e}^{X_F(t_l)} & \text{for } \mathrm{e}^{X_F(t_l)} < \alpha, \\ \max(\beta(t_l) + \alpha, \gamma \mathrm{e}^{X_F(t_l)}) & \text{for } \bar{C}_{\text{call}}(t_l, X_F(t_l)) \geq \beta(t_l) + \alpha \text{ and } \mathrm{e}^{X_F(t_l)} \geq \alpha, \\ \gamma \mathrm{e}^{X_F(t_l)} & \text{for } \bar{C}(t_l, X_F(t_l)) \geq \gamma \mathrm{e}^{X_F(t_l)},\ \mathrm{e}^{X_F(t_l)} \geq \alpha \\ & \qquad \text{and } \bar{C}_{\text{call}}(t_l, X_F(t_l)) < \beta(t_l) + \alpha, \\ \alpha + \bar{C}(x_j, y_k) & \text{otherwise,} \end{cases}$$

where $\alpha \in [0, 1]$ is the coupon rate, γ is the conversion rate, β is the call penalty function and the continuation values for the convertible $C(t_l, A(t_l))$ and the call option $C^{\text{call}}(t_l, X_F(t_l))$ are given in Definition 1.4 and Definition 1.5 respectively.

Similar to the convertible bond, the callable convertible bond values will be obtained by using the relation given in Equation (4.14). Therefore, a recursive scheme backward in time will be presented to obtain the cosine coefficients for $V_k(t_1)$. Note that the coefficients at maturity are known via Theorem 4.8. To obtain the recursion scheme, assume that $V_k(t_{l+1})$ and $V_k(t_l + t_c)$ are known. The objective is to find cosine coefficients for $t_l \in \mathcal{T}^{\text{coup}}$ such that,

$$V_k(t_l) = D_k(t_l, \mathcal{D}) + G_k^{\text{call}}(t_l, \mathcal{G}_{\text{call}}) + G_k^{\text{conv}}(t_l, \mathcal{G}_{\text{conv}}) + P_k(t_l, \mathcal{C}),$$

for default interval $\mathcal{D}$, call interval $\mathcal{G}_{\text{call}}$, converting interval $\mathcal{G}_{\text{conv}}$ and continuation interval $\mathcal{C}$ such that,

$$\mathcal{D} \cup \mathcal{G}_{\text{call}} \cup \mathcal{G}_{\text{conv}} \cup \mathcal{C} = [a, b],$$

and where,

$$\begin{cases} D_k(t_l, x_1, x_2) & = \dfrac{2}{b-a} \displaystyle\int_{x_1}^{x_2} \mathrm{e}^y \cos\left(k\pi \frac{y-a}{b-a}\right) \mathrm{d}y, \\ G_k^{\text{call}}(t_l, x_1, x_2) & = \dfrac{2\alpha}{b-a} \displaystyle\int_{x_1}^{x_2} \max(\beta(t) + \alpha, \gamma \mathrm{e}^y) \cos\left(k\pi \frac{y-a}{b-a}\right) \mathrm{d}y, \\ G_k^{\text{conv}}(t_l, x_1, x_2) & = \dfrac{2}{b-a} \displaystyle\int_{x_1}^{x_2} \gamma \mathrm{e}^y \cos\left(k\pi \frac{y-a}{b-a}\right) \mathrm{d}y, \\ P_k(t_l, x_1, x_2) & = \dfrac{2}{b-a} \displaystyle\int_{x_1}^{x_2} \alpha \cos\left(k\pi \frac{y-a}{b-a}\right) \mathrm{d}y + C_k(t_l, x_1, x_2), \\ C_k(t_l, x_1, x_2) & = \dfrac{2}{b-a} \displaystyle\int_{x_1}^{x_2} C(t_l, x) \cos\left(k\pi \frac{y-a}{b-a}\right) \mathrm{d}y. \end{cases} \tag{4.17}$$

At time steps $t_l \notin \mathcal{T}^{\text{coup}}$, the cosine coefficients can be computed according to Equation (4.16).

Note that, given the intervals $\mathcal{D}$ and $\mathcal{C}$ the coefficients D_k and P_k can be directly obtained by applying Lemma 4.3 and Lemma 4.5 directly. The coefficients G_k^{call} and G_k^{conv} can also be obtained by straightforward integration. The result is given in Lemma 4.6.

Lemma 4.6. *The cosine coefficients G_k^{call} and G_k^{conv}, given in Equation* (4.17)*, can be determined in closed form and are given by*

$$G_k^{\text{call}}(t_l, x_1, x_2) = \frac{2}{b-a} \begin{cases} \left[(\beta(t_l) + \alpha) \psi_k \left(x_1, \log\left(\dfrac{\beta(t_l) + \alpha}{\gamma} \right) \right) \right. & \text{for } x_1 < \log\left(\frac{\beta(t_l)+\alpha}{\gamma}\right), \\ \qquad \left. + \gamma \chi_k \left(\log\left(\dfrac{\beta(t_l) + \alpha}{\gamma} \right), x_2 \right) \right] & \\ \gamma \chi_k(x_1, x_2) & \text{otherwise,} \end{cases}$$

and

$$G_k^{\text{conv}}(t_l, x_1, x_2) = \frac{2\gamma\chi_k(x_1, x_2)}{b-a},$$

respectively, where the functions ψ_k and χ_k are given in Lemma C.1.

The proof of Lemma 4.6 is given in Appendix D.14. It is therefore left to obtain the interval $\mathcal{G}_{\text{call}}$ and $\mathcal{G}_{\text{conv}}$.

Recall the optimal exercising policy for the call option and the convertible option discussed in Section 1.2.4. Given the call option, the management of the company will exercise the call option if,

$$K(t) + \alpha F - C^{\text{call}}(t_l, x_{\text{call}}) = 0. \tag{4.18}$$

Hence, solving Equation (4.18) yields the early exercise point $x_{\text{call}} \in [a, b]$ for which it is optimal to call the bond. On the other hand, by solving,

$$C(t_l, X(t_l)) - \gamma \mathrm{e}^{X(t_l)} = 0, \tag{4.19}$$

the early exercise point x_{conv} can be found for which it is optimal to exercise the convertible option. As the options within the bond cannot be exercised both, the corresponding intervals $\mathcal{G}_{\text{call}}$ and $\mathcal{G}_{\text{conv}}$ need to be mutually exclusive. Moreover, both options will be exercised only if the value at immediate exercise is large enough. Therefore, it will hold that,

$$\mathcal{D} = [a, \log(\alpha)],\ \mathcal{C} = [\log(\alpha), x_*],\ \mathcal{G}_{\text{call}} = [x_*, x^*],\ \mathcal{G}_{\text{conv}} = [x^*, b],$$

where

$$x^* = \max(x_{\text{conv}}, x_{\text{call}}),\ x_* = \min(x_{\text{conv}}, x_{\text{call}}).$$

The early exercise points x_{call} and x_{conv} can be found using a numerical solver for Equation (4.18) and Equation (4.19). Note that the continuation value corresponding to the call option and the continuation value corresponding to the convertible option can be approximated by,

$$\begin{cases} C_{\text{call}}(t_l, x) & \approx \mathrm{e}^{-r(\Delta_t + t_c)} \bar{G}[V_k(t_l + t_c)](x), \\ C(t_l, x) & \approx \mathrm{e}^{-r\Delta_t} \bar{G}[V_k(t_{l+1})](x), \end{cases}$$

respectively, using the COS approximation method given in Theorem 4.6. Hence, the derivative for x can be determined. Using the derivatives for $\dfrac{\partial C_{\text{call}}}{\partial x}$ and $\dfrac{\partial C}{\partial x}$, Newton-Raphson can be used to obtain the values for x_{call} and x_{conv} respectively.

4.4.3. Zero-coupon convertible bond under stochastic interest rates

The two-dimensional COS method for Bermudan-type options may become significantly more complex and is considered beyond the scope of this thesis. However, if the convertible bond is considered without any coupon payments, the problem reduces to one that is similar to the valuation of a European type of option. Consider the transformation X_F given in Equation (4.12). If no coupon payments are assumed, i.e. $\alpha = 0$, the value of the convertible option $V(t_0, X(t_0))$ without any call option becomes,

$$V_{\text{conv}}(t_0, X_F(t_0)) = \mathbb{E}^T \left[\mathrm{e}^{-\int_{t_0}^T r(s)\mathrm{d}s} H_{\text{conv}}(T, X_F(T)) \,\middle|\, \mathcal{F}_0 \right],$$

under the equivalent martingale measure $\mathbb{Q}$, where H is the payoff function of the convertible bond,

$$H_{\text{conv}}(T, X_F(T)) = F \min(\mathrm{e}^{X_F(T)}, \max((1+\alpha), \gamma \mathrm{e}^{X_F(T)})),$$

where γ is the conversion factor and F is the face value of the bond. The value of the option can then be approximated via the two-dimensional COS method, which is given in Theorem 4.7, with cosine coefficients,

$$H_{k_1,k_2} = \int_{a_1}^{b_1} \int_{a_1}^{b_2} H_{\text{conv}}(T, y) \cos\left(k_1 \pi \frac{x - a_1}{b_1 - a_1}\right) \cos\left(k_2 \pi \frac{y - a_2}{b_2 - a_2}\right) \mathrm{d}x\mathrm{d}y, \tag{4.20}$$

for $N_1, N_2 \in \mathbb{N}$, $(k_1, k_2) \in \{0, \ldots, N_1\} \times \{0, \ldots, N_2\}$ and bounds $a_1, b_1, a_2, b_2 \in \mathbb{R}$. In particular, the cosine coefficients H_{k_1,k_2} can be computed in closed form.

Lemma 4.7. *Let H_{k_1,k_2} be given as in Equation* (4.20)*, then the integrals can be computed in closed form. In particular, it holds that*

$$H_{k_1,k_2} = F\psi_{k_1}(a_1,b_1)\left[\chi_{k_2}(a_2,0) + \psi_{k_2}\left(0,\log\left(\frac{1}{\gamma}\right)\right) + \gamma\chi_{k_2}\left(\log\left(\frac{1}{\gamma}\right),b_2\right)\right],$$

where F is the face value of the bond, γ is the conversion factor and ψ_k and χ_k are given in Lemma C.1.

Proof. The proof of Lemma 4.7 follows directly by observing that the integrals can be split and therefore can be written as multiplication. Applying Lemma C.1 for the interest rate part and using the results of Theorem 4.8 yields the expression for H_{k_1,k_2}. □

Lemma 4.7 and Theorem 4.7, yield the possibility to obtain the the value of the convertible bond with coupon rate $\alpha = 0$ numerically, where the characteristic function $\phi_{\boldsymbol{X}}$ of the process $\{X_F(t)\}_{t\geq 0}$ under the Black-Scholes Hull-White model is used as given in Theorem 2.3.

5

Numerical results

5.1. Introduction

The previous chapters have described the modeling of the callable convertible bond and numerical methods to evaluate them. In this section, the numerical results will be presented. Numerical results will be obtained for both the one-dimensional approach and the two-dimensional approach. First, to verify the correctness of the characteristic functions, the two-dimensional densities will be approximated. In particular, two density planes will be demonstrated; that of the Black-Scholes Hull-White model defined in Section 2 and that of the transformed Black-Scholes Hull-White model under the transformation defined in Theorem 2.5 (see Section 2.3.3). The density approximations are obtained by using the Fourier approximation and are verified by the Monte Carlo techniques discussed in Chapter 3.

In the subsequent section, the value of different bonds will be computed. First, the convertible and callable convertible bonds will be approximated under constant interest rates. A convergence scheme will be presented to analyze the convergence of the COS and the CONV method. Moreover, the sensitivity to the hyper-parameters will be discussed. Finally, the two-dimensional approach will be presented. Numerical results under the CONV method will be given. To conduct a convergence analysis, the non-coupon paying convertible bond under the Black-Scholes Hull-White model will be discussed. Also in the two-dimensional case, the hyper-parameters will be discussed.

5.2. Black-Scholes Hull-White density recovery

Using the CONV method, the density of both the jump-diffusion and the Black-Scholes Hull-White model can easily be obtained. Let $f_{\tilde{r},X}$ be the density function of the Black-Scholes Hull-White (BSHW) model under the decomposed interest rate process $\{\tilde{r}(t)\}_{t\geq 0}$ as defined in Section 2. Recall the definition of the inverse Fourier transform, given in Definition 4.3. Following Theorem 4.2, the one-dimensional density of the model can be obtained by computing the approximation $\bar{\mathrm{F}}_2^{-1}[\phi_{\tilde{r},X}](x_j)$ over the grids $x_j = (j - \frac{N}{2})\Delta_x$ and $y_k = (k - \frac{N}{2})\Delta_y$, where is used that $N_0 = N_1 \equiv N$, where $\phi_{\tilde{r},X}$ is obtained by combining Theorem 2.4 and the results given in Section 2.3.4. Specific values for a density $f_{\tilde{r},X}(x)$ can be obtained using the approximation of the Fourier transform, given in Theorem 4.2 and via Monte Carlo techniques discussed in Chapter 3. It remains to choose a grid size Δ_x and Δ_y for the approximations under the Fourier approach. Note that if the grid size is too small, significant values of the density may lay outside the grid which may result in bad approximations. On the other hand, if the grid size is chosen too large, a denser integration grid will be needed to obtain accurate results. As a consequence, the following rule of thumb will be followed.

Remark 5.1 (Simple integration grid size)**.** *Consider the two-dimensional grid given in Theorem 4.2. To obtain sufficiently accurate results for the Fourier approximations, the grid size for Δ_x and Δ_y is chosen such that*

$$\Delta_x = \frac{2L_x\eta\sqrt{T}}{N},\ \Delta_y = \frac{2L_y\sigma\sqrt{T}}{N},$$

for $L_x, L_y \in \mathbb{R}_+$ the number of standard deviations away from the mean, σ the volatility of the process $\{X(t)\}_{t\geq 0}$, η the volatility of the process $\{\tilde{r}(t)\}_{t\geq 0}$ and T the maturity. Furthermore, $N \in \mathbb{N}$ is the number of nodes included in the grid, where the same number of nodes is chosen for both grids.

When the grid sizes are set, the Fourier transform can be applied on the characteristic function $\phi_{\tilde{r},X}$ to obtain a transition density from the initial values $(\tilde{r}(t_0), X(t_0))$ to the value $(\tilde{r}(T), X(T))$. Moreover, when values for $(\tilde{r}(T), X(T))$ are directly sampled using the Monte Carlo method discussed in Chapter 3, a reference density can be obtained to check the results. The obtained density is given in Figure 5.1. To better compare the results, Figure 5.2 yields the contour plots of both the simulation

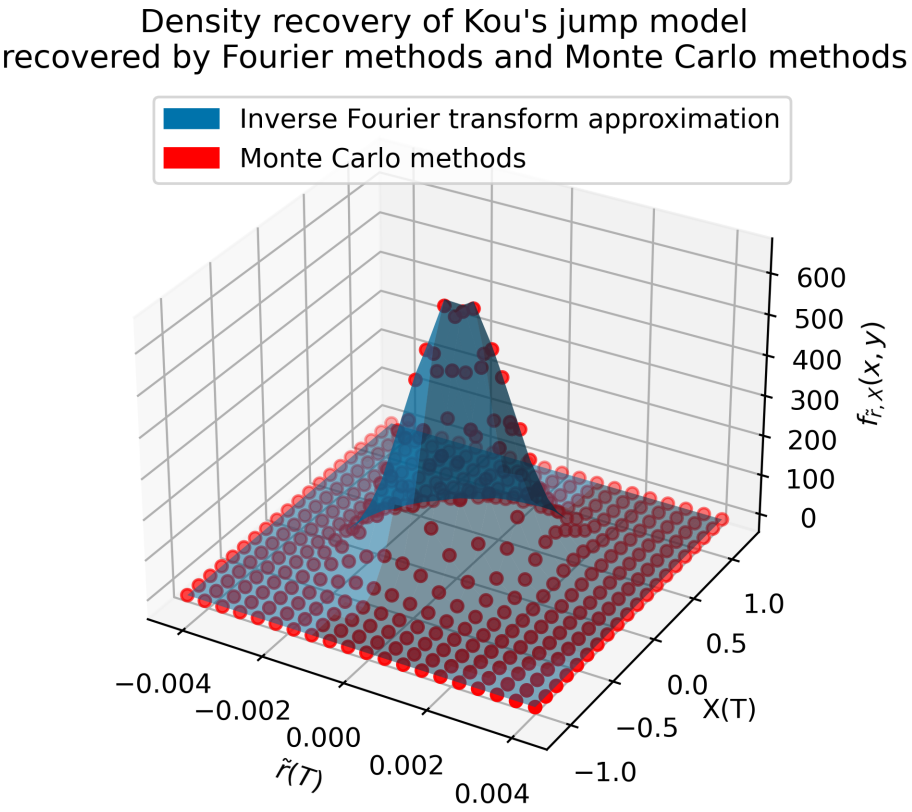
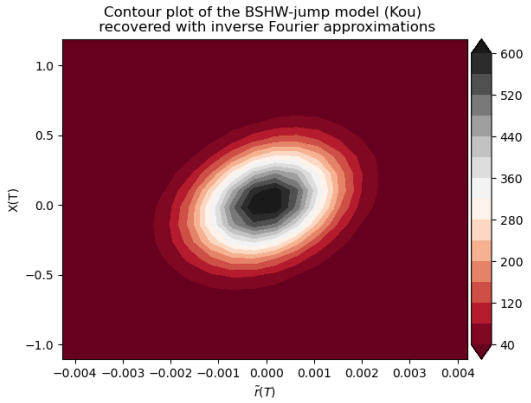


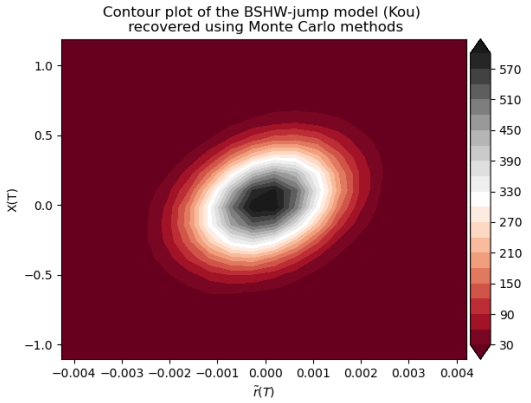
Figure 5.1: Three-dimensional plot of the density plane of the BSHW model with jumps distributed following Kou's model. The blue surface represents the results obtained via the inverse Fourier approximation. The red dots represent the results obtained via simulation and binned such that a scatter plot in three dimensions could be made. The parameters of the BSHW model used were; $S_0 = 1, \sigma = 0.25, \rho = 0.3$, where the Hull-White model is given for parameters $\eta = 0.001$ and $\kappa = 0.05$ and with maturity $T = 1$. The jump parameters under Kou's model are given by $\lambda = 3, \alpha_1 = 50, \alpha_2 = 33$ and $p_1 = 0.4$. The parameters used for the inverse Fourier approximation were set to $N = 2^9 = 512$ and $L_x = L_y = 10$. The simulations were computed using $M = 500000$ paths with each $N = 50$ nodes.

and the Fourier inversion results. From the contour plots in Figure 5.2 it can be seen that the density function is sufficiently captured by the Fourier approximation approach. Besides the jump-diffusion proposed by Kou, also the jump model proposed by Merton is verified. The corresponding figures are included in Appendix E.1. From these figures, it can also be seen that the BSHW model extended with a jump-diffusion model proposed by Merton can be sufficiently captured using the Fourier approximation techniques [29].

In addition, the transformed model described in Section 2.3.3 can be obtained by both Fourier and simulation techniques. As the jump part of the model is independent of the diffusion part of the model, the characteristic $\phi_{Z_r,Z}$ after the transformation (Z_r, Z) can be multiplied with the characteristic function concerning the specified jump model to obtain the transformed model with jumps. Moreover, due to



(a) Contour plot of the results obtained using the inverse Fourier approximation.



(b) Contour plot of the results obtained using the Monte Carlo techniques.

Figure 5.2: Both the contour plots of using the Fourier approximation and simulation techniques. The parameters used were described in Figure 5.1.

the independence of the jumps, the transformation can be written as

$$Z_r(T) = Z_r^C(T),\ Z(T) = Z^C(T) + Z^J(T),$$

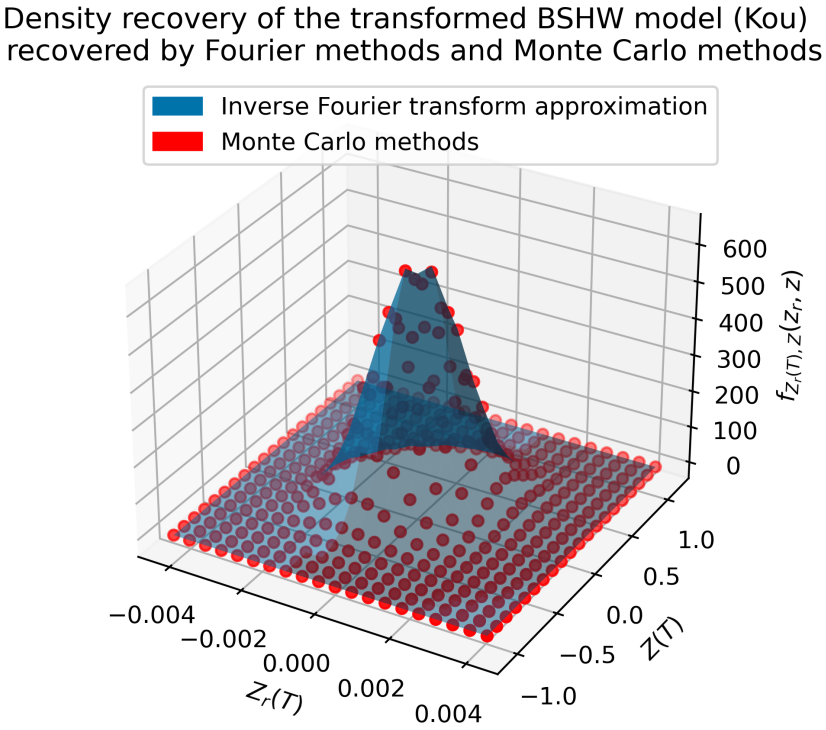
where $Z_J(T) = X^J(T)$, as no transformation will be applied to the jump part. The simulation results can therefore be obtained by applying the transformation,

$$\begin{cases} Z_r(T) & = \tilde{r}(T) - g_r(\tilde{r}(t_0)), \\ Z^C(T) & = X^C(T) - g(\tilde{r}(t_0), X^C(t_0)), \end{cases}$$

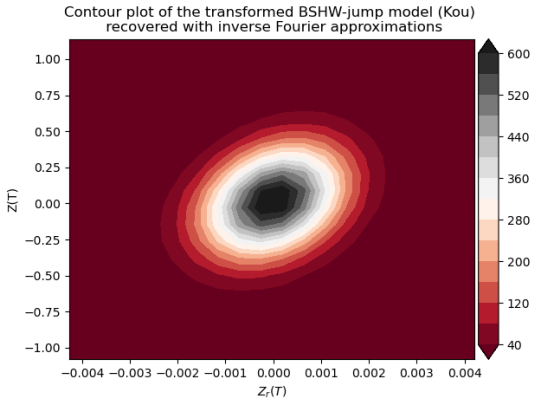
where the functions g_r and g are given by,

$$g_r(x) = x\mathrm{e}^{-\kappa\tau},\ g(x,y) = y + B(\tau)x,$$

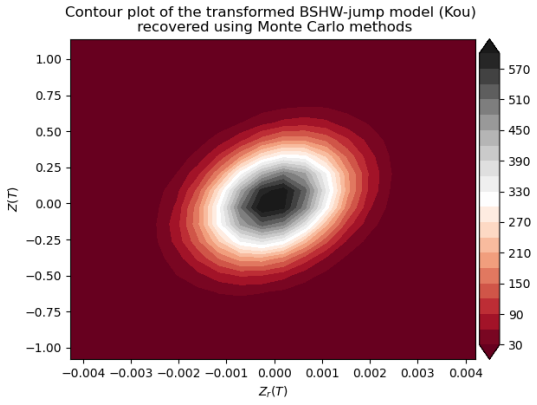
for $B(\tau) = \frac{1}{\kappa}(1 - \mathrm{e}^{-\kappa\tau})$. The resulting density recovery is given in Figure 5.3 for the transformation extended with the jump model proposed by Kou [22].



(a) Density recovery of the transformed variables (Z_r, Z) using Fourier and simulation techniques. The red dots yield the, binned, simulation results. The blue surface corresponds to the numerical Fourier approximation.



(b) Contour plot of the numerical results obtained by using the Fourier approximation approach.



(c) Contour plot of the numerical results obtained using the Monte Carlo techniques.

Figure 5.3: Numerical recovery of the density of the transformed random variables (Z_r, Z) with both the numerical approximation of the continuous Fourier transform and simulation. The transformed BSHW model is extended with Kou's jump model described in Section 2.3.4. The parameters of the BSHW model are given by; $S_0 = 1, \sigma = 0.25, \rho = 0.3, \eta = 0.001, \kappa = 0.05$ and $T = 1$. Furthermore, the jump parameters used were $\lambda = 3, \alpha_1 = 50, \alpha_2 = 33$ and $p_1 = 0.4$. The hyper-parameters of the Fourier approximation method used are $N = 2^9 = 512$ and $L_x = L_y = L = 10$. The simulation is conducted with $M = 500000$ paths each containing $N = 50$ nodes.

From Figure 5.3 it can be seen that the transformed density can be sufficiently captured using Fourier techniques. Figure 5.3 also shows that the split of the jump and continuous part due to the independence of the jump part is well captured under the transformation. For the transformed model, also the density recovery is done when the Merton jump-model is incorporated. The results are included in

Appendix E.2.

5.3. Numerical results under constant interest rates

Using the results of Chapter 3 and Chapter 4, the convertible bond under the structural model defined in Chapter 2 can be numerically obtained. Assume a constant interest rate r and the process $\{X(t)\}_{t\geq 0}$ as defined in Section 2.2. First, the less complex convertible bond without the call option will be computed. The numerical results will be obtained under Kou's jump model described in Chapter 2. As the jump additions are independent, the results can be interchanged with each other jump model. The parameters corresponding to the jump model are given in Table 5.1.

λ	p_1	α_1	α_2
3	0.4	50	33

Table 5.1: Parameters used for the jump component in the structural model. The jumps are assumed to follow a double exponential distribution with parameters as given. The parameters chosen are based on the parameters given by Ballotta and Kyriakou [3].

The diffusion and the interest parameter are given in Table 5.2.

r	σ	S_0
0.05	0.25	10

Table 5.2: The parameters used for the diffusion part of the structural model. Moreover, the interest rate is assumed to be constant and given by r.

Following Section 4.3.1 and Section 4.4.1 the convertible bond prices can be obtained under different numerical approximation methods. Furthermore, the results can be checked using the results for simulation obtained in Chapter 3. The bond characteristics are given in Table 5.3

α	γ	T
0.01	0.3	5

Table 5.3: Parameters for the convertible bond characteristics.

The integration grids for the COS and CONV method are based on cumulants to sufficiently capture the tails of the jump distribution.

Remark 5.2 (Cumulant based integration grid size)**.** *Consider the one-dimensional grid given in Theorem 4.1. To obtain sufficiently accurate results for the Fourier approximations, the grid size for Δ_x is chosen such that*

$$\Delta_x = \frac{2L_x\sqrt{\zeta_2(u,T) + \sqrt{\zeta_4(u,T)}}}{N},$$

for $L_x \in \mathbb{R}_+$, N the number of grid points, ζ_2 and ζ_4 the second and fourth cumulant function which is given by

$$\zeta_j(u,T) = \frac{1}{\mathrm{i}^j}\frac{\partial^j \log(\phi_X(u,T))}{\partial u^j}\bigg|_{u=0},$$

for $j \in \mathbb{N}$. Furthermore, the integration grid for the COS method can be derived in a similar way,

$$\begin{cases} a &= -\hat{L}\sqrt{\zeta_2(u,T) + \sqrt{\zeta_4(u,T)}}, \\ b &= \hat{L}\sqrt{\zeta_2(u,T) + \sqrt{\zeta_4(u,T)}}, \end{cases}$$

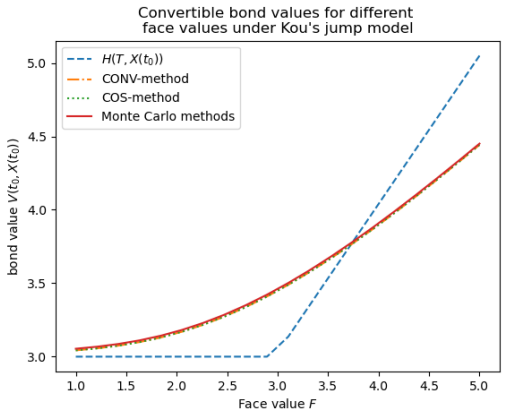
for $\hat{L} \in \mathbb{R}_+$. Note that $\hat{L}$ used for the COS method and L_x used for the CONV method, may differ.

Cumulants for often used jump models are given by Oosterlee and Grzelak [32].

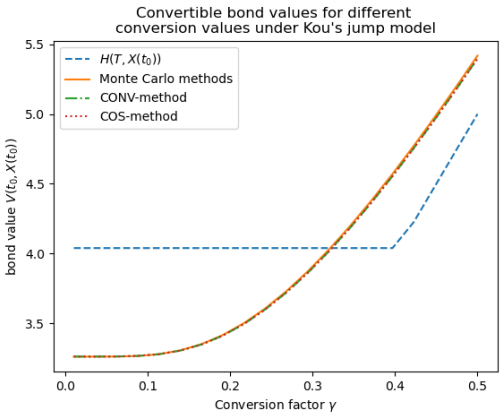
The coupon scheme is set semi-annual, which means that a yearly coupon payment results in two coupon payments per year of half the coupon rate. Each year, the first payment of the bond is halfway through the year and the second payment is at the end of each year. As a consequence, the coupon and exercise dates are given by

$$\mathcal{T}^{\text{coup}} = \mathcal{T}^{\text{conv}} = \left\{ \tau_l = \frac{1}{2} + \frac{1}{2}l \;\middle|\; \text{where } t_0 < \tau_l \leq T, \text{ for } l \in \{0, \ldots, \mathrm{M}-1\} \right\}.$$

Note that the set $\mathcal{T}^{\text{coup}}$ does not necessarily coincide with a chosen time grid. When the time steps do not coincide with $\mathcal{T}^{\text{coup}}$, the first date is chosen that is included in the time grid and is directly after the actual coupon date. The results of the CONV and simulation method are given in Figure 5.4. Often, convertible bonds are based on the conversion factor γ. In 5.4 b, a single value $F = 4$ is chosen to compare the value of the convertible bond for different values of γ.



(a) A plot of the convertible bond values for different values of F. The CONV method is based on a grid with $L_x = 10$ and $N = 2^9 = 512$ grid nodes. The COS method is performed with $N = 2^6 = 64$ grid nodes and $\hat{L} = 10$.



(b) A plot of the convertible bond for different values of γ. The CONV method is based on a grid with $L_x = 10$ and $N = 2^9 = 512$ grid nodes. The COS method is performed with $N = 2^6 = 64$ grid nodes and $\hat{L} = 10$.

Figure 5.4: A plot of a convertible bond approximated by the CONV method. The approximated values using the CONV method are compared to the approximated values using simulation as discussed in Chapter 3. For all methods $\mathrm{M} = 51$ time steps are considered. The simulations are conducted under $\mathrm{N} = 500000$ paths. The blue dashed line corresponds with the payoff function at maturity $H(T, \cdot)$ given in Equation (4.6), and therefore shows the value of the bond at immediate payoff.

From Figure 5.4 it can be seen that the CONV method and the Monte Carlo method give very similar results. In particular, the maximum absolute error of the comparison between the CONV and the simulation results is 0.0125 over the different values of F. Similar results were obtained for the COS method. Note that for small face values the bond will be converted almost definitely. When the face value exceeds the conversion value $\gamma A(t)$, the payoff becomes linear in the face value F. Hence, for larger face values the value of the bond will show more linearity. Figure 5.4 b shows that for low conversion factors, the value of the convertible bond is almost equal to the discounted face value F, which corresponds to accepting the face value at maturity T. For large values of the conversion factor γ, it becomes more and more interesting for the investor to convert the bond. As a consequence, linearity can be seen in 5.4 b for large values of the conversion factor γ.

Subsequently, values for the callable convertible bond will be computed. Using the approach described in Chapter 4, the callable convertible bond values can be approximated using the CONV and the COS method. Furthermore, the simulation method described in Chapter 3 can be used to verify the results. The call option is assumed to be exercisable only on the coupon dates, i.e. $\mathcal{T}^{\text{call}} = \mathcal{T}^{\text{coup}}$. Furthermore, it is assumed that the call payoff function is given by

$$K(t) = 1.2F,$$

where F is the face value of the bond. The results are given in Figure 5.5

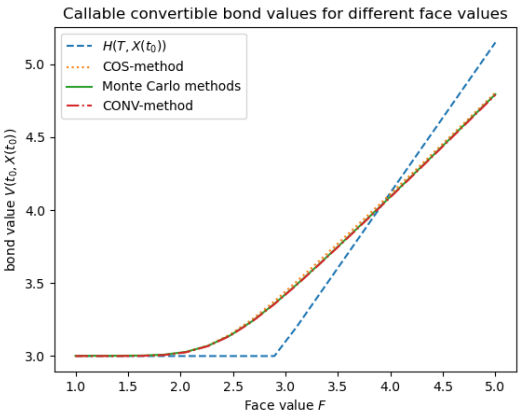


Figure 5.5: A plot of the convertible bond for different values of F. The CONV method is based on a grid with $L_x = 10$ and $N = 2^9 = 512$ grid nodes. The COS method is performed with $N = 2^6 = 64$ grid nodes and $\hat{L} = 10$. For all methods $\mathrm{M} = 51$ time steps are considered. The simulations are conducted under $\mathrm{N} = 500000$ paths. The blue dashed line corresponds with the payoff function at maturity $H(T, \cdot)$ given in Equation (4.6), and therefore shows the value of the bond at immediate payoff.

The results given in Figure 5.5 are very similar to the results for the Convertible bond. For small values of the face value F, the bond will be converted almost certainly. Since the bond also contains a callable option, the value of the callable convertible is lower than the value of a convertible bond without a callable option. When the face value F becomes larger, it becomes more and more interesting for the investor to not convert and accept the call price upon call. If the bond is not called, the bond will still not be converted and the face value will be accepted upon maturity.

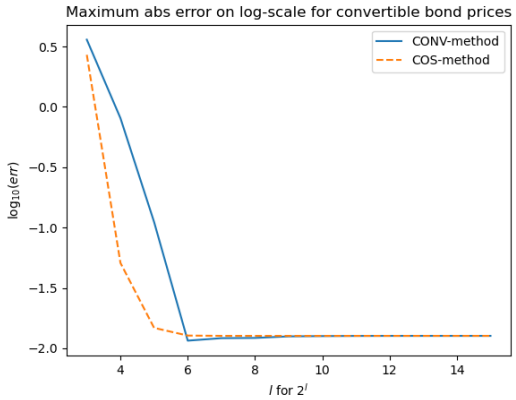
The call interval $\mathcal{G}_{\text{call}}$ and the converting interval $\mathcal{G}_{\text{conv}}$ are computed numerically. Both intervals are determined by the early exercise point x_{conv} and x_{call}, which are approximated using the Newton-Raphson method. During the implementation of the COS method, it appeared that in particular the computation of the early exercise point x_{call} is highly unstable. Moreover, wrong estimations of the early exercise point x_{call} will result in wrong approximations of the intermediate cosine coefficients which then will result in wrong approximations by the COS method. As a consequence, for the callable convertible bond, the convergence of the COS method is far less stable than the convergence of the CONV method.

To have a better insight into the convergence of both methods, the results for the CONV en COS method are compared based on the convergence rate. As can be seen in Figure 5.4 and Figure 5.5, the approximation approaches yield good results for the Kou jump-model. The convergence of the algorithm is numerically checked for both the CONV and the COS method. To compare the convergence, the maximum absolute error between the simulation results and the COS and CONV method is taken. The results are given in Figure 5.6. 5.6 a shows that the COS method does converge a bit faster to the simulation results. It is known of the COS method, that it has exponential convergence, which is clear from the figure [32]. Furthermore, both methods are not becoming more accurate when compared to the Monte Carlo results. However, note that the simulation results are also approximations and therefore contain an approximation error with respect to the true results. From Figure 5.6 b it can be seen that the COS and the CONV method also converge to each other, which together with the simulation results, shows that both the CONV and the COS method converge to the true price.

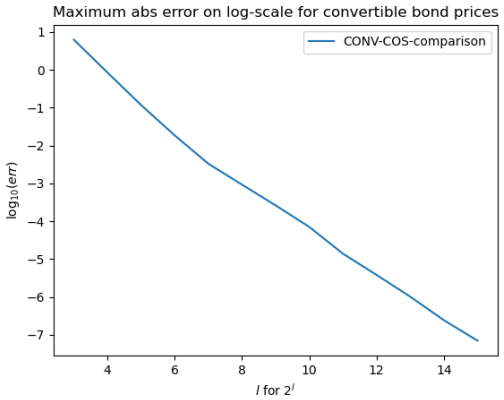
A similar comparison can be made for the callable convertible bond approximations. From Figure 5.7 a it can be seen that the COS method only has a slightly faster convergence for small d, where $N = 2^d$ and $d \in \mathbb{N}$. However, when $d > 6$, the CONV method shows much better convergence. In Figure 5.7 b, it can be seen that the COS method shows no stable convergence over L, which may be related to the unstable approximations of the early exercise point x_{call}.

5.4. Numerical results under stochastic interest rates

In this section, the model is extended with stochastic interest rates. Using the results of Chapter 3, the actual price can be approximated using simulation. Furthermore, also the CONV algorithm can be



(a) Error analysis of convertible bond prices obtained with the COS and CONV method and compared with the simulation results.



(b) Comparison between the COS and the CONV for convertible bond approximations.

Figure 5.6: A plot of the convergence of the CONV and COS method. Both the COS and the CONV method were used to compute values of a convertible bond for different values of F. The CONV method is based on a grid with $L_x = 10$. The COS method is performed with $\hat{L} = 10$. For all methods $\mathrm{M} = 51$ time steps are considered. The simulations are conducted under $\mathrm{N} = 500000$ paths. The error was computed by taking the maximum of the absolute difference between the Fourier methods and the simulation results respectively.

κ	η	f_r	ρ
0.05	0.001	0.05	0.3

Table 5.4: Parameters used for the Hull-White process. The parameters are chosen based on expert judgment.

used to obtain approximations, where the approach described in Chapter 4 is used. The stochastic interest rates will be computed under the Heath-Jarrow-Morton framework, which requires market data as input for the forward curve. In this thesis, an artificial forward curve is assumed which is given by

$$P_{\text{mkt}}(t,T) = \mathrm{e}^{-r_f(T-t)}, \tag{5.1}$$

where $r_f \in \mathbb{R}$ is a constant rate. The parameters for the Hull-White process are given in Table 5.4. The remaining parameters are set equal to the ones given in Section 5.3.

Under the assumption that the zero-coupon bonds follow the price given in Equation (5.1), the function P_{mkt} becomes smooth. The differentiability of P_{mkt} enables the use of numerical differentiation to obtain the forward rate function $f_r(0,\cdot)$, which is also smooth. Using the numerical differentiation techniques, the function θ from Equation (2.2) can be approximated. The corresponding approximations enable the simulation, numerical integration, and computations of the interest rate process $\{\tilde{r}(t)\}_{t\geq 0}$, $\theta(t)$ within the characteristics function and intermediate bond prices via Equation (2.7).

The grid size in the two-dimensional approach is taken similarly to the one-dimensional approximations.

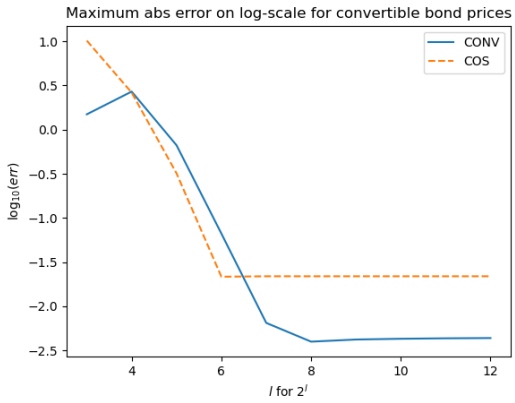
Remark 5.3 (Cumulant based integration grid size)**.** *Consider the two-dimensional grid given in Theorem 4.2. For both of the grids, an equal number of integration points is chosen. To obtain sufficiently accurate results for the Fourier approximations, the grid size for Δ_x and Δ_y is chosen such that*

$$\Delta_x = \frac{2L_x\sqrt{\zeta^r(T)}}{N},\ \Delta_y = \frac{2L_y\sqrt{\zeta_2(w,T)+\sqrt{\zeta_4(w,T)}}}{N},$$

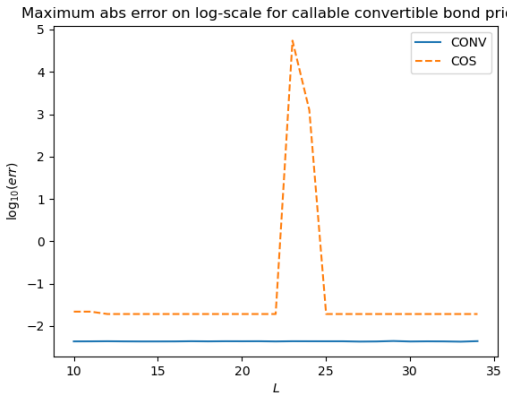
for $L_x, L_y \in \mathbb{R}_+$, N the number of grid points, ζ_r the variance of the interest rate process, ζ_2 and ζ_4 the second and fourth cumulant function of the individual process $\{X(t)\}_{t\geq 0}$ which is given by

$$\zeta_j(w,T) \equiv \zeta_j(w,T) = \frac{1}{\mathrm{i}^j}\frac{\partial^j \log(\phi_X(w,T))}{\partial w^j}\bigg|_{w=0},$$

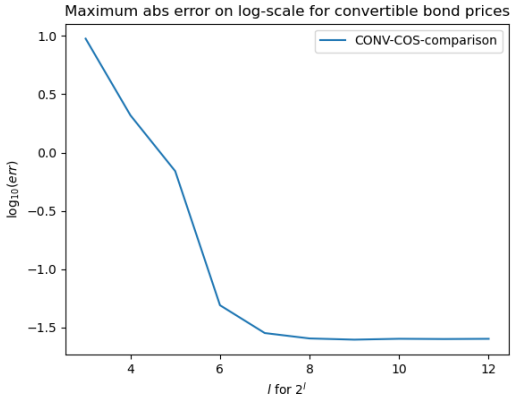
for $j \in \{2,4\}$ and the characteristics function $\phi_X(w,T)$ corresponding to the one-dimensional process $\{X(t)\}_{t\geq 0}$ under constant interest rates. The variance of the interest rate process $\{\tilde{r}(t)\}_{t\geq 0}$ is given



(a) Error analysis of convertible bond prices obtained with the COS and CONV method and compared with the simulation results for different grid sizes N.



(b) The convergence scheme of the COS and the CONV method for different values of $\hat{L}$ and L_x respectively. The computations were conducted using $N = 2^{13} = 8192$ grid points for both the COS and the CONV method.



(c) Error analysis of convertible bond prices obtained with the COS and compared with the CONV method for different grid sizes N.

Figure 5.7: A plot of the convergence of the CONV and COS method. Both the COS and the CONV method were used to compute values of a callable convertible bond for different values of F. The CONV method is based on a grid with $L_x = 10$. The COS method is performed with $\hat{L} = 10$. For all methods $\mathrm{M} = 51$ time steps are considered. The simulations are conducted under $\mathrm{N} = 500000$ paths. The error was computed by taking the maximum of the absolute difference between the Fourier methods and the simulation results respectively.

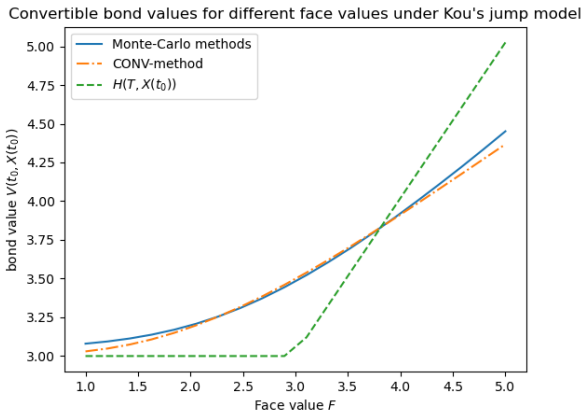


Figure 5.8: A plot of the convertible bond for different values of F. The CONV method is based on a grid with $L_x = 2$, $L_y = 8$ and $N = 2^{13} = 8192$ grid nodes. For the CONV $\mathrm{M}_{\text{conv}} = 20$ time steps were considered. The Monte Carlo method was conducted under $\mathrm{N} = 500000$ paths with each $\mathrm{M}_{\text{MC}} = 51$ nodes. The blue dashed line corresponds with the payoff function at maturity $H(T, \cdot)$ given in Equation (4.6), and therefore shows the value of the bond at immediate payoff.

by [32].

$$\zeta^r(t) = \frac{\eta^2}{2\kappa}(1 - \mathrm{e}^{-2\kappa t}).$$

As in the one-dimensional case, first, the convertible bond without the callable feature is considered. The results are given in Figure 5.8.

From Figure 5.8 it can be seen that, when $N = 2^{13}$ grid points were considered, the resulting approximations using the CONV method are still considerably off. To obtain better approximations, the number of grid points should be considered greater than $N = 2^{13}$. However, due to the interpolation step in the algorithm, higher N cannot be approximated as this will result in a memory error. The interpolation needs to be done on the surface with more than $2^{13} \times 2^{13}$ nodes, which becomes too large. Using different configurations for L_x and L_y yields similar results. For the callable convertible bond, similar behavior is observed. The results are given in Figure 5.9.

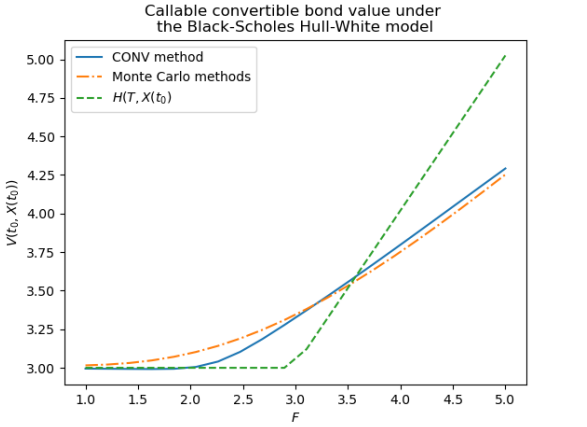


Figure 5.9: A plot of the callable convertible bond for different values of F. The CONV method is based on a grid with $L_x = 1$, $L_y = 8$ and $N = 2^{13} = 8192$ grid nodes. For the CONV method $\mathrm{M}_{\text{conv}} = 20$ time steps were considered. The Monte Carlo method was conducted under $\mathrm{N} = 500000$ paths of each $\mathrm{M}_{\text{MC}} = 51$ nodes. The blue dashed line corresponds with the payoff function at maturity $H(T, \cdot)$ given in Equation (4.6), and therefore shows the value of the bond at immediate payoff.

If the coupon rate is chosen equal to zero, such that the bond does not pay any coupon payments, the pricing problem of the convertible bond becomes one that is very similar to a European type of option. As discussed in Section 4.4.3, the bond value can then be obtained via the COS method in

two dimensions. Furthermore, the CONV method can also be used where the intermediate steps will disappear and the value of the bond can be directly obtained via Theorem 4.5 and Equation (4.5). The results are given in Figure 5.10.

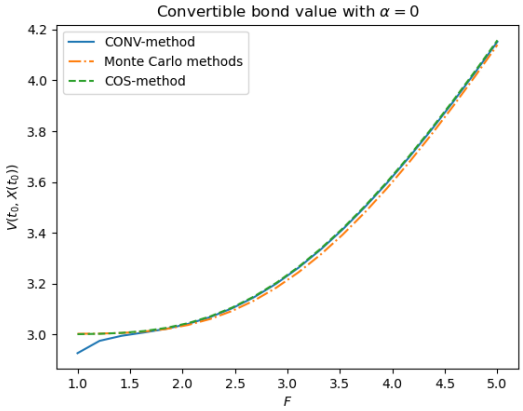
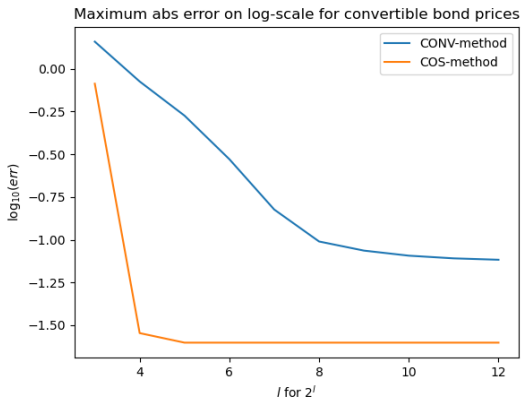
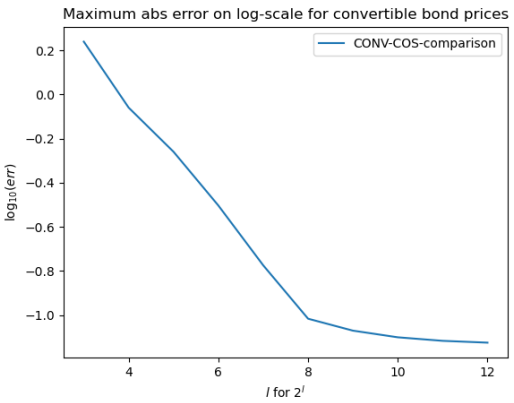


Figure 5.10: A plot of the value of a non-coupon paying convertible bond for different values of F. The CONV method is based on a grid with $L_x = 1$, $L_y = 8$ and $N = 2^{13} = 8192$ grid nodes. The COS method is applied with $N = 2^{13} = 8192$, $\hat{L}_1 = 10$ and $\hat{L}_2 = 10$. The Monte Carlo method was conducted under $\mathrm{N} = 500000$ paths, each with $\mathrm{M} = 51$ nodes.

The possibility to use the COS method allows for a more extensive convergence analysis. These results are given in Figure 5.11.



(a) Error analysis of convertible bond prices with $\alpha = 0$ obtained with the COS and CONV method and compared with the Monte Carlo results for different grid sizes N.



(b) Error analysis of convertible bond prices with $\alpha = 0$ obtained with the COS and compared with the CONV method for different grid sizes N.

Figure 5.11: A plot of the convergence of the CONV and COS method. Both the COS and the CONV method were used to compute values of a convertible bond that pays no coupon payments. The CONV method is based on a grid with $L_x = 1$ and $L_y = 9$. The COS method is performed with $\hat{L}_1 = \hat{L}_2 = 10$. For all methods $\mathrm{M} = 51$ time steps are considered. The simulations are conducted under $\mathrm{N} = 500000$ paths. The error was computed by taking the maximum of the absolute difference between the Fourier methods and the simulation results respectively.

From Figure 5.11 a it can be seen that the COS method converges much more rapidly than the CONV method. The rapid conversion was expected from the one-dimensional case. Contrary to the one-dimensional case, Figure 5.11 a shows a much slower convergence of the CONV method in the two-dimensional case. The slow conversion is consistent with the observation done in Figure 5.9 and explains the insufficient approximations. Figure 5.11 b seem to indicate that the COS and the CONV method are converging, however very slow due to the slow convergence of the CONV method. At the tail of Figure 5.11 b the convergence seems to stop. Yet, it is known that the CONV method does not

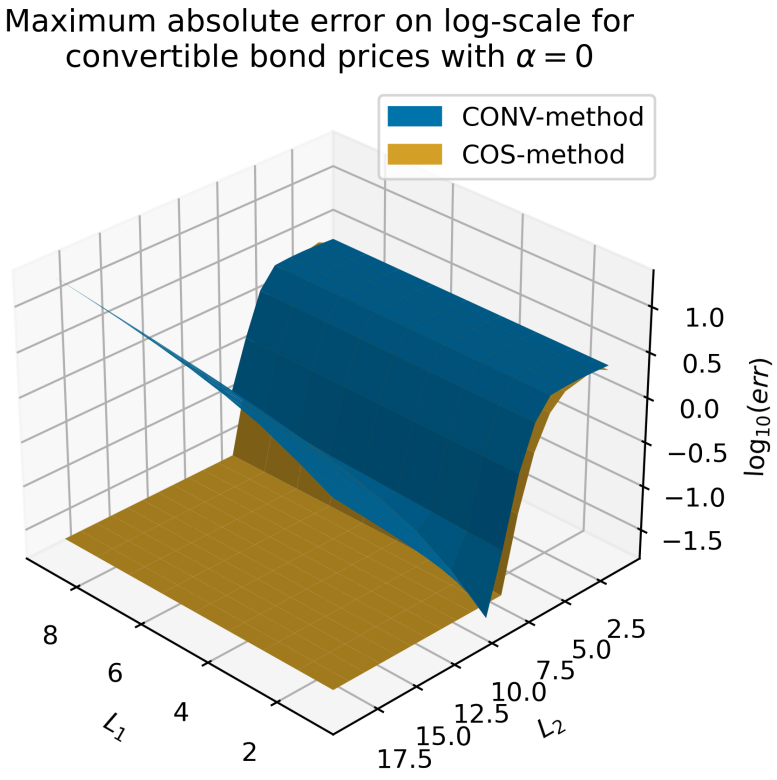


Figure 5.12: A plot of the maximum absolute error between the CONV and COS obtained values compared to the values obtained by Monte Carlo methods for a zero-coupon convertible bond. The values are computed for different $L_x, L_y, \hat{L}_1$ and $\hat{L}_2$ as given in Remark 5.3 and are given by L_1 and L_2 over the x- and y-axis. The amount of nodes is set to $N = 2^{10} = 1024$. The simulations are conducted under $\mathrm{N} = 500000$ paths, each with $\mathrm{N} = 51$ nodes.

necessarily yield exponential convergence, and hence does not necessarily yield a linear relationship between the log of the absolute error and the number of grid points N [14].

To ensure the correct chosen hyperparameter L_x and L_y, the convergence of the CONV method for the zero-coupon convertible bond is computed. The results are given in Figure 5.12. From Figure 5.12, the best hyper-parameters are indeed low values for L_x and L_y around 10.

6

Conclusion

6.1. Conclusion

In this thesis, a structural default model is used to value callable convertible bonds. The structural model in this thesis is proposed to follow a jump-diffusion process and is extended with stochastic interest rates. In particular, the jump models proposed by Kou and Merton are discussed. For the model including stochastic interest rates, the Heath, Jarrow, and Morton framework was introduced to enable the possibility to directly model the forward rates from quotes observed in the market. Within this thesis, artificial zero-coupon bond quotes were used to compute the parameters of the zero-coupon bonds. Though, given a sufficient smooth curve for the forward rate quotes, the model can also be used directly on market quotes.

This thesis discussed different valuation methods as proposed by Longstaff and Schwartz, Lord et al. and Oosterlee and Grzelak applied to the valuation of callable convertible bond. The proposed methods include Monte Carlo methods and Fourier methods. In particular, the Fourier techniques, which include a convolution method (CONV) and a cosine-based (COS) method, enable fast and efficient computation possibilities. Whereas the Monte Carlo methods and the CONV method are already used in different articles, the COS method was not yet applied to the callable convertible bond valuation [3, 1]. As the COS method has been proven to be an efficient algorithm for approximating the values of financial derivatives, this thesis uses the COS method to compare its convergence to seek more insights into the convergence of the CONV method. Monte Carlo methods are used to verify the obtained approximations.

For the problem considered under constant interest rates, this thesis shows that the COS method can be derived and applied. For the COS method to be applied to the callable convertible bond problem, exercise and continuation cosine coefficients need to be computed for intermediate steps. The computation of the cosine coefficients, to be able to approximate the bond value, entails some more complex steps and makes the total approximation method more error-prone. The cosine coefficients depend on the early exercise points, which are the boundary points between early exercise intervals and the continuation interval. The early exercise points are obtained using a numerical root-finding algorithm, which was chosen to be the Newton-Raphson method. Contrary to the conversion interval, the early exercise point corresponding to the call interval was observed to be sensitive to the initial value of the root-finding algorithm. The sensitivity will make the COS method less robust with respect to the CONV method.

For small values of the grid size, the COS method showed to converge much faster than the CONV method. For larger values of the grid size, the CONV method showed to catch up with the COS method to become almost equally accurate. Especially when the early exercise points are not considered and a convertible bond without a call option is proposed, the CONV and the COS method showed to converge very rapidly to a common price, which can be assumed to be the true price of the convertible bond. The results also showed that, contrary to the COS method, the CONV method was robust under the choice of the hyper-parameter concerning the integration grid. For the COS method, it was shown that

a bad choice of the hyper-parameter could lead to a bad approximation.

Including stochastic interest rates will make the valuation problem more complex. An algorithm for using the CONV method in two dimensions was proposed by Ballotta and Kyriakou to be able to compute the bond values efficiently [3]. Although the two-dimensional CONV method shows to converge to the value obtained under Monte Carlo simulation, the results are still off for grid sizes of intermediate size, e.g. $N = 2^{13}$. Results seem to indicate that the amount of grid points is not sufficient for the proposed integration interval and that therefore more grid points are needed. As the two-dimensional algorithm depends on the use of interpolation heavily, and the CONV method is based on the assumption that the grid points are a power of two, each higher order grid will result in an interpolation over four times the size of the previous grid surface. As a consequence, a much greater amount of computer memory is necessary to compute these interpolations and can therefore be a problem for computations under limited resources.

To conduct a more thorough error analysis in the two-dimensional case, the problem was reduced to the valuation of a zero-coupon convertible bond. This reduction yielded the possibility to omit the computation of the intermediate bond values and introduced a problem similar to a European-type option. From the results of the zero-coupon case, the COS method showed a more rapid convergence towards the approximations obtained by the Monte Carlo methods than the approximations obtained using the CONV method. Furthermore, when taking the hyper-parameters into account, the CONV method showed less robust features than the COS method. Only for a small range of L_x and L_y convergence is obtained for the CONV method. On the other hand, the COS method clearly shows convergence for a much wider range of hyper-parameters $\hat{L}_1$ and $\hat{L}_2$. Though the COS method shows a much better and more robust convergence, it should be mentioned that the one-dimensional case showed that for early exercise products the COS method shows less robustness which is not taken into account in this comparison.

6.2. Further research

Different aspects of this thesis yield possibilities for further research. First, the results of this thesis give rise to a calibration method for the calibration of structural models based on the COS method. As the COS method shows to be robust and efficient for non-early exercise bonds, more advanced structural default models such as the CGMYB and the NIGB jump-diffusion models can be calibrated with bond data. Whereas Fang et al. showed the application of the COS method to the calibration of structural default models by using credit default spreads, this thesis shows a way to also calibrate structural default models based on bond quotes. However, as only a small selection of companies have credit default swaps quoted, this may not be an ideal way to calibrate structural default models in general. A greater selection of companies finances their investments with debt. Therefore, there is a greater market for bond data. Using the COS method approach as presented in this thesis, further research can be done in light of the calibration of different structural Levy models. Although this thesis yields the tools to be able to calibrate structural default models based on bond quotes, the actual calibration algorithm was not presented and is therefore left for further research.

Secondly, as discussed in Section 6.1, the COS method applied to the valuation of callable convertible bonds does indicate some sensitivity with respect to the approximation of the early exercise point concerning the call option. In this thesis, the Newton-Raphson method was used to obtain early exercise points for constructing intervals used in the computations of the cosine coefficients. Though the Newton-Raphson method does yield sufficient results, its use affects the robustness of the algorithm due to the sensitivity to the initial guess necessary for the Newton-Raphson method. Further research is necessary to overcome the problem of sensitivity with respect to the initial guess. Possibly more robust root-finding algorithms can be used to obtain a more robust COS algorithm for the callable convertible bond.

Finally, as the COS method seem to have a better convergence, it may be interesting to derive the two-dimensional COS algorithm for valuing callable convertible bonds. Ruijter and Oosterlee has shown a two-dimensional algorithm that can be used to obtain continuation coefficients for the Heston model. A similar approach can be used to obtain coefficients for the Black Scholes Hull-White model. The two-dimensional COS approach will enable a way to use the COS method in approximating (callable)

convertible bond values under the Black Scholes Hull-White model. Being able to do efficient computations under the Black-Scholes Hull-White model, enables the application of efficient computations to the class of early-exercise hybrid products.

A

Concepts

A.1. Economic assumptions

In general, three different categories of assumptions concerning the general economy can be distinguished, namely behavioral, market, and model assumptions [21]. Behavioral assumptions describe how the participants within a market will act on changes within the economy. On the other hand, market assumptions prescribe the market conditions and the characteristics of the securities which are traded on the market. Finally, model assumptions concern all assumptions that apply directly to the modeling of the assets within the market [21]. Different parties on the contract of a contingent claim may have different interests. This discrepancy in interest will result in different behavior concerning the possibilities within the contract. A fair price for the claim may therefore be difficult to derive. To exclude any ambiguity in the behavior of different parties concerning a contract, this section will describe the behavior of the market and its participants. Particularly, this section will conclude with the assumption regarding the capital structure of companies participating in the market. The capital structure is an important assumption within structural default models, as it describes how debt and assets are related and how the capital structure may change upon conversion. Furthermore, specific model assumptions and the dynamics of the underlying assets, are discussed in Chapter 2 and will therefore not be repeated here.

Suppose an economy that includes a market on which securities can be traded. Furthermore, assume investors who trade in that market and suppose there exist companies that can participate in this market to be able to finance investments.

Assumption A.1 (Rational investors)**.** *The economy consists only of rational investors. A rational investor will always try to enhance her wealth.*

One of the main assumptions within derivative valuation is the no-arbitrage argument. The no-arbitrage argument will prevent the existence of the so-called free lunch, i.e. there are no risk-free profits for the considered stochastic quantities [32]. A more detailed discussion on arbitrages is given by, for instance, Shreve [35].

Assumption A.2 (No arbitrage)**.** *The market does not allow for an arbitrage opportunity.*

As the investor's main concern will be enhancing her wealth, the company will be interested in the wealth of its shareholders. Here, the assumption from Ingersoll Jr is followed:

Assumption A.3 (Rational management)**.** *The company is managed rationally, that is, the management of the company will always pursue to enhance the wealth of its common shareholders within its possibilities.*

Prices of claims can be dependent on the costs within the economy in a complicated way. Therefore, a perfect market is assumed.

Assumption A.4 (Perfect markets)**.** *The market within the economy is perfect, i.e. there are no taxes, and transaction costs and all investors have equal information [21].*

When an investor takes a long position in a security, she buys the security. Buying the security will yield the value of the security for which she has bought the security as equity in her portfolio. On the other hand, when an investor takes a short position in a security, she borrows the asset and immediately sells it. Selling the asset will yield the value of the asset for which it was borrowed as debt within her portfolio.

Assumption A.5 (Frictionless markets)**.** *Within the market, there are no restrictions on borrowing money and taking so-called short positions.*

The securities and claims on the securities within the market will be modeled in a continuous-time model.

Assumption A.6 (Continuous time trading)**.** *Within the market, securities can be traded in continuous time.*

Finally, there is also an assumption made on the capital structure of the company. As the company value may depend on the amount of leverage it uses to finance its investments (for instance due to taxes), the value of the company may not entirely be determined by the liabilities and assets alone. To overcome this problem, the assumption better known as the Modiagli-Miller theorem will be assumed.

Assumption A.7 (Capital structure)**.** *It is assumed that the Modiagli-Miller proposition 1 holds, that is, the value of the company does not depend on its capital structure [21, 31, 29].*

Assumption A.7 will assure that due the limited liability of corporations, the company value $A(t)$ at time t can be decomposed in,

$$A(t) = V^{\text{debt}}(t) + V^{\text{assets}}(t), \tag{A.1}$$

where $V^{\text{debt}}(t)$ is the market value of the debt of the company at time t and $V^{\text{assets}}(t)$ is market value of the equity of the company at time t [21]. Equation (A.1) states that the value of the company can be decomposed into the market value of the debt and the market value of the assets of the company (for example its shares) [21]. Hence, it is assumed that the company is financed out of debt (for instance bonds) and equity (for instance shares).

B

Derivations

B.1. Kou's density function

In this section, it will be derived that the density function proposed by Kou and Wang as density for the jump size

$$f_J(x) = p_1\alpha_1 \mathrm{e}^{-\alpha_1 x}\mathbb{1}_{\{x\geq 0\}} + p_2\alpha_2\mathrm{e}^{\alpha_2 x}\mathbb{1}_{\{x<0\}},$$

such that $p_1+p_2 = 1,\ p_1 > 0, p_2 > 0,\ \alpha_1 > 1$ and $\alpha_2 > 0$, is indeed a density function [23]. First note that f_J is a positive function, as p_1, p_2, α_1 and α_2 are strictly positive numbers. Moreover, $g(x) = \mathrm{e}^{ax}, a \in \mathbb{R}$ is a positive function and the sum over positive functions is also a positive function. As the indicator function, $\mathbb{1}_{\{A\}}(x)$ over a set A is also a positive function and the multiplication over two positive functions is also positive, it follows that f_J is a positive function. Next, consider the integral of f_J over the support;

$$\int_{-\infty}^{\infty} f_J(x)\mathsf{d}x = \int_{-\infty}^{0} p_2\alpha_2\mathrm{e}^{\alpha_2 x}\mathsf{d}x + \int_{0}^{\infty} p_1\alpha_1\mathrm{e}^{-\alpha_1 x}\mathsf{d}x. \tag{B.1}$$

Using that

$$\frac{\mathsf{d}g}{\mathsf{d}x}(x) = a\mathrm{e}^{ax},$$

for $g(x) = \mathrm{e}^{ax}$, it follows that Equation (B.1) can be rewritten as

$$-p_1\left[\mathrm{e}^{-\alpha_1 x}\right]_{x=0}^{x=\infty} + p_2\left[\mathrm{e}^{\alpha_2 x}\right]_{x=-\infty}^{x=0} = p_1 + p_2 = 1,$$

which concludes the argument and therefore f_J is a density.

B.1.1. The cumulative distribution function

By definition, the cumulative distribution function is given by

$$F_J(x) = \int_{-\infty}^{x} f_J(s)\mathsf{d}s. \tag{B.2}$$

To compute the integral given in Equation (B.2), the density function f_J is rewritten in two cases;

$$f_J(x) = \begin{cases} p_1\alpha_1\mathrm{e}^{-\alpha_1 x} & \text{if} \quad x \geq 0 \\ p_2\alpha_2\mathrm{e}^{\alpha_2 x} & \text{if} \quad x < 0 \end{cases},$$

where $p_1 + p_2 = 1$. First, the case $x < 0$ is integrated. Integrating $f_J(x)$ for $x < 0$ yields,

$$\int_{-\infty}^{x} f_J(s)\mathsf{d}s = \int_{-\infty}^{x} p_2\alpha_2\mathrm{e}^{\alpha_2 s}\mathsf{d}s = p_2\left[\mathrm{e}^{\alpha_2 s}\right]_{s=-\infty}^{s=x} = p_2\mathrm{e}^{\alpha_2 x}.$$

On the other hand if $x \geq 0$

$$\begin{aligned}\int_{-\infty}^{x} f_J(s)\mathsf{d}s &= \int_{-\infty}^{0} f_J(s)\mathsf{d}s + \int_{0}^{x} f_J(s)\mathsf{d}s \\ &= p_2 + \int_0^x p_1\alpha_1 \mathsf{e}^{-\alpha_1 s}\mathsf{d}s \\ &= 1 - p_1 + p_1\left[-\mathsf{e}^{\alpha_1 s}\right]_{s=0}^{s=x} \\ &= 1 - p_1\mathsf{e}^{\alpha_1 x},\end{aligned}$$

where it was used that $p_1 + p_2 = 1$. This yields the following CDF;

$$F_J(x) = \begin{cases} 1 - p_1\mathsf{e}^{-\alpha_1 x} & \text{if} \quad x \geq 0, \\ p_2\mathsf{e}^{\alpha_2 x} & \text{if} \quad x < 0. \end{cases} \tag{B.3}$$

B.1.2. Simulating from Kou's density

There are different ways to simulate from a known density function, e.g. by rejection sampling or directly from a known CDF. In this section, the CDF $F_J(x)$ in Equation (B.3) will be used to sample realizations from the random variable J. To be able to sample from $F_J(x)$, the following known result will be used (see for example [28]). Denote the uniform distribution over the interval $[0, 1]$ by unif $(0, 1)$, such that the CDF of $U \sim$ unif $(0, 1)$ is given by

$$F_U(u) = \mathbb{Q}(U \leq u) = u.$$

Let X be a continuous random variable and let the general the inverse of the CDF of X be given as

$$F_X^{-1}(u) = \inf\{x : F_X(x) \geq u\}.$$

Write $X = F_X^{-1}(U)$, the general inverse of the CDF of X over the continuous uniform random variable U, then for the random variable X holds that (see [28])

$$\mathbb{Q}\left(X \leq x\right) = \mathbb{Q}\left(F_X^{-1}(U) \leq x\right) = \mathbb{Q}\left(U \leq F_X(x)\right) = F_X(x).$$

As a consequence, realizations of X can be sampled using the uniform distribution. To sample from the distribution of X, the following algorithm can be used;

1. Sample n_r realizations $u_0, u_1, \ldots, u_{n_r-1}$ of the uniform distribution U.
2. calculate $x_i = F^{-1}(u_i)$ for $i \in 0, \ldots, n_r - 1$.

The computed value of x_i is a realization of X.

It is therefore left to find the general inverse of F_J. Write $F_J(x) = y$, and distinguish between the cases $x < 0$ and $x \geq 0$. First, let $x < 0$, then $F_J(x) = p_2\mathsf{e}^{\alpha_2 x}$ and therefore

$$y = p_2\mathsf{e}^{\alpha_2 x} \iff \frac{1}{\alpha_2}\log\left(\frac{y}{p_2}\right).$$

As $x < 0$, it has to hold that

$$\frac{1}{\alpha_2}\log\left(\frac{y}{p_2}\right) < 0 \iff \frac{y}{p_2} < 1 \iff y < p_2 = 1 - p_1.$$

On the other hand, if $x \geq 0$,

$$y = 1 - p_1\mathsf{e}^{-\alpha_1 x} \iff \frac{1-y}{p_1} = \mathsf{e}^{\alpha_1 x} \iff \frac{1}{\alpha_1}\log\left(\frac{p_1}{1-y}\right),$$

and therefore

$$\frac{1}{\alpha_1}\log\left(\frac{p_1}{1-y}\right) \geq 0 \iff \frac{p_1}{1-y} \geq 1 \iff p_1 \geq 1 - y \iff y \geq 1 - p_1.$$

As a consequence, the generalized inverse of F_J is given by

$$F_J^{-1}(y) = \begin{cases} \frac{1}{\alpha_1} \log\left(\frac{p_1}{1-y}\right) & \text{if} \quad y \geq p_2, \\ \\ \frac{1}{\alpha_2} \log\left(\frac{y}{p_2}\right) & \text{if} \quad y < p_2, \end{cases}$$

with $p_2 = 1 - p_1$.

B.2. Log transform of the jump diffusion process

In this section, the derivation of the log-transform of the jump-diffusion process $\{A(t)\}_{t\geq 0}$ will be given. To find the log-transform, Ito's lemma will be applied to find the dynamics of the log-transformed process $\{X(t)\}_{t\geq 0}$. First, Ito's lemma for jump-diffusion processes will be given. Here, only the lemma will be stated. Proofs can be found in the rich literature about the subject, see for instance [35]. Afterward, the specific derivation for the log-transform of the process $\{A(t)\}_{t\geq 0}$ will be derived.

Theorem B.1 (Ito's lemma for jump-diffusion processes)**.** *Assume that $\mathcal{P}(t)$ follows a Poisson process with intensity parameter λ. Let $A(t)$ follow a jump diffusion process given as*

$$\mathbf{d}A(t) = \mu(t, A(t))\mathbf{d}t + \sigma(t, A(t))\mathbf{d}W_A(t) + g(t, A(t))\mathbf{d}\mathcal{P}(t),$$

where $\mu(t, A(t))$, $\sigma(t, A(t))$ and $g(t, A(t))$ are adapted functions and $W_A(t)$ is a Brownian motion under the measure $\mathbb{Q}$ [32]. Let $f \equiv f(t, A(t))$ be a sufficiently differentiable function. Then for $A \equiv A(t)$ it holds that

$$\mathbf{d}f = \left[\frac{\partial f}{\partial t} + \mu(t, A(t))\frac{\partial f}{\partial A} - \frac{1}{2}\sigma^2(t, A(t))\frac{\partial^2 f}{\partial A^2}\right]\mathbf{d}t + \sigma(t, A(t))\frac{\partial f}{\partial A}\mathbf{d}W_A(t) + \left[f(g(t, A_-) + A_-) - f(A_-)\right]\mathbf{d}\mathcal{P}(t), \quad \text{(B.4)}$$

where A_- is the left limit of the process $A(t)$ at time t [32]. Let $A^C(t)$ denote the continuous part of the process $A(t)$, such that

$$\mathbf{d}A^C(t) = \mu(t, A(t))\mathbf{d}t + \sigma(t, A(t))\mathbf{d}W_A(t).$$

Then Equation (B.4) *can be rewritten to [35]*

$$\mathbf{d}f = \frac{\partial f}{\partial t}\mathbf{d}t + \frac{\partial f}{\partial A}\mathbf{d}A^C + \frac{1}{2}\frac{\partial^2 f}{\partial A^2}(\mathbf{d}A^C)^2 + \left[f(g(t, A_-) + A_-) - f(A_-)\right]\mathbf{d}\mathcal{P}(t),$$

where the last term concerns the discontinuities.

An intuitive explanation of Theorem B.1 is given by Oosterlee and Grzelak [32]. Using Theorem B.1, the log-transform of the process $\{A(t)\}_{t\geq 0}$ will be derived. Consider the jump-diffusion process as given in Equation (2.1), i.e.

$$\frac{\mathrm{d}A(t)}{A(t)} = \left(r(t) - \lambda\mathbb{E}\left[\mathrm{e}^J - 1\right]\right)\mathrm{d}t + \sigma\mathrm{d}W_A(t) + \left(\mathrm{e}^J - 1\right)\mathrm{d}\mathcal{P}(t), \quad \text{(B.5)}$$

where $\{\mathcal{P}(t)\}_{t\geq 0}$ is a Poisson process with intensity parameter λ, J is a random variable representing the jump size and $\{W_A(t)\}_{t\geq 0}$ is a Brownian motion under the measure $\mathbb{Q}$. Assume that $\{W_A(t)\}_{t\geq 0}$, J and $\{\mathcal{P}(t)\}_{t\geq 0}$ are mutually independent. Note that $\{r(t)\}_{t\geq 0}$ is also modelled as a stochastic process, with Brownian motion $\{W_r(t)\}_{t\geq 0}$. $\{W_A(t)\}_{t\geq 0}$ may therefore be a dependent Brownian motion that can be decomposed in

$$W_A(t) = \rho\bar{W}_r(t) + \sqrt{1-\rho^2}\bar{W}_A(t),$$

with $\left\{\bar{W}_r(t)\right\}_{t\geq 0}$ and $\left\{\bar{W}_A(t)\right\}_{t\geq 0}$ two independent Brownian motions under the measure $\mathbb{Q}$ (see Section 2.2). Finally, recall the continuous part of Equation (B.5), which is defined as

$$\frac{\mathrm{d}A^C(t)}{\mathrm{d}A(t)} = \left(r(t) - \lambda\mathbb{E}\left[\mathrm{e}^J - 1\right]\right)\mathrm{d}t + \sigma\mathrm{d}W_A(t),$$

as discussed in Section 2.2.

Let $f(t, A(t)) = X(t) = \log(A(t))$, then following Ito's lemma for a jump-diffusion processes yields

$$\mathsf{d}f = \left[\frac{\partial f}{\partial t} + \mu(t, A(t))\frac{\partial f}{\partial A} - \frac{1}{2}\sigma^2(t, A(t))\frac{\partial^2 f}{\partial A^2}\right]\mathsf{d}t + \sigma(t, A(t))\frac{\partial f}{\partial A}\mathsf{d}W_A(t) + [f(g(t, A_-) + A_-) - f(A_-)]\,\mathsf{d}\mathcal{P}(t),$$

where $g(t, A(t)) = \mathrm{e}^J - 1$. Note that

$$\frac{\partial f}{\partial t} = 0 \text{ and } \frac{\partial^2 f}{\partial A^2} = 0,$$

and hence for $A \equiv A(t)$ holds that,

$$\mathsf{d}f(t, A(t)) = \frac{\partial f}{\partial A}\mathsf{d}A^C + \frac{1}{2}\frac{\partial^2 f}{\partial A^2}(\mathsf{d}A^C)^2 + [f(g(t, A_-) + A_-) - f(A_-)]\,\mathsf{d}\mathcal{P}(t).$$

Computing the function f in the discontinuity of the process $\{A(t)\}_{t\geq 0}$ yields,

$$f(g(t, A_-) + A_-) - f(A_-) = \log\left(A_-\left(\mathrm{e}^J - 1\right) + A_-\right) - \log(A_-) = J.$$

Therefore the system $\mathsf{d}f(t, A(t)) = \mathsf{d}X(t)$ can be written as

$$\mathsf{d}X(t) = \left(r(t) - \lambda\mathbb{E}\left[\mathrm{e}^J - 1\right] - \frac{1}{2}\sigma^2\right)\mathsf{d}t + \sigma\mathsf{d}W_A(t) + J\mathsf{d}\mathcal{P}(t),$$

C

Additional Lemmas

C.1. Functions χ and ψ

Lemma C.1. *Let the coefficients functions χ_k and ϕ_k be given as*

$$\begin{cases} \chi_k(x_1, x_2) &= \int_{x_1}^{x_2} \mathrm{e}^y \cos\left(k\pi \frac{y-a}{b-a}\right) \mathrm{d}y, \\ \psi_k(x_1, x_2) &= \int_{x_1}^{x_2} \mathrm{e}^y \cos\left(k\pi \frac{y-a}{b-a}\right) \mathrm{d}y, \end{cases}$$

then both functions are known in closed form and given by

$$\begin{cases} \chi_k(x_1, x_2) &= \dfrac{1}{1+\left(\dfrac{k\pi}{b-a}\right)^2} \left[\begin{array}{r} \cos\left(k\pi \dfrac{d-a}{b-a}\right) \mathrm{e}^d - \cos\left(k\pi \dfrac{c-a}{b-a}\right) \mathrm{e}^c \\ + \dfrac{k\pi}{b-a} \sin\left(k\pi \dfrac{d-a}{b-a}\right) \mathrm{e}^d + \dfrac{k\pi}{b-a} \sin\left(k\pi \dfrac{c-a}{b-a}\right) \mathrm{e}^c \end{array} \right] \\ \psi_k(x_1, x_2) &= \begin{cases} \dfrac{b-a}{k\pi} \left[\sin\left(k\pi \dfrac{d-a}{b-a}\right) - \sin\left(k\pi \dfrac{c-a}{b-a}\right) \right] & k \neq 0, \\ d-c & k = 0. \end{cases} \end{cases}$$

The validity of Lemma C.1 follows directly from a result given by Oosterlee and Grzelak [32].

D

Proofs of theorems

D.1. Proof of Theorem 2.3

The proof of Theorem 2.3 includes two steps. First, the model is placed inside the class of affine diffusion models. Secondly, the characteristic function is derived.

Proof of Theorem 2.3. Recall the system of stochastic differential equations defined by the Black-Scholes Hull-White model,

$$\begin{bmatrix} \mathsf{d}r(t) \\ \mathsf{d}X^C(t) \end{bmatrix} = \begin{bmatrix} \kappa(\theta(t) - r(t)) \\ r(t) - \frac{1}{2}\sigma^2 \end{bmatrix} \mathsf{d}t + \begin{bmatrix} \eta & 0 \\ \sigma\rho & \sigma\sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} \mathsf{d}\bar{W}_r^{\mathbb{Q}}(t) \\ \mathsf{d}\bar{W}_A^{\mathbb{Q}}(t) \end{bmatrix}, \tag{D.1}$$

where $\left\{\bar{W}_A^{\mathbb{Q}}(t)\right\}_{t\geq 0}$ and $\{\bar{W}_r^{\mathbb{Q}}(t)\}_{t\geq 0}$ are two independent Brownian motions under the equivalent martingale measure $\mathbb{Q}$, $\kappa, \eta, \sigma, \rho$ are the speed of mean reversion, interest diffusion, volatility and correlation parameter respectively, and where θ is the mean reversion function given in Equation (2.2). The system in Equation (D.1) can be written in the form,

$$\mathsf{d}\boldsymbol{X} = \boldsymbol{\mu}(t, \boldsymbol{X}(t))\mathsf{d}t + \boldsymbol{\sigma}(t, \boldsymbol{X}(t))\mathsf{d}\bar{\boldsymbol{W}}(t),$$

where

$$\boldsymbol{X} = [r(t), X^C(t)]^{\mathsf{T}},\ \boldsymbol{\mu}(t, \boldsymbol{X}(t)) = \begin{bmatrix} \kappa(\theta(t) - r(t)) \\ r(t) - \frac{1}{2}\sigma^2 \end{bmatrix},$$
$$\boldsymbol{\sigma}(t, \boldsymbol{X}(t)) = \begin{bmatrix} \eta & 0 \\ \sigma\rho & \sigma\sqrt{1-\rho^2} \end{bmatrix},\ \bar{\boldsymbol{W}}(t) = \begin{bmatrix} \mathsf{d}\bar{W}_r^{\mathbb{Q}}(t) \\ \mathsf{d}\bar{W}_A^{\mathbb{Q}}(t) \end{bmatrix},$$

and such that the affinity conditions are given by (see Theorem 2.1)

$$\begin{aligned} \boldsymbol{\mu}(t, \boldsymbol{X}(t)) &= a_0 + a_1\boldsymbol{X}(t), \\ \boldsymbol{r}(t, \boldsymbol{X}(t)) &= r_0 + r_1^T\boldsymbol{X}(t), \\ \left(\boldsymbol{\sigma}(t, \boldsymbol{X}(t))\boldsymbol{\sigma}(t, \boldsymbol{X}(t))^T\right)_{ij} &= (c_0)_{ij} + (c_1)_{ij}^T\boldsymbol{X}_j(t), \end{aligned}$$

for $r_0 \in \mathbb{R}$, $r_1, a_0 \in \mathbb{R}^n$, $a_1, c_0 \in \mathbb{R}^{n\times n}$ and $c_1 \in \mathbb{R}^{n\times n\times n}$ (see Section 2.3). Furthermore, $\boldsymbol{r}$ is the interest rate process with respect to the discounting within the system. Under the BSHW model, the discounting is done by the Hull-White process and can be written in the form

$$\boldsymbol{r}(t, \boldsymbol{X}(t)) = r_0 + r_1^T\boldsymbol{X}(t) = [1\ 0]^T \begin{bmatrix} r(t) \\ X(t) \end{bmatrix}.$$

Hence $r_0 = 0$ and $r_1 = [1\ 0]^T$. Furthermore, note that $\boldsymbol{\sigma}(t, \boldsymbol{X}(t))$ does not contain any factors in $\boldsymbol{X}$ and hence, $c_1 = 0$. Moreover, for $\boldsymbol{\sigma} \equiv \boldsymbol{\sigma}(t, \boldsymbol{X}(t))$,

$$c_0 = \boldsymbol{\sigma}\boldsymbol{\sigma}^T = \begin{bmatrix} \eta & 0 \\ \sigma\rho & \sigma\sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} \eta & \sigma\rho \\ 0 & \sigma\sqrt{1-\rho^2} \end{bmatrix} = \begin{bmatrix} \eta^2 & \eta\sigma\rho \\ \eta\sigma\rho & \sigma^2\rho^2 + \sigma^2(1-\rho^2) \end{bmatrix} = \begin{bmatrix} \eta^2 & \eta\sigma\rho \\ \eta\sigma\rho & \sigma^2 \end{bmatrix}.$$

The affinity condition for the drift term can be written as

$$\boldsymbol{\mu}(t, \boldsymbol{X}(t)) = \begin{bmatrix} \kappa(\theta(t) - r(t)) \\ r(t) - \frac{1}{2}\sigma^2 \end{bmatrix} = \begin{bmatrix} \kappa\theta(t) \\ -\frac{1}{2}\sigma^2 \end{bmatrix} + \begin{bmatrix} -\kappa & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} r(t) \\ X(t) \end{bmatrix} = a_0 + a_1 \boldsymbol{X}(t).$$

It can therefore be concluded that the system defined in Equation (D.1)is indeed affine. Moreover, the affine coefficients are given by

$$\begin{aligned} a_0 &= \begin{bmatrix} \kappa\theta(t) \\ -\frac{1}{2}\sigma^2 \end{bmatrix}, & a_1 &= \begin{bmatrix} -\kappa & 0 \\ 1 & 0 \end{bmatrix}, \\ r_0 &= 0, & r_1 &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \\ c_0 &= \begin{bmatrix} \eta^2 & \eta\sigma\rho \\ \eta\sigma\rho & \sigma^2 \end{bmatrix}, & c_1 &= 0. \end{aligned} \tag{D.2}$$

The discounted characteristic function under the measure $\mathbb{Q}$ can be derived by solving the complex ordinary differential equations. This concludes the proof of the inclusion of the model within the affine class of diffusion models.

The proof continues with the derivation of the discounted characteristic function. Let ϕ denote the characteristic function of the system, where ϕ is given by

$$\phi(u, t, T) = \mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{-\int_t^T r(s)\mathrm{d}s + \mathrm{i}u^T \boldsymbol{X}(T)} \;\middle|\; \mathcal{F}_t \right],$$

for $u = [u_1\ u_2]^T, \boldsymbol{X}(t) \in \mathbb{R}^2$ and $\boldsymbol{X}(t) = [r(t)\ X(t)]$. As shown in [32], the discounted characteristic can be written in closed form, such that it is given by

$$\phi(u, t, T) = \mathrm{e}^{A_1(u_1,u_2,\tau) + A_2(u_1,u_2,\tau)r(t) + A_3(u_1,u_2,\tau)X(t)},$$

where $\tau = T - t$ and the functions $A = [A_2\ A_3]^T$ and A_1 are given by the complex ordinary differential equations (ODEs),

$$\begin{aligned} \frac{\mathrm{d}A_1}{\mathrm{d}\tau} &= -r_0 + A^T a_0 + \tfrac{1}{2} A^T c_0 A, \\ \frac{\mathrm{d}A}{\mathrm{d}\tau} &= -r_1 + a_1^T A + \tfrac{1}{2} A^T c_1 A, \end{aligned}$$

such that the initial conditions are given by $A_1(u_1, u_2, 0) = 0$, $A_2(u_1, u_2, 0) = \mathrm{i}u_1$ and $A_3(u_1, u_2, 0) = \mathrm{i}u_2$. The affine parameters (r_j, a_j and c_j for $j \in \{1, 2\}$) necessary for these derivations are known and given in Equation (D.2). First, the latter multi-dimensional ODE will be computed. Denote $A_2 \equiv A_2(u_1, u_2, \tau)$ and $A_3 \equiv A_3(u_1, u_2, \tau)$, and write the system of ODEs in matrix notation,

$$\begin{bmatrix} \frac{\mathrm{d}A_2}{\mathrm{d}\tau} \\ \\ \frac{\mathrm{d}A_3}{\mathrm{d}\tau} \end{bmatrix} = -\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -\kappa & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_2 \\ A_3 \end{bmatrix},$$

and recall that $c_1 = 0$ and therefore the last term vanishes. The ODE concerning A_3 is therefore given by

$$\frac{\mathrm{d}A_3}{\mathrm{d}\tau} = 0 \implies A_3 = C,$$

for a constant $C \in \mathbb{R}$. Using the initial condition yields,

$$A_3(u_1, u_2, \tau) = \mathrm{i}u_2.$$

Using this result, the ODE concerning A_2 is given by

$$\frac{\mathrm{d}A_2}{\mathrm{d}\tau} = -1 + -\kappa A_2 + A_3 = -1 - \kappa A_2 + \mathrm{i}u_2 \iff \frac{\mathrm{d}A_2}{\mathrm{d}\tau} + \kappa A_2 = -1 + \mathrm{i}u_2.$$

Multiply both sides by a test function φ and recognize that

$$\frac{\mathsf{d}\varphi A_2}{\mathsf{d}\tau} = A_2\frac{\mathsf{d}\varphi}{\mathsf{d}\tau} + \varphi\frac{\mathsf{d}A_2}{\mathsf{d}\tau}.$$

As a consequence, it has to hold that

$$\frac{\mathsf{d}\varphi}{\mathsf{d}\tau}A_2 = \kappa\varphi A_2 \iff \int\frac{1}{\varphi}\mathsf{d}\varphi = \int\kappa\mathsf{d}\tau \iff \log(\varphi) = \kappa\tau + \bar{C} \iff \varphi = C_0\mathrm{e}^{\kappa\tau},$$

for $C_0 = \mathrm{e}^{\bar{C}}$ and $\bar{C} \in \mathbb{R}$ a constant. Hence, the original ODE can be written as

$$\varphi\frac{\mathsf{d}A_2}{\mathsf{d}\tau} + \varphi\kappa A_2 = \varphi(-1+\mathrm{i}u_2) \iff \frac{\mathsf{d}\varphi A_2}{\mathsf{d}\tau} = \varphi(-1+\mathrm{i}u_2).$$

Integrating both sides, yield,

$$\begin{aligned}
\varphi A_2 &= \int C_0 e^{\kappa\tau}(-1+\mathrm{i}u_2)\mathsf{d}\tau \\
&= C_0(-1+\mathrm{i}u_2)\int \mathrm{e}^{\kappa\tau}\mathsf{d}\tau \\
&= (-1+\mathrm{i}u_2)\left[\frac{1}{\kappa}C_0\mathrm{e}^{\kappa\tau} + C_0C_1\right] \\
\iff \quad C_0e^{\kappa\tau}A_2 &= (-1+\mathrm{i}u_2)\left[\frac{1}{\kappa}C_0\mathrm{e}^{\kappa\tau} + C_0C_1\right] \\
\iff \quad A_2 &= \frac{(-1+\mathrm{i}u_2)}{\kappa} + C_2\mathrm{e}^{-\kappa\tau},
\end{aligned}$$

where $C_2 = (-1+\mathrm{i}u_2)C_1$. Using the initial condition $A_2(u_1,u_2,0) = \mathrm{i}u_1$, yields

$$\begin{aligned}
\mathrm{i}u_1 &= \frac{(-1+\mathrm{i}u_2)}{\kappa} + C_2 \\
\iff \quad C_2 &= \mathrm{i}u_1 + \frac{(1-\mathrm{i}u_2)}{\kappa}.
\end{aligned}$$

Hence, the function A_2 is given by

$$A_2(u_1,u_2,\tau) = \left[\mathrm{i}u_1 + \frac{(1-\mathrm{i}u_2)}{\kappa}\right]\mathrm{e}^{-\kappa\tau} + \frac{(\mathrm{i}u_2-1)}{\kappa} = \mathrm{i}u_1\mathrm{e}^{-\kappa\tau} + \frac{(\mathrm{i}u_2-1)}{\kappa}(1-\mathrm{e}^{-\kappa\tau}).$$

It remains to solve the ODE corresponding to the function A_1,

$$\begin{aligned}
\frac{\mathsf{d}A_1}{\mathsf{d}\tau} &= -r_0 + A^Ta_0 + \frac{1}{2}A^Tc_0A \\
&= [A_2\ A_3]\begin{bmatrix}\kappa\theta(t)\\ -\frac{1}{2}\sigma^2\end{bmatrix} + \frac{1}{2}[A_2\ A_3]\begin{bmatrix}\eta^2 & \eta\sigma\rho\\ \eta\sigma\rho & \sigma^2\end{bmatrix}\begin{bmatrix}A_2\\ A_3\end{bmatrix} \\
&= \kappa\theta(t)A_2 - \frac{1}{2}\sigma^2A_3 + \frac{1}{2}\left[\eta^2A_2^2 + 2\eta\sigma\rho A_2A_3 + \sigma^2A_3^2\right],
\end{aligned}$$

which will be written in the convenient form;

$$\frac{\mathsf{d}A_1}{\mathsf{d}\tau} = \kappa\theta(t)A_2 + \frac{1}{2}\sigma^2(A_3^2 - A_3) + \frac{1}{2}\eta^2A_2^2 + \eta\sigma\rho A_2A_3, \tag{D.3}$$

and where θ is given in Equation (2.2). Note that the RHS of Equation (D.3) is not dependent on a function of A_1, hence integrating both sides yield the indefinite integral terms,

$$\begin{aligned}
I_1(\tau) &= \kappa\int\theta(t)A(\tau)\mathsf{d}\tau, \\
I_2(\tau) &= \tfrac{1}{2}\sigma^2(A_3^2 - A_3)\int 1\mathsf{d}\tau, \\
I_3(\tau) &= \tfrac{1}{2}\eta^2\int A_2^2(\tau), \\
I_4(\tau) &= \eta\sigma\rho A_3\int A_2(\tau)\mathsf{d}\tau.
\end{aligned}$$

For the moment, $I_1(\tau)$ will not be derived further. The second integral, $I_2(\tau)$, can be computed and is given by

$$I_2(\tau) = \frac{1}{2}\sigma^2(-u_2^2 - \mathrm{i}u_2)\tau + C_2 = \frac{1}{2}\sigma^2\mathrm{i}u_2(\mathrm{i}u_2 - 1)\tau + C_2,$$

for the corresponding integration constant $C_2 \in \mathbb{R}$. For the third integral, $I_3(\tau)$, first compute $A_2^2(\tau)$,

$$\begin{aligned} A_2^2(\tau) &= \left(\mathrm{i}u_1\mathrm{e}^{-\kappa\tau} + \frac{\mathrm{i}u_2 - 1}{\kappa}(1 - \mathrm{e}^{-\kappa\tau})\right)^2 \\ &= -u_1^2\mathrm{e}^{-2\kappa\tau} + 2\frac{\mathrm{i}u_1(\mathrm{i}u_2 - 1)}{\kappa}(\mathrm{e}^{-\kappa\tau} - \mathrm{e}^{-2\kappa\tau}) + \frac{(\mathrm{i}u_2 - 1)^2}{\kappa^2}(1 - 2e^{-\kappa\tau} + \mathrm{e}^{-2\kappa\tau}). \end{aligned}$$

Hence, integrating yields,

$$\begin{aligned} \int A_2^2(\tau)\mathsf{d}\tau &= -u_1^2\int \mathrm{e}^{-2\kappa\tau}\mathsf{d}\tau + 2\frac{\mathrm{i}u_1(\mathrm{i}u_2 - 1)}{\kappa}\int(\mathrm{e}^{-\kappa\tau} - \mathrm{e}^{-2\kappa\tau})\mathsf{d}\tau \\ &\quad + \frac{(\mathrm{i}u_2 - 1)^2}{\kappa^2}\int(1 - 2e^{-\kappa\tau} + \mathrm{e}^{-2\kappa\tau})\mathsf{d}\tau \\ &= -\frac{u_1^2}{2\kappa}\mathrm{e}^{-2\kappa\tau} + 2\frac{\mathrm{i}u_1(\mathrm{i}u_2 - 1)}{\kappa}\left(-\frac{1}{\kappa}\mathrm{e}^{-\kappa\tau} + \frac{1}{2\kappa}\mathrm{e}^{-2\kappa\tau}\right) \\ &\quad + \frac{(\mathrm{i}u_2 - 1)^2}{\kappa^2}\left(\tau + \frac{2}{\kappa}\mathrm{e}^{-\kappa\tau} - \frac{1}{2\kappa}\mathrm{e}^{-2\kappa\tau}\right) + \hat{C}_3 \\ &= -\frac{u_1^2}{2\kappa}\mathrm{e}^{-2\kappa\tau} + \frac{\mathrm{i}u_1(\mathrm{i}u_2 - 1)}{\kappa^2}\left(\mathrm{e}^{-2\kappa\tau} - 2\mathrm{e}^{-\kappa\tau}\right) + \frac{(\mathrm{i}u_2 - 1)^2}{2\kappa^3}\left(2\kappa\tau + 4\mathrm{e}^{-\kappa\tau} - \mathrm{e}^{-2\kappa\tau}\right) \\ &\quad + \hat{C}_3, \end{aligned}$$

for the integration constant $\hat{C}_3 \in \mathbb{R}$. As a consequence, the integral $I_3(\tau)$ is now given by

$$\begin{aligned} I_3(\tau) = \frac{\eta^2}{2}\Bigg[-\frac{u_1^2}{2\kappa}\mathrm{e}^{-2\kappa\tau} + \frac{\mathrm{i}u_1(\mathrm{i}u_2 - 1)}{\kappa^2}\left(\mathrm{e}^{-2\kappa\tau} - 2\mathrm{e}^{-\kappa\tau}\right) \\ + \frac{(\mathrm{i}u_2 - 1)^2}{2\kappa^3}\left(2\kappa\tau + 4\mathrm{e}^{-\kappa\tau} - \mathrm{e}^{-2\kappa\tau}\right)\Bigg] + C_3, \end{aligned}$$

where $C_3 = \frac{\eta^2}{2}\hat{C}_3$. The fourth integral, $I_4(\tau)$, is given by

$$\begin{aligned} \eta\sigma\rho A_3\int A_2(\tau)\mathsf{d}\tau &= \eta\sigma\rho\mathrm{i}u_2\left[\mathrm{i}u_1\int \mathrm{e}^{-\kappa\tau}\mathsf{d}\tau + \frac{(\mathrm{i}u_2 - 1)}{\kappa}\int(1 - \mathrm{e}^{-\kappa\tau})\mathsf{d}\tau\right] \\ &= \eta\sigma\rho\mathrm{i}u_2\left[-\frac{\mathrm{i}u_1}{\kappa}\mathrm{e}^{-\kappa\tau} + \frac{(\mathrm{i}u_2 - 1)}{\kappa}\tau + \frac{(\mathrm{i}u_2 - 1)}{\kappa^2}\mathrm{e}^{-\kappa\tau}\right] + C_4, \end{aligned}$$

for the integration constant $C_4 \in \mathbb{R}$ and which can be rewritten in the form,

$$I_4(\tau) = \frac{\eta\sigma\rho\mathrm{i}u_2}{\kappa^2}\left[(\mathrm{i}u_2 - 1)\left(\kappa\tau + \mathrm{e}^{-\kappa\tau}\right) - \kappa\mathrm{i}u_1\mathrm{e}^{-\kappa\tau}\right] + C_4.$$

It remains to compute the first integral $I_1(\tau)$. Rewriting the integral, yields,

$$\begin{aligned} I_1(\tau) &= \kappa\mathrm{i}u_1\int \theta(t)\mathrm{e}^{-\kappa\tau}\mathsf{d}\tau + (\mathrm{i}u_2 - 1)\int \theta(t)(1 - \mathrm{e}^{-\kappa\tau})\mathsf{d}\tau \\ &= (\kappa\mathrm{i}u_1 - (\mathrm{i}u_2 - 1))\int \theta(t)\mathrm{e}^{-\kappa\tau}\mathsf{d}\tau + (\mathrm{i}u_2 - 1)\int \theta(t)\mathsf{d}\tau. \end{aligned}$$

First, the indefinite integral in the first term will be evaluated,

$$\int \theta(t)\mathrm{e}^{-\kappa\tau}\mathsf{d}\tau = \int \theta(T - \tau)\mathrm{e}^{-\kappa\tau}\mathsf{d}\tau, \tag{D.4}$$

where the definition for $\tau = T - t$ is used. Make the following observation. For an integrable function g, its indefinite integral is given by

$$\int g(t)\mathsf{d}t = G(t) + C,$$

for the corresponding integration constant and the anti-derivative G of g. Furthermore, by the fundamental theorem of calculus, it also holds that

$$G(t) = \int_0^t g(z)\mathsf{d}z.$$

Hence, by substitution,

$$\int g(t)\mathsf{d}t = \int_0^t g(z)\mathsf{d}z + C, \tag{D.5}$$

for corresponding integration constant $C \in \mathbb{R}$. It follows from the results of Equation (D.4) and Equation (D.5) that,

$$\int \theta(T-\tau)\mathrm{e}^{-\kappa\tau}\mathsf{d}\tau = \int_0^\tau \theta(T-\tau)\mathrm{e}^{-\kappa\tau}\mathsf{d}\tau + C_5,$$

for some integration constant $C_5 \in \mathbb{R}$. The integral $\int_0^\tau \theta(T-\tau)\mathrm{e}^{-\kappa\tau}\mathsf{d}\tau$ can be computed numerically. Moreover, by directly applying Equation (D.5) to the second integral term given in $I_1(\tau)$, it holds that,

$$\int \theta(t)\mathsf{d}\tau = \int_0^\tau \theta(T-\tau)\mathsf{d}\tau + C_6 = -\int_T^{T-\tau} \theta(s)\mathsf{d}s + C_6 = \int_t^T \theta(s)\mathsf{d}s + C_6,$$

for integration constant $C_6 \in \mathbb{R}$ and using the variable transformation $s = T - \tau$. Recall the definition of θ under the Heath-Jarrow-Morton framework as given in Equation (2.2),

$$\theta(t) = \frac{1}{\kappa}\frac{\partial}{\partial t}f_r(0,t) + f_r(0,t) + \frac{\eta^2}{2\kappa^2}(1-\mathrm{e}^{-2\kappa t}),$$

where $f_r(0,t)$ is the instantaneous forward rate at time $t_0 = 0$ and maturity t. The integral with respect to θ is therefore given by

$$\int_t^T \theta(s)\mathsf{d}s = \frac{1}{\kappa}\int_t^T \frac{\partial}{\partial s}f_r(0,s)\mathsf{d}s + \int_t^T f_r(0,s)\mathsf{d}s + \frac{\eta^2}{2\kappa^2}\int_t^T (1-\mathrm{e}^{-2\kappa s})\mathsf{d}s,$$

where

$$f_r(0,s) = -\frac{\partial \log P_{\mathsf{mkt}}}{\partial s}(0,s). \tag{D.6}$$

From the fundamental theorem of calculus, it holds that,

$$\frac{1}{\kappa}\int_t^T \frac{\partial}{\partial s}f_r(0,s)\mathsf{d}s = \frac{1}{\kappa}(f_r(0,T) - f_r(0,t)). \tag{D.7}$$

By substitution of the relation given in Equation (D.6), the fundamental theorem of calculus can be applied once more,

$$\int_t^T f_r(0,s)\mathsf{d}s = -\int_t^T \frac{\partial \log P_{\mathsf{mkt}}}{\partial s}(0,s)\mathsf{d}s = -\log\left[\frac{P_{\mathsf{mkt}}(0,T)}{P_{\mathsf{mkt}}(0,t)}\right]. \tag{D.8}$$

The last integral can also be evaluated in closed form,

$$\begin{aligned}\int_t^T \frac{\eta^2}{2\kappa^2}(1-\mathrm{e}^{-2\kappa t})\mathsf{d}t &= \frac{\eta^2}{2\kappa^2}\left[\int_t^T 1\mathsf{d}t - \int_t^T \mathrm{e}^{-2\kappa t}\mathsf{d}t\right] \\ &= \frac{\eta^2}{2\kappa^2}\left[(T-t) + \frac{1}{2\kappa}\left(\mathrm{e}^{-2\kappa T} - \mathrm{e}^{-2\kappa t}\right)\right].\end{aligned} \tag{D.9}$$

Combining the results of Equation (D.7), Equation (D.8) and Equation (D.9) yield the closed-form expression for the indefinite integral over θ,

$$\int_t^T \theta(s)\mathsf{d}s = \frac{1}{\kappa}(f_r(0,T) - f_r(0,t)) - \log\left[\frac{P_{\mathsf{mkt}}(0,T)}{P_{\mathsf{mkt}}(0,t)}\right] + \frac{\eta^2}{4\kappa^3}\left(2\kappa(T-t) + \mathrm{e}^{-2\kappa T} - \mathrm{e}^{-2\kappa t}\right).$$

Combining all the results, yields,

$$\begin{aligned}
A_1(u_1,u_2,\tau) &= I_1(\tau) + I_2(\tau) + I_3(\tau) + I_4(\tau) \\
&= (\kappa \mathrm{i}u_1 - (\mathrm{i}u_2 - 1))\int_0^\tau \theta(T-z)\mathrm{e}^{-\kappa z}\mathsf{d}z + (\mathrm{i}u_2 - 1)\int_t^T \theta(s)\mathsf{d}s \\
&+ \frac{1}{2}\sigma^2 \mathrm{i}u_2(\mathrm{i}u_2 - 1)\tau \\
&+ \frac{\eta^2}{2}\left[-\frac{u_1^2}{2\kappa}\mathrm{e}^{-2\kappa\tau} + \frac{\mathrm{i}u_1(\mathrm{i}u_2-1)}{\kappa^2}\left(\mathrm{e}^{-2\kappa\tau} - 2\mathrm{e}^{-\kappa\tau}\right)\right. \\
&\left.+\frac{(\mathrm{i}u_2-1)^2}{2\kappa^3}\left(2\kappa\tau + 4\mathrm{e}^{-\kappa\tau} - \mathrm{e}^{-2\kappa\tau}\right)\right] \\
&+ \frac{\eta\sigma\rho\mathrm{i}u_2}{\kappa^2}\left[(\mathrm{i}u_2 - 1)\left(\kappa\tau + \mathrm{e}^{-\kappa\tau}\right) - \kappa\mathrm{i}u_1\mathrm{e}^{-\kappa\tau}\right] + C,
\end{aligned}$$

where $C \in \mathbb{R}$ is the integration constant resulting from combining all individual integration constants. Using the initial condition, $A(u_1,u_2,0) = 0$, only the terms containing an exponent will remain, from which the integration constant can be derived. Note, that the terms obtained from the integration constant will have an opposite sign in the term. Moreover, the integrals will be zero. Hence,

$$\begin{aligned}
A_1(u_1,u_2,\tau) &= (\kappa \mathrm{i}u_1 - (\mathrm{i}u_2 - 1))\int_0^\tau \theta(T-z)\mathrm{e}^{-\kappa z}\mathsf{d}z + (\mathrm{i}u_2 - 1)\int_t^T \theta(s)\mathsf{d}s \\
&+ \frac{1}{2}\sigma^2 \mathrm{i}u_2(\mathrm{i}u_2 - 1)\tau \\
&+ \frac{\eta^2}{2}\left[-\frac{u_1^2}{2\kappa}(1-\mathrm{e}^{-2\kappa\tau}) + \frac{\mathrm{i}u_1(\mathrm{i}u_2-1)}{\kappa^2}\left(1 + \mathrm{e}^{-2\kappa\tau} - 2\mathrm{e}^{-\kappa\tau}\right)\right. \\
&\left.+\frac{(\mathrm{i}u_2-1)^2}{2\kappa^3}\left(2\kappa\tau + 4\mathrm{e}^{-\kappa\tau} - \mathrm{e}^{-2\kappa\tau} - 3\right)\right] \\
&+ \frac{\eta\sigma\rho\mathrm{i}u_2}{\kappa^2}\left[(\mathrm{i}u_2 - 1)\left(\kappa\tau + \mathrm{e}^{-\kappa\tau} - 1\right) - \kappa\mathrm{i}u_1(1-\mathrm{e}^{-\kappa\tau})\right],
\end{aligned}$$

where the integral over θ is given by

$$\int_t^T \theta(s)\mathsf{d}s = \frac{1}{\kappa}(f_r(0,T) - f_r(0,t)) - \log\left[\frac{P_{\mathsf{mkt}}(0,T)}{P_{\mathsf{mkt}}(0,t)}\right] + \frac{\eta^2}{4\kappa^3}\left(2\kappa(T-t) + \mathrm{e}^{-2\kappa T} - \mathrm{e}^{-2\kappa t}\right),$$

and the integral term $\int_0^\tau \theta(T-\tau)\mathrm{e}^{-\kappa\tau}\mathsf{d}\tau$ can be computed numerically. Furthermore, the function A_2 and A_3 are given by,

$$\begin{aligned}
A_2(u_1,u_2,\tau) &= \mathrm{i}u_1\mathrm{e}^{-\kappa\tau} + \frac{(\mathrm{i}u_2-1)}{\kappa}(1-\mathrm{e}^{-\kappa\tau}), \\
A_3(u_1,u_2,\tau) &= \mathrm{i}u_2.
\end{aligned}$$

□

D.2. Proof of Theorem 2.4

The proof of Theorem 2.4 is very similar to the proof of Theorem 2.3. Though, for completeness of the thesis, the whole derivations will be given in this section. Also, the proof of Theorem 2.4 is divided into two steps. First, the affinity conditions are verified, then the characteristic function is derived.

Proof of Theorem 2.4. The affinity conditions for the decomposed interest rate system are very similar to the derivations under the original Hull-White dynamics. Recall the Hull-White interest rate dynamics decomposed into a stochastic mean reversion process $\{\tilde{r}(t)\}_{t\geq 0}$ and deterministic function ψ;

$$\begin{cases} r(t) &= \tilde{r}(t) + \psi(t), \\ \psi(t) &= r_0 \mathsf{e}^{-\kappa t} + \kappa \int_0^t \theta(z) \mathsf{e}^{-\kappa(t-z)} \mathsf{d}z, \\ \mathsf{d}\tilde{r}(t) &= -\kappa \tilde{r}(t) \mathsf{d}t + \eta \mathsf{d}W_r^{\mathbb{Q}}(t). \end{cases}$$

The BSHW system under the decomposed dynamics is therefore given by,

$$\begin{bmatrix} \mathsf{d}\tilde{r}(t) \\ \mathsf{d}X^C(t) \end{bmatrix} = \begin{bmatrix} \kappa\tilde{r}(t) \\ \tilde{r}(t) + \psi(t) - \frac{1}{2}\sigma^2 \end{bmatrix} \mathsf{d}t + \begin{bmatrix} \eta & 0 \\ \sigma\rho & \sigma\sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} \mathsf{d}\bar{W}_r(t) \\ \mathsf{d}\bar{W}_A(t) \end{bmatrix}, \tag{D.10}$$

for independent Brownian motions $\left\{\bar{W}_r\right\}_{t\geq 0}$ and $\left\{\bar{W}_A\right\}_{t\geq 0}(t)$. Define the vector $\boldsymbol{X}(t) = [\tilde{r}(t), X^C(t)]$ containing the stochastic factors of the system defined in Equation (2.14). The affinity condition for the discounting part of the system is

$$\boldsymbol{r}(t, \boldsymbol{X}(t)) = r_0 + r_1^T \boldsymbol{X}(t) = [1\ 0]^T \begin{bmatrix} \tilde{r}(t) \\ X(t) \end{bmatrix}.$$

Hence $r_0 = 0$ and $r_1 = [1\ 0]^T$. Furthermore, the drift condition is given as

$$\boldsymbol{\mu}(t, \boldsymbol{X}(t)) = \begin{bmatrix} \kappa\tilde{r}(t) \\ \tilde{r}(t) + \psi(t) - \frac{1}{2}\sigma^2 \end{bmatrix} = \begin{bmatrix} 0 \\ \psi(t) - \frac{1}{2}\sigma^2 \end{bmatrix} + \begin{bmatrix} -\kappa & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \tilde{r}(t) \\ X(t) \end{bmatrix} = a_0 + a_1 \boldsymbol{X}(t).$$

As the system in Equation (2.14) has a similar diffusion matrix as the system under the original Hull-White dynamics, the affinity coefficients corresponding to the system under the decomposed interest rate are the same. Hence, the system under the decomposed interest rate, defined in Equation (2.14) is included in the affine class of models. Moreover, its affinity coefficients are given by

$$a_0 = \begin{bmatrix} 0 \\ \psi(t) - \frac{1}{2}\sigma^2 \end{bmatrix}, \qquad a_1 = \begin{bmatrix} -\kappa & 0 \\ 1 & 0 \end{bmatrix},$$
$$r_0 = 0, \qquad r_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$
$$c_0 = \begin{bmatrix} \eta^2 & \eta\sigma\rho \\ \eta\sigma\rho & \sigma^2 \end{bmatrix}, \qquad c_1 = 0.$$

Hence, the model under the decomposed interest rate process $\{\tilde{r}(t)\}_{t\geq 0}$ is also included in the affine class of diffusion models. This concludes the first part of the proof.

In the second part of the proof, the discounted characteristic function corresponding to the system given in Equation (D.10) will be derived. The corresponding complex differential equations, defining the discounted characteristic function, are then given by;

$$\begin{cases} \dfrac{\mathsf{d}A}{\mathsf{d}\tau} &= -r_0 + B^{\mathsf{T}} a_0 + \frac{1}{2} B^{\mathsf{T}} c_0 B, \\ \dfrac{\mathsf{d}B}{\mathsf{d}\tau} &= -r_1 + a_1^{\mathsf{T}} B + \frac{1}{2} B^{\mathsf{T}} c_1 B, \end{cases} \tag{D.11}$$

for functions $A^{\mathsf{T}} \equiv A(u, w, \tau)$ and $B^{\mathsf{T}} = [B_1\ B_2]^{\mathsf{T}} = [B_1(u, w, \tau)\ B_2(u, w, \tau)]^{\mathsf{T}}$ with initial conditions $A(u, w, 0) = 0$, $B_1(u, w, 0) = \mathrm{i}u$ and $B_2(u, w, 0) = \mathrm{i}w$ and where $\tau = T - t$ the time till maturity T. Substitution of the results obtained in Equation (D.11), yields therefore the following expression for the function B;

$$\begin{bmatrix} \dfrac{\mathsf{d}B_1}{\mathsf{d}\tau} \\ \dfrac{\mathsf{d}B_2}{\mathsf{d}\tau} \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \begin{bmatrix} -\kappa & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} B_1 \\ B2 \end{bmatrix} + 0.$$

Hence, for the function B_2 the following relation must be satisfied,

$$\frac{\mathsf{d}B_2}{\mathsf{d}\tau} = 0,$$

with initial condition $B_2 = \mathrm{i}w$ and therefore it holds that

$$B_2(u, w, \tau) = \mathrm{i}w.$$

Using the closed form function for B_2, the relation for the function B_1 can be written as,

$$\frac{\mathsf{d}B_1}{\mathsf{d}\tau} = -1 - \kappa B_1 + \mathrm{i}w \iff \frac{\mathsf{d}B_1}{\mathsf{d}\tau} + \kappa B_1 = \mathrm{i}w - 1 \iff \varphi\frac{\mathsf{d}B_1}{\mathsf{d}\tau} + \varphi\kappa B_1 = \varphi(\mathrm{i}w - 1), \tag{D.12}$$

for a test function $\varphi \equiv \varphi(\tau)$. From the chain rule for differentiation, it is known that

$$\frac{\mathsf{d}B_1\varphi}{\mathsf{d}\tau} = \frac{\mathsf{d}\varphi}{\mathsf{d}\tau}B_1 + \varphi\frac{\mathsf{d}B_1}{\mathsf{d}\tau},$$

from which it can be seen that

$$\frac{\mathsf{d}\varphi}{\mathsf{d}\tau} = \varphi\kappa,$$

must hold in order to solve Equation (D.12). As a consequence, it must hold that

$$\varphi(\tau) = \mathrm{e}^{\kappa\tau}C, \tag{D.13}$$

for some constant $C \in \mathbb{R}$. Substitution of the result in Equation (D.13) into Equation (D.12) yields,

$$\frac{\mathsf{d}B_1\varphi}{\mathsf{d}\tau} = \mathrm{e}^{\kappa\tau}C(\mathrm{i}w - 1).$$

Integrating with respect to τ,

$$\varphi B_1 = C(\mathrm{i}w - 1)\int \mathrm{e}^{\kappa\tau}\mathsf{d}\tau = \frac{C(\mathrm{i}w - 1)}{\kappa}\mathrm{e}^{\kappa\tau} + C_1,$$

for a constant $C_1 \in \mathbb{R}$ and rearranging the terms, yields

$$B_1 = \frac{\mathrm{i}w - 1}{\kappa} + C_2\mathrm{e}^{-\kappa\tau},$$

with $C_2 \in \mathbb{R}$ the combined constant following from C_1 and C. Using the initial condition $B_1 = \mathrm{i}u$, the following expression for $B_1(u, w, \tau)$ can be derived,

$$B_1(u, w, \tau) = \frac{\mathrm{i}w - 1}{\kappa} + \left[-\frac{\mathrm{i}w - 1}{\kappa} + \mathrm{i}u\right]\mathrm{e}^{-\kappa\tau} = \frac{\mathrm{i}w - 1}{\kappa}\left[1 - \mathrm{e}^{-\kappa\tau}\right] + \mathrm{i}u\mathrm{e}^{-\kappa\tau}.$$

It is left to obtain an expression for A. The differential equation for the function A states that

$$\frac{\mathsf{d}A}{\mathsf{d}\tau} = [B_1\ B_2]\begin{bmatrix}0\\ \psi(t) - \frac{1}{2}\sigma^2\end{bmatrix} + \frac{1}{2}[B_1\ B_2]\begin{bmatrix}\eta^2 & \eta\sigma\rho\\ \eta\sigma\rho & \sigma^2\end{bmatrix}\begin{bmatrix}B_1\\ B_2\end{bmatrix},$$

which can be rewritten in a more convenient form,

$$A(u, w, \tau) = \int B_2\psi(t)\mathsf{d}\tau + \frac{1}{2}\sigma^2\int B_2^2 - B_2\mathsf{d}\tau + \frac{1}{2}\eta^2\int B_1^2\mathsf{d}\tau + \eta\sigma\rho\int B_1B_2\mathsf{d}\tau. \tag{D.14}$$

It, therefore, remains to compute the remaining integrals. Note that, with respect to τ, B_2 forms a constant function, hence,

$$\int B_2\psi(t)\mathsf{d}\tau = \mathrm{i}w\int \psi(t)\mathsf{d}\tau.$$

Note that using Equation (D.5) from the proof given in Appendix D.1, it holds for the indefinite integral,

$$\int \psi(t)\mathsf{d}\tau = \int \psi(T - \tau)\mathsf{d}\tau = \int_0^\tau \psi(T - z)\mathsf{d}z + C_3, \tag{D.15}$$

for some integration constant $C_3 \in \mathbb{R}$ and where the relation $\tau = T - t$ was used. Under the HJM framework, the function ψ is known in closed-form and given by [32]

$$\psi(t) = \mathrm{e}^{-\kappa t} r_0 + f_r(0,t) - f_r(0,0)\mathrm{e}^{-\kappa t} + \frac{\eta^2}{\kappa^2}\mathrm{e}^{-\kappa t}(\cosh(\kappa t) - 1),$$

for the initial value of the short rate process $r(t_0) = r_0$ and $\{f_r(t)\}_{t \geq 0}$ the forward rate process. This enables the possibility to compute the integral in Equation (D.15) numerically. For the second term in Equation (D.14), it holds that

$$\frac{1}{2}\sigma^2 \int B_2^2 - B_2 \mathsf{d}\tau = \frac{1}{2}\sigma^2 \mathrm{i}w(\mathrm{i}w - 1)\tau + C_4,$$

for a constant $C_4 \in \mathbb{R}$. Furthermore, the integral with respect to the third term in Equation (D.14) yields,

$$\begin{aligned} \frac{1}{2}\eta^2 \int B_1^2 \mathsf{d}\tau &= \frac{1}{2}\eta^2 \int \left[\frac{\mathrm{i}w - 1}{\kappa}(1 - \mathrm{e}^{-\kappa\tau}) + \mathrm{i}u\mathrm{e}^{-\kappa\tau}\right]^2 \mathsf{d}\tau \\ &= \frac{1}{2}\eta^2 \int \frac{(\mathrm{i}w - 1)^2}{\kappa^2}(1 - \mathrm{e}^{-\kappa\tau})^2 + \frac{\mathrm{i}w - 1}{\kappa}(1 - \mathrm{e}^{-\kappa\tau})\mathrm{i}u\mathrm{e}^{-\kappa\tau} + (\mathrm{i}u)^2\mathrm{e}^{-2\kappa\tau}\mathsf{d}\tau. \end{aligned} \tag{D.16}$$

The first term in the integral, yields,

$$\int (1 - \mathrm{e}^{-\kappa\tau})\mathsf{d}\tau = \int 1 - 2\mathrm{e}^{-\kappa\tau} + \mathrm{e}^{-2\kappa\tau}\mathsf{d}\tau = \tau + \frac{2}{\kappa}\mathrm{e}^{-\kappa\tau} - \frac{1}{2\kappa}\mathrm{e}^{-2\kappa\tau} + \bar{C}_1,$$

for $\bar{C}_1 \in \mathbb{R}$. Moreover, the third term within the integral given in Equation (D.16) yields,

$$\int (1 - \mathrm{e}^{-\kappa\tau})\mathrm{e}^{-\kappa\tau}\mathsf{d}\tau = \int \mathrm{e}^{-\kappa\tau} - \mathrm{e}^{-2\kappa\tau}\mathsf{d}\tau = -\frac{1}{\kappa}\mathrm{e}^{-\kappa\tau} + \frac{1}{2\kappa}\mathrm{e}^{-2\kappa\tau} + \bar{C}_2,$$

for $\bar{C}_2 \in \mathbb{R}$. Furthermore, the last term from the integral in Equaiton (D.16), can be written as

$$\int \mathrm{e}^{-2\kappa\tau}\mathsf{d}\tau = -\frac{1}{2\kappa}\mathrm{e}^{-2\kappa\tau} + \bar{C}_3,$$

for $\bar{C}_3 \in \mathbb{R}$. Combining the results for the individual terms of the integral given in Equation (D.16), yields,

$$\begin{aligned} &\frac{1}{2}\eta^2 \int B_1^2 \mathsf{d}\tau = \\ &\qquad \frac{\eta^2}{2}\left[\frac{(\mathrm{i}w - 1)^2}{2\kappa^3}\left(2\tau\kappa + 4\mathrm{e}^{-\kappa\tau} - \mathrm{e}^{-2\kappa\tau}\right) + \frac{(\mathrm{i}w - 1)\mathrm{i}u}{2\kappa^2}\left[\mathrm{e}^{-2\kappa\tau} - 2\mathrm{e}^{-\kappa\tau}\right] - \frac{(\mathrm{i}u)^2}{2\kappa}\mathrm{e}^{-2\kappa\tau}\right] + C_5, \end{aligned}$$

combining the individual constants $\bar{C}_1,\ \bar{C}_2,\ \bar{C}_3$. The last term of the integral given in Equation (D.14) is given by,

$$\begin{aligned} \eta\sigma\rho\mathrm{i}w \int B_1 \mathsf{d}\tau &= \eta\sigma\rho\mathrm{i}w \int \frac{\mathrm{i}w - 1}{\kappa}\left[1 - \mathrm{e}^{-\kappa\tau}\right] + \mathrm{i}u\mathrm{e}^{-\kappa\tau}\mathsf{d}\tau \\ &= \eta\sigma\rho\mathrm{i}w\left[\frac{\mathrm{i}w - 1}{\kappa^2}(\kappa\tau + \mathrm{e}^{-\kappa\tau}) - \frac{\mathrm{i}u}{\kappa}\mathrm{e}^{-\kappa\tau}\right] + C_6, \end{aligned}$$

for $C_5 \in \mathbb{R}$. Combining all the results for the function A, yields,

$$\begin{aligned} A(u,w,\tau) &= \int B_2\psi(t)\mathsf{d}\tau + \frac{1}{2}\sigma^2 \int B_2^2 - B_2\mathsf{d}\tau + \frac{1}{2}\eta^2 \int B_1^2 \mathsf{d}\tau + \eta\sigma\rho \int B_1 B_2 \mathsf{d}\tau \\ &= \mathrm{i}w \int_0^\tau \psi(T - z)\mathsf{d}z + \frac{1}{2}\sigma^2 \mathrm{i}w(\mathrm{i}w - 1)\tau \\ &\quad + \frac{\eta^2}{2}\left[\frac{(\mathrm{i}w - 1)^2}{2\kappa^3}\left(2\tau\kappa + 4\mathrm{e}^{-\kappa\tau} - \mathrm{e}^{-2\kappa\tau}\right) + \frac{(\mathrm{i}w - 1)\mathrm{i}u}{2\kappa^2}\left[\mathrm{e}^{-2\kappa\tau} - 2\mathrm{e}^{-\kappa\tau}\right] - \frac{(\mathrm{i}u)^2}{2\kappa}\mathrm{e}^{-2\kappa\tau}\right] \\ &\quad + \eta\sigma\rho\mathrm{i}w\left[\frac{\mathrm{i}w - 1}{\kappa^2}(\kappa\tau + \mathrm{e}^{-\kappa\tau}) - \frac{\mathrm{i}u}{\kappa}\mathrm{e}^{-\kappa\tau}\right] + C, \end{aligned}$$

where $C \in \mathbb{R}$ follows from combining all the individual integration constants. Using the initial value $A(u, w, 0) = 0$, corresponding to the differential equation, yields that A is given by,

$$\begin{aligned} A(u,w,\tau) = \mathrm{i}w \int_0^\tau \psi(T-z)\mathrm{d}z + \frac{1}{2}\sigma^2 \mathrm{i}w(\mathrm{i}w-1)\tau \\ + \frac{\eta^2}{2}\left[\frac{(\mathrm{i}w-1)^2}{2\kappa^3}\left(2\tau\kappa + 4\mathrm{e}^{-\kappa\tau} - \mathrm{e}^{-2\kappa\tau} - 3\right) + \frac{(\mathrm{i}w-1)\mathrm{i}u}{2\kappa^2}\left[\mathrm{e}^{-2\kappa\tau} - 2\mathrm{e}^{-\kappa\tau} + 1\right]\right. \\ \left. -\frac{(\mathrm{i}u)^2}{2\kappa}(\mathrm{e}^{-2\kappa\tau} - 1)\right] + \eta\sigma\rho\mathrm{i}w\left[\frac{\mathrm{i}w-1}{\kappa^2}(\kappa\tau + \mathrm{e}^{-\kappa\tau} - 1) - \frac{\mathrm{i}u}{\kappa}(\mathrm{e}^{-\kappa\tau} - 1)\right], \end{aligned} \tag{D.17}$$

where the integral $\int_0^\tau \psi(T-z)\mathrm{d}z$ can be computed numerically using the expression [32],

$$\psi(t) = \mathrm{e}^{-\kappa t} r_0 + f_r(0,t) - f_r(0,0)\mathrm{e}^{-\kappa t} + \frac{\eta^2}{\kappa^2}\mathrm{e}^{-\kappa t}(\cosh(\kappa t) - 1),$$

for the function ψ. Therefore, the discounted characteristic function of the system defined in Equation (D.10) can be written in the form,

$$\phi_{\tilde{r}(t),X^C}(u,w,t,T) = \mathrm{e}^{A(u,w,\tau) + B_1(u,w,\tau)\tilde{r}(t) + B_2(u,w,\tau)X^C(t)}.$$

Moreover, the function A_1 is given as in Equation (D.17) and B_1 and B_2 are given by

$$\begin{cases} B_1(u,w,\tau) &= \dfrac{\mathrm{i}w-1}{\kappa}\left[1 - \mathrm{e}^{-\kappa\tau}\right] + \mathrm{i}u\mathrm{e}^{-\kappa\tau}, \\ B_2(u,w,\tau) &= \mathrm{i}w. \end{cases}$$

□

D.3. Proof of Lemma 2.2

The proof will be based on a slightly simpler lemma.

Lemma D.1. *The characteristic function under the T-forward measure $\mathbb{Q}^T$ is related to the discounted characteristic function under the measure $\mathbb{Q}$ via the value of a zero-coupon bond at time t maturing at time T, i.e.*

$$\mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\mathrm{d}s + \mathrm{i}\left(ur(T) + wX^C(T)\right)} \;\middle|\; \mathcal{F}_t\right] = P(t,T)\mathbb{E}^T\left[\mathrm{e}^{\mathrm{i}\left(ur(T) + wX^C(T)\right)} \;\middle|\; \mathcal{F}_t\right].$$

Proof of Lemma D.1. The conditional expectation of a function under the equivalent martingale measure can be transformed into a conditional expectation under a different, equivalent measure. Consider the discounted characteristic function,

$$\mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\mathrm{d}s + \mathrm{i}\left(ur(T) + wX^C(T)\right)} \;\middle|\; \mathcal{F}_t\right] = M(t)\mathbb{E}\left[\frac{\mathrm{e}^{\mathrm{i}\left(ur(T) + wX^C(T)\right)}}{M(T)} \;\middle|\; \mathcal{F}_t\right], \tag{D.18}$$

in terms of the money-market account $\{M(t)\}_{t\geq 0}$ under the equivalent martingale measure $\mathbb{Q}$. Define the Radon-Nikodym derivative as

$$\frac{\mathrm{d}\mathbb{Q}^T}{\mathrm{d}\mathbb{Q}} = \frac{M(t)P(T,T)}{M(T)P(t,T)}, \tag{D.19}$$

for the T-forward measure $\mathbb{Q}^T$ and the value of a zero-coupon bond $P(t,T)$ at time t with maturity T.

Substitution of the Radon-Nikodym derivative given in Equation (D.19) into Equation (D.18), yields,

$$\begin{aligned} M(t)\mathbb{E}\left[\frac{\mathrm{e}^{\mathrm{i}\left(ur(T)+wX^C(T)\right)}}{M(T)}\;\middle|\;\mathcal{F}_t\right] &= M(t)\int_{\mathbb{R}^2}\frac{\mathrm{e}^{\mathrm{i}\left(ur(T)+wX^C(T)\right)}}{M(T)}\mathsf{d}\mathbb{Q} \\ &= M(t)\int_{\mathbb{R}^2}\frac{\mathrm{e}^{\mathrm{i}\left(ur(T)+wX^C(T)\right)}}{M(T)}\frac{M(T)P(T,T)}{M(t)P(t,T)}\mathsf{d}\mathbb{Q}^T \\ &= \int_{\mathbb{R}^2}\mathrm{e}^{\mathrm{i}\left(ur(T)+wX^C(T)\right)}\frac{P(t,T)}{P(T,T)}\mathsf{d}\mathbb{Q}^T \\ &= P(t,T)\mathbb{E}^T\left[\mathrm{e}^{\mathrm{i}\left(ur(T)+wX^C(T)\right)}\;\middle|\;\mathcal{F}_t\right], \end{aligned}$$

where was used that $P(T,T)=1$ and that $P(t,T)$ is $\mathcal{F}_t$ measurable. □

Using the Lemma D.1, the proof of Lemma 2.2 is straightforward.

Proof of Lemma 2.2. Given Lemma 2.1, the discounted characteristic function under the interest rate process $\{r(t)\}_{t\geq 0}$ can be decomposed, such that the resulting expression is expressed in terms of the process $\{\tilde{r}(t)\}_{t\geq 0}$, i.e.

$$\mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\mathsf{d}s+\mathrm{i}\left(ur(T)+wX^C(T)\right)}\;\middle|\;\mathcal{F}_t\right] = \mathrm{e}^{-\int_t^T\psi(s)\mathsf{d}s+\mathrm{i}u\psi(T)}\mathbb{E}\left[\mathrm{e}^{-\int_t^T\tilde{r}(s)\mathsf{d}s+\mathrm{i}\left(u\tilde{r}(T)+wX^C(T)\right)}\;\middle|\;\mathcal{F}_t\right].$$

Applying Lemma 2.2 to the left hand side of Equation D.3 yields,

$$\mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\mathsf{d}s+\mathrm{i}\left(ur(T)+wX^C(T)\right)}\;\middle|\;\mathcal{F}_t\right] = P(t,T)\mathbb{E}^T\left[\mathrm{e}^{\mathrm{i}\left(ur(T)+wX^C(T)\right)}\;\middle|\;\mathcal{F}_t\right].$$

Also at the right hand side of Equation D.3, the decomposition given in Lemma 2.1 can be applied,

$$\mathbb{E}^T\left[\mathrm{e}^{\mathrm{i}\left(ur(T)+wX^C(T)\right)}\;\middle|\;\mathcal{F}_t\right] = \mathrm{e}^{-\mathrm{i}u\psi(T)}\mathbb{E}^T\left[\mathrm{e}^{\mathrm{i}\left(ur(T)+wX^C(T)\right)}\;\middle|\;\mathcal{F}_t\right].$$

Combining the results yield,

$$\begin{aligned} \mathbb{E}\left[\mathrm{e}^{-\int_t^T\tilde{r}(s)\mathsf{d}s+\mathrm{i}\left(u\tilde{r}(T)+wX^C(T)\right)}\;\middle|\;\mathcal{F}_t\right] &= \mathrm{e}^{\int_t^T\psi(s)\mathsf{d}s-\mathrm{i}u\psi(T)}\mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\mathsf{d}s+\mathrm{i}\left(ur(T)+wX^C(T)\right)}\;\middle|\;\mathcal{F}_t\right] \\ &= \mathrm{e}^{\int_t^T\psi(s)\mathsf{d}s-\mathrm{i}u\psi(T)}P(t,T)\mathbb{E}^T\left[\mathrm{e}^{\mathrm{i}\left(ur(T)+wX^C(T)\right)}\;\middle|\;\mathcal{F}_t\right] \\ &= \mathrm{e}^{\int_t^T\psi(s)\mathsf{d}s}P(t,T)\mathbb{E}^T\left[\mathrm{e}^{\mathrm{i}\left(u\tilde{r}(T)+wX^C(T)\right)}\;\middle|\;\mathcal{F}_t\right]. \end{aligned}$$

Hence, it holds that

$$\mathbb{E}^T\left[\mathrm{e}^{\mathrm{i}\left(u\tilde{r}(T)+wX^C(T)\right)}\;\middle|\;\mathcal{F}_t\right] = \frac{\mathrm{e}^{-\int_t^T\psi(s)\mathsf{d}s}\mathbb{E}\left[\mathrm{e}^{-\int_t^T\tilde{r}(s)\mathsf{d}s+\mathrm{i}\left(u\tilde{r}(T)+wX^C(T)\right)}\;\middle|\;\mathcal{F}_t\right]}{P(t,T)},$$

which concludes the proof. □

D.4. Proof of Lemma 2.3

The proof follows from the results of Section 2.3.1, Theorem 2.4 and Lemma 2.2.

Proof. From the results of Lemma 2.2, the characteristic function of the system given in Equation 2.14 can be written in the known, closed form, solution for the discounted characteristic function under the equivalent martingale measure $\mathbb{Q}$. Direct substitution for the terms $P(t,T)$ and the discounted characteristic function as given in Theorem 2.4 yields,

$$\begin{aligned} \mathbb{E}^T\left[\mathrm{e}^{\mathrm{i}\left(u\tilde{r}(T)+wX^C(T)\right)}\;\middle|\;\mathcal{F}_t\right] &= \mathrm{e}^{-\int_t^T\psi(s)\mathsf{d}s+\bar{A}(u,w,\tau)+A_2(u,w,\tau)\tilde{r}(t)+\mathrm{i}wX^C(t)} \\ &\quad\cdot\mathrm{e}^{\int_t^T\psi(s)\mathsf{d}s-A(\tau)+B(\tau)\tilde{r}(t)} \\ &= \mathrm{e}^{-A(\tau)+B(\tau)\tilde{r}(t)+\bar{A}(u,w,\tau)+A_2(u,w,\tau)\tilde{r}(t)+\mathrm{i}wX^C(t)}. \end{aligned} \tag{D.20}$$

The expression in Equation (D.20) can be rewritten in a deterministic part and a stochastic part depending on $\tilde{r}(t)$ and $X^C(t)$. By substitution of the expression for A_2 in terms of the function B, Equation (D.20) becomes

$$\begin{aligned}\mathbb{E}^T\left[\mathrm{e}^{\mathrm{i}\left(u\tilde{r}(T)+wX^C(T)\right)}\;\middle|\;\mathcal{F}_t\right] &= \mathrm{e}^{-A(\tau)+B(\tau)\tilde{r}(t)+\bar{A}(u,w,\tau)+A_2(u,w,\tau)\tilde{r}(t)+\mathrm{i}wX^C(t)}\\ &= \mathrm{e}^{-A(\tau)+\bar{A}(u,w,\tau)}\mathrm{e}^{B(\tau)\tilde{r}(t)+A_2(u,w,\tau)\tilde{r}(t)+\mathrm{i}wX^C(t)}\\ &= \mathrm{e}^{-A(\tau)+\bar{A}(u,w,\tau)}\mathrm{e}^{\mathrm{i}ue^{-\kappa\tau}\tilde{r}(t)+\mathrm{i}w(X^C(t)+B(\tau)\tilde{r}(t))},\end{aligned}$$

where the decomposition of A_2 in terms of $B(\tau)$ given in Equation (2.13) from Section 2.3.1 was used. □

D.5. Proof of Theorem 2.5

Proof of Theorem 2.5. Recall the definition for the functions g_r and g, given by

$$g_r(x) = x\mathrm{e}^{-\kappa\tau},\; g(x,y) = y + B(\tau)x,$$

and define the variable transformation

$$Z_r(T) = \tilde{r}(T) - g_r(\tilde{r}(t)),\; Z(T) = X^C(T) - g(\tilde{r}(t), X^C(t)).$$

From the result of Lemma 2.3, the characteristic function of the variable transformation is know. As the values $g_r(\tilde{r}(t))$ and $g(\tilde{r}(t), X^C(t))$ are $\mathcal{F}_t$ measurable, they can be taken outside of the expectation. Computing the characteristic function of the proposed variable transformation under the T-forward measure yields,

$$\begin{aligned}\phi_{Z_r,Z}(u,w,t,T) &= \mathbb{E}^T\left[\mathrm{e}^{\mathrm{i}(uZ_r(T)+wZ(T))}\;\middle|\;\mathcal{F}_t\right]\\ &= \mathrm{e}^{-\mathrm{i}ug_r(\tilde{r}(t))-\mathrm{i}wg(\tilde{r}(t),X^C(t))}\mathbb{E}^T\left[\mathrm{e}^{\mathrm{i}(u\tilde{r}(T)+wX^C(T))}\;\middle|\;\mathcal{F}_t\right]\\ &= \mathrm{e}^{-\mathrm{i}ug_r(\tilde{r}(t))-\mathrm{i}wg(\tilde{r}(t),X^C(t))}\mathrm{e}^{-A(\tau)+\bar{A}(u,w,\tau)}\mathrm{e}^{\mathrm{i}ug_r(\tilde{r}(t))+\mathrm{i}wg(\tilde{r}(t),X^C(t))}\\ &= \mathrm{e}^{\bar{A}(u,w,\tau)-A(\tau)},\end{aligned}$$

which shows that the variable transformation is chosen in such a way that the resulting functions $g_r(\tilde{r}(t))$ and $g(\tilde{r}(t), X^C(t))$ cancel out the requirement of the state variables to compute the characteristic function. □

D.6. Proof of Theorem 4.1

In this section the proof of Theorem 4.1 will be given. As the Fourier transform is considered in a probability sense, the Fourier transform will be addressed as the characteristic function of a density.

Proof of Theorem 4.1. Let f be a density function and consider Definition 4.1 for its continuous Fourier transform. Following the described approach, define the the truncated interval $[-L, L]$ for which $f(x) \approx 0$ if $x \in \mathbb{R}\backslash[-L, L]$. Furthermore, let $N \in \mathbb{N}$ and consider a partition $x_0, \ldots, x_{N-1}$ of $N-1$ intervals $[x_j, x_{j+1}]$ for $j \in \{0, \ldots, N-1\}$ such that the Newton-Cotes integration rule with trapezoidal integration weights is given by

$$\begin{aligned}\mathrm{F}[f](u) &\approx \sum_{j=0}^{N-1}\int_{x_j}^{x_{j+1}} e^{iux}f(x)\mathrm{d}x \qquad (D.21)\\ &\approx \Delta_x\left[\sum_{j=0}^{N-1} e^{iux_j}f(x_j) - \frac{1}{2}\left(e^{iux_0}f(x_0) + e^{iux_{N-1}}f(x_{N-1})\right)\right],\end{aligned}$$

where $\Delta_x = x_{j+1} - x_j$ is the interval length. Next, specify the partition as an equally spaced grid $x_j = \left(j - \frac{N}{2}\right)\Delta_x$ for a chosen $\Delta_x \in \mathbb{R}_+$. Let $u_n = \left(n - \frac{N}{2}\right)\Delta_u$ and impose the restriction,

$$\Delta_u\Delta_x = \frac{2\pi}{N}. \qquad (D.22)$$

Note that by definition the multiplication of the grid nodes now yields,

$$\begin{aligned} x_j u_n &= \left(j - \frac{N}{2}\right)\Delta_x \left(n - \frac{N}{2}\right)\Delta_u \\ &= \frac{2\pi}{N}\left(nj - n\frac{N}{2} - j\frac{N}{2} + \frac{N^2}{4}\right) \\ &= 2\pi\frac{nj}{N} - \left(j + n - \frac{N}{2}\right)\pi. \end{aligned} \tag{D.23}$$

Recall the discrete Fourier transform, defined as

$$D_n\left(x_j\right) = \sum_{j=0}^{N-1} e^{i2\pi\frac{nj}{N}} x_j$$

Substitution of Equation (D.23)into Equation (D.21), gives

$$\begin{aligned} \mathrm{F}[f](u_n) &\approx \Delta_x \left[\sum_{j=0}^{N-1} e^{i\frac{nj}{N}2\pi}(-1)^{j+n-\frac{N}{2}} f(x_j) - \frac{1}{2}\left(e^{iu_n x_0} f(x_0) + e^{iu_n x_{N-1}} f(x_{N-1})\right)\right] \\ &= \Delta_x \left[(-1)^{n-\frac{N}{2}} D_n\left(f(x_j)(-1)^j\right) - \frac{1}{2}\left(e^{iu_n x_0} f(x_0) + e^{iu_n x_{N-1}} f(x_{N-1})\right)\right] \\ &:= \bar{\mathrm{F}}[f](u_n), \end{aligned}$$

where the Euler identity $e^{i\pi} - 1 = 0$ is used.

An approximation of the inverse Fourier transform is derived in a very similar way. Denote the Fourier transform of an integrable function as $\phi(u) = \mathrm{F}[f](u)$ and recall the inverse of the Fourier transform from Definition 4.1,

$$\mathrm{F}^{-1}[\phi](x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-iux}\phi(u)\mathsf{d}u.$$

First, a truncated interval $[-L, L]$ is determined such that the Fourier transform ϕ is almost zero outside the interval. Let $N \in \mathbb{N}$. Define a partition over the interval, u_n for $n \in \{0, \ldots, N-1\}$, with interval length $\Delta_u = u_{n+1} - u_n$. Next, the Newton-Cotes integral rule is used to obtain an approximation of the integral,

$$\begin{aligned} 2\pi\mathrm{F}^{-1}[\phi](x) &\approx \sum_{n=0}^{N-1}\int_{u_n}^{u_{n+1}} e^{-iux}\phi(u)\mathsf{d}u \\ &\approx \Delta_u\left[\sum_{n=0}^{N-1} e^{iu_n x}\phi(u_n)\mathsf{d}u - \frac{1}{2}\left(e^{iu_0 x}\phi(u_0) + e^{iu_{N-1}x}\phi(u_{N-1})\right)\right] \end{aligned}$$

Specifically define the grid $\Delta_x = x_{j+1} - x_j$ is the interval length. Next, specify the partition as an equally spaced grid $u_n = \left(n - \frac{N}{2}\right)\Delta_u$ and let $x_j = \left(j - \frac{N}{2}\right)\Delta_x$ for a chosen $\Delta_x \in \mathbb{R}_+$. Furthermore, impose the relation specified in Equation (D.22). As the same argument concerning Equation (D.23) holds, it can be derived that

$$\begin{aligned} 2\pi\mathrm{F}^{-1}[\phi](x_j) &\approx \Delta_u\left[(-1)^{j-\frac{N}{2}} D_j^{-1}\left(\phi(u_n)(-1)^n\right) - \frac{1}{2}\left(e^{iu_0 x_j}\phi(u_0) + e^{iu_{N-1}x_j}\phi(u_{N-1})\right)\right] \\ &:= 2\pi\bar{\mathrm{F}}^{-1}[f](u_n), \end{aligned}$$

where the inverse discrete Fourier transform is defined as

$$D_j^{-1}\left(u_n\right) = \sum_{n=0}^{N-1} e^{-i2\pi\frac{jn}{N}} u_n.$$

□

D.7. Proof of Theorem 4.2

Proof of Theorem 4.2. Recall the Fourier transform in two dimensions,

$$\mathrm{F}_2[f](\boldsymbol{u}) = \int_{\mathbb{R}^2} e^{i\boldsymbol{u}^T\boldsymbol{x}} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x},$$

for a density function f and $\boldsymbol{x} = [x\ y], \boldsymbol{u} = [u\ w], \in \mathbb{R}^2$. Consider an interval $[-L_0, L_0] \times [-L_1, L_1]$ for $L_0, L_1 \in \mathbb{R}_+$ for which the density function is almost zero outside the interval. Furthermore, consider partitions $x_j = \left(j - \frac{N_0}{2}\right)\Delta_x,\ \Delta_x \in \mathbb{R}_+$ and $y_k = \left(k - \frac{N_1}{2}\right)\Delta_y,\ \Delta_y \in \mathbb{R}_+$ where the partition (x_j, y_k) spans the interval $[-L_0, L_0] \times [-L_1, L_1]$ and $N_0, N_1 \in \mathbb{N}$. The truncated multidimensional integral can be approximated by the use of the Newton-Cotes integration rule. Applying trapezoidal weights results in

$$\begin{aligned}
\mathrm{F}_k[f](u,w) &\approx \sum_{j=0}^{N_0-1} \int_{x_j}^{x_{j+1}} e^{iux} \sum_{k=0}^{N_1-1} \int_{y_k}^{y_{k+1}} e^{iwy} f(x,y)\mathrm{d}y\mathrm{d}x \qquad \text{(D.24)}\\
&\approx \Delta_x \sum_{j=0}^{N_0-1} e^{iux_j} \Delta_y \left[\sum_{k=0}^{N_1-1} e^{iwy_k} f(x_j, y_k) - \frac{1}{2}\left(e^{iwy_0} f(x_j, y_0) + e^{iwy_{N_1-1}} f(x_j, y_{N_1-1})\right)\right]\\
&- \frac{\Delta_x\Delta_y}{2}\left[e^{iux_0} \sum_{k=0}^{N_1-1} e^{iwy_k} f(x_0, y_k) - \frac{e^{iux_0}}{2}\left(e^{iwy_0} f(x_0, y_0) + e^{iwy_{N_1-1}} f(x_0, y_{N_1-1})\right)\right.\\
&\left.+ e^{iux_{N_0-1}} \sum_{k=0}^{N_1-1} e^{iwy_k} f(x_{N_0-1}, y_k) - \frac{e^{iux_{N_0-1}}}{2}\left(e^{iwy_0} f(x_{N_0-1}, y_0) + e^{iwy_{N_1-1}} f(x_{N_0-1}, y_{N_1-1})\right)\right].
\end{aligned}$$

Here, the results of the one dimensional given in Equation (D.21) is applied in both the x_j and the y_k grid. Rearranging the terms in Equation (D.24)yields,

$$\begin{aligned}
\frac{\phi(u,w)}{\Delta_x\Delta_y} &\approx \sum_{j=0}^{N_1-1}\sum_{k=0}^{N_1-1} e^{i(ux_j+wy_k)} f(x_j, y_k) \qquad \text{(D.25)}\\
&- \frac{1}{2}\left(e^{iux_0} \sum_{k=0}^{N_1-1} e^{iwy_k} f(x_0, y_k) + e^{iux_{N_0-1}} \sum_{k=0}^{N_1-1} e^{iwy_k} f(x_{N_0-1}, y_k)\right.\\
&\left.+ e^{iwy_0} \sum_{j=0}^{N_0-1} e^{iux_j} f(x_j, y_0) + e^{iwy_{N_1-1}} \sum_{j=0}^{N_0-1} e^{iux_j} f(x_j, y_{N_1-1})\right)\\
&+ \frac{1}{4}\left(e^{iux_0}e^{iwy_0} f(x_0, y_0) + e^{iux_0}e^{iwy_{N_1-1}} f(x_0, y_{N_1-1})\right.\\
&\left.+ e^{iux_{N_0-1}}e^{iwy_0} f(x_{N_0-1}, y_0) + e^{iux_{N_0-1}}e^{iwy_{N_1-1}} f(x_{N_0-1}, y_{N_1-1})\right).
\end{aligned}$$

To incorporate the FFT, the discrete Fourier transform is incorporated in the approximation. Define two Fourier grids u_n and w_m such that $u_n = \left(n - \frac{N_1}{2}\right)\Delta_u$ and $w_m = \left(m - \frac{N_1}{2}\right)\Delta_w$ for Δ_u and Δ_w subject to the condition

$$\Delta_x\Delta_u = \frac{2\pi}{N_0}, \text{ and } \Delta_y\Delta_w = \frac{2\pi}{N_1}.$$

For $u_n x_j$ and $y_k w_m$ the same derivation can be made as in the one dimensional case and is given in Equation (D.23). Substituting the expression in Equation (D.23) for $u_n x_j$ and $y_k w_m$ into Equation (D.25)

and using the Euler identity yields,

$$\begin{aligned}
\frac{\phi(u_n,w_m)}{\Delta_x\Delta_y} &\approx (-1)^{n+m+\frac{N_1+N_0}{2}}\bar{D}_{n,m}\left(f(x_j,y_k)\right)\\
&\quad -\frac{1}{2}\Big((-1)^{m+\frac{N_1}{2}}e^{iux_0}\bar{D}_m\left(f(x_0,y_k)\right)+(-1)^{m+\frac{N_1}{2}}e^{iux_{N_0-1}}\bar{D}_m\left(f(x_{N_0-1},y_k)\right)\\
&\quad +(-1)^{n+\frac{N_0}{2}}e^{iw_my_0}\bar{D}_n\left(f(x_j,y_0)\right)+(-1)^{n+\frac{N_0}{2}}e^{iw_my_{N_1-1}}\bar{D}_n\left(f(x_j,y_{N_1-1})\right)\Big)\\
&\quad +\frac{1}{4}\big(e^{iux_0}e^{iwy_0}f(x_0,y_0)+e^{iux_0}e^{iwy_{N_1-1}}f(x_0,y_{N_1-1})\\
&\quad +e^{iux_{N_0-1}}e^{iwy_0}f(x_{N_0-1},y_0)+e^{iux_{N_0-1}}e^{iwy_{N_1-1}}f(x_{N_0-1},y_{N_1-1})\big)\\
&:= \frac{\bar{\mathrm{F}}_2[f](u_n,w_m)}{\Delta_x\Delta_y},
\end{aligned}$$

where an adjusted discrete Fourier transform is used, which is given by

$$\bar{D}_{n,m}\left(f(x_j,y_k)\right)=D_{n,m}\left(f(x_j,y_k)(-1)^{k+j}\right)=\sum_{j=0}^{N_0-1}\sum_{k=0}^{N_1-1}e^{i2\pi\left(\frac{nj}{N_0}+\frac{mk}{N_1}\right)}f(x_j,y_k)(-1)^{k+j}.$$

Using the same steps in the derivation, also an approximation of the two-dimensional inverse Fourier approximation can be given in terms of the discrete Fourier transform. By following these steps, the corresponding approximation is then given by

$$\begin{aligned}
4\pi^2\frac{f(x_j,y_k)}{\Delta_u\Delta_w} &\approx (-1)^{k+j+\frac{N_0+N_1}{2}}\bar{D}^{-1}_{j,k}\left(\phi(u_n,w_m)\right)\\
&\quad -\frac{1}{2}\Big((-1)^{j+\frac{N_0}{2}}\bar{D}^{-1}_{j}\left(\phi(u_n,w_0)\right)+(-1)^{j+\frac{N_0}{2}}\bar{D}^{-1}_{j}\left(\phi(u_n,w_{N_1-1})\right)\\
&\quad +(-1)^{k+\frac{N_1}{2}}\bar{D}^{-1}_{k}\left(\phi(u_0,w_m)\right)+(-1)^{k+\frac{N_1}{2}}\bar{D}^{-1}_{k}\left(\phi(u_{N_0-1},w_m)\right)\Big)\\
&\quad +\frac{1}{4}\big(e^{-i(u_0x+w_0y)}\phi(u_0,w_0)+e^{-i(u_0x+w_{N_1-1}y)}\phi(u_0,w_{N_1-1})\\
&\quad +e^{-i(u_{N_0-1}x+w_0y)}\phi(u_{N_0-1},w_0)+e^{-i(u_{N_0-1}x+w_{N_1-1}y)}\phi(u_{N_0-1},w_{N_1-1})\big)\\
&:= \frac{4\pi^2}{\Delta_u\Delta_v}\bar{\mathrm{F}}_2^{-1}[\phi](x_j,y_k).
\end{aligned}\tag{D.26}$$

Within Equation (D.26), the adjusted inverse discrete Fourier transform is used, which is given by

$$\bar{D}^{-1}_{j,k}\left(\phi(u_n,w_m)\right)=D^{-1}_{j,k}\left(\phi(u_n,w_m)(-1)^{n+m}\right)=\sum_{n=0}^{N_0-1}\sum_{m=0}^{N_1-1}e^{-i2\pi\left(\frac{nj}{N_0}+\frac{mk}{N_1}\right)}\phi(u_n,w_m)(-1)^{n+m}.$$

□

Higher order Fourier transforms can be obtained by extending the derivation to higher order integrals.

D.8. Proof of Lemma 4.1

The proof of Lemma 4.1 is very similar to the proof of Lemma D.1.

Proof. Recall the time grid $t_l = l\Delta_t$ defined in Section 4.3. The continuation value $C(t_l, X_F(t_l))$ at time t_l of a convertible bond is defined as

$$C(t_l,X_F(t_l))=\mathbb{E}\left[\mathrm{e}^{-\int_{t_l}^{t_{l+1}}r(s)\mathrm{d}s}V(t_{l+1},X_F(t_{l+1}))\;\middle|\;\mathcal{F}_{t_l}\right],\tag{D.27}$$

where $\{r(t)\}_{t\geq 0}$ is the interest rate process, $\{X_F(t)\}_{t\geq 0}$ is the company process under the log-face transform defined in Equation (4.1) and $V(t_{l+1}, X_F(t_{l+1}))$ is the value of the convertible bond at time $t_{l+1} \in (0, T]$ for maturity T. Recall the value of the money-market account $M(t_l)$ at time t_l defined in Definition 1.2,

$$\mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\mathsf{d}s} V(t_{l+1}, X_F(t_{l+1})) \;\middle|\; \mathcal{F}_{t_l}\right] = M(t_l)\mathbb{E}\left[\frac{V(t_{l+1}, X_F(t_{l+1}))}{M(t_{l+1})} \;\middle|\; \mathcal{F}_{t_l}\right].$$

Define the Radon-Nikodym derivative as

$$\frac{\mathsf{d}\mathbb{Q}^{t_{l+1}}}{\mathsf{d}\mathbb{Q}} = \frac{M(t_l)P(t_{l+1}, t_{l+1})}{M(t_{l+1})P(t_l, t_{l+1})}, \tag{D.28}$$

for the t_{l+1}-forward measure $\mathbb{Q}^{t_{l+1}}$ and the value of a zero-coupon bond $P(t_l, t_{l+1})$ at time t_l with maturity t_{l+1}. Substitution of the Radon-Nikodym derivative given in Equation (D.28) into Equation (D.27), yields,

$$\begin{aligned}
M(t_l)\mathbb{E}\left[\frac{V(t_{l+1}, X_F(t_{l+1}))}{M(t_{l+1})} \;\middle|\; \mathcal{F}_{t_l}\right] &= M(t_l)\int_{\mathbb{R}^2} \frac{V(t_{l+1}, X_F(t_{l+1}))}{M(t_{l+1})}\mathsf{d}\mathbb{Q} \\
&= M(t_l)\int_{\mathbb{R}^2} \frac{V(t_{l+1}, X_F(t_{l+1}))}{M(t_{l+1})}\frac{M(t_{l+1})P(t_{l+1}, t_{l+1})}{M(t_l)P(t_l, t_{l+1})}\mathsf{d}\mathbb{Q}^{t_{l+1}} \\
&= \int_{\mathbb{R}^2} V(t_{l+1}, X_F(t_{l+1}))\frac{P(t_l, t_{l+1})}{P(t_{l+1}, t_{l+1})}\mathsf{d}\mathbb{Q}^{t_{l+1}} \\
&= P(t_l, t_{l+1})\mathbb{E}^{t_{l+1}}\left[V(t_{l+1}, X_F(t_{l+1})) \mid \mathcal{F}_{t_l}\right],
\end{aligned}$$

where was used that $P(t_{l+1}, t_{l+1}) = 1$ and that $P(t_l, t_{l+1})$ is $\mathcal{F}_{t_l}$ measurable. □

D.9. Proof of Theorem 4.3

Proof. Assume that the condition

$$f_{X_F}(x|x_l) = f_{X_F}(x - x_l),$$

holds for the discounted density $f_{X_F}(x)$ corresponding to the process $\{X_F(t)\}_{t\geq 0}$. Furthermore, define the variable transform $z_l = x - x_l$. Combining the expression $f_{X_F}(x - x_l)$ with the variable transformation, yields,

$$\begin{aligned}
\mathrm{e}^{r\Delta_t}\mathrm{F}[\bar{C}_l](u) &= \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}x_l u}\bar{C}_l(t_l, x_l)\mathsf{d}x_l \\
&= \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}x_l u}\int_{\mathbb{R}} \bar{V}(t_{l+1}, x)f_{X_F}(x - x_l)\mathsf{d}x\mathsf{d}x_l \\
&= \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}x_l u}\int_{\mathbb{R}} \bar{V}(t_{l+1}, z_l + x_l)f_{X_F}(z_l)\mathsf{d}z_l\mathsf{d}x_l \\
&= \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}(x - z_l)u}\int_{\mathbb{R}} \bar{V}(t_{l+1}, x)f_{X_F}(z_l)\mathsf{d}z_l\mathsf{d}x \\
&= \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}xu}\bar{V}(t_{l+1}, x)\mathsf{d}x\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}z_l u}f_{X_F}(z_l)\mathsf{d}z_l \\
&= \mathrm{F}[\bar{V}_{l+1}](u)\phi_{X_F}(-u).
\end{aligned}$$

□

D.10. Proof of Theorem 4.4

Proof. Writing the continuation value of the corporate bond under the transformation defined in Equation (4.3) with respect to the log-face form yields [3],

$$\begin{aligned}\frac{C(t_l,\tilde{r}(t_l),X_F(t_l))}{P(t_l,t_{l+1})} &= \mathbb{E}^{t_{l+1}}\left[V(t_{l+1},\tilde{r}(t_{l+1}),X_F(t_{l+1}))\mid \mathcal{F}_{t_l}\right]\\ &= \mathbb{E}^{t_{l+1}}\left[V(t_{l+1},Z_r+g_r(\tilde{r}(t_{l+1})),Z+g(\tilde{r}(t_l),X_F(t_l)))\mid \mathcal{F}_{t_l}\right]\\ &= \int_{\mathbb{R}^2} V(t_{l+1},z_r+g_r(\tilde{r}(t_l)),z+g(\tilde{r}(t_l),X_F(t_l)))f_{Z_r,Z}(z,z_r)\mathsf{d}(z_r,z)\\ &=: \frac{C(t_l,g_r(\tilde{r}(t_l)),g(\tilde{r}(t_l),X_F(t_l)))}{P(t_l,t_{l+1})}.\end{aligned}$$

□

The transformed continuation value forms a convolution between the future bond value and the transition density of the transformed variables Z and Z_r. The formed convolution can be computed in Fourier space and the corresponding continuation value can therefore be recovered using the inverse Fourier transform.

D.11. Proof of Theorem 4.5

Proof. Recall the results of Theorem 4.4, which will be given in shorthand notation,

$$C(t_l,x,y) := \int_{\mathbb{R}^2} V(t_{l+1},z_r+x,z+y)f_{Z_r,Z}(z,z_r)\mathsf{d}(z_r,z).$$

Denote $\bar{V}_l \equiv \bar{V}(t_l,x,y)$. Indeed, if considered $x_0 = z_r + x$ and $y_0 = z + y$,

$$\begin{aligned}\frac{\mathrm{F}_2[C_l](u,w)}{P(t_l,t_{l+1})} &= \frac{\int_{\mathbb{R}^2}\mathrm{e}^{\mathrm{i}(ux+wy)}C(t_l,x,y)\mathsf{d}(x,y)}{P(t_l,t_{l+1})}\\ &= \int_{\mathbb{R}^2}\mathrm{e}^{\mathrm{i}(ux+wy)}\int_{\mathbb{R}^2}\bar{V}_{\mathsf{conv}}(t_{l+1},z_r+x,z+y)f_{Z_r,Z}(z_r,z)\mathsf{d}(z_r,z)\mathsf{d}(x,y)\\ &= \int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\mathrm{e}^{\mathrm{i}(u(x_0-z_r)+w(y_0-z))}\bar{V}_{\mathsf{conv}}(t_{l+1},x_0,y_0)f_{Z_r,Z}(z_r,z)\mathsf{d}(z_r,z)\mathsf{d}(x_0,y_0)\\ &= \int_{\mathbb{R}^2}\mathrm{e}^{\mathrm{i}(ux_0+wy_0)}\bar{V}_{\mathsf{conv}}(t_{l+1},x_0,y_0)\mathsf{d}(x_0,y_0)\int_{\mathbb{R}^2}\mathrm{e}^{-\mathrm{i}(uz_r+wz)}f_{Z_r,Z}(z_r,z)\mathsf{d}(z_r,z)\\ &= \mathrm{F}_2[\bar{V}_{l+1}](u,w)\phi_{Z_r,Z}(-u,-w),\end{aligned}$$

where $\phi_{Z_r,Z}(u,w) = \phi_{Z_r,Z}(u,w,t,T)$ is the characteristic function corresponding to the stochastic variables (Z_r,Z). □

D.12. Proof of Lemma 4.4

Proof. Recall the definition for the continuation coefficients C_k,

$$C_k(t_l,x_1,x_2) = \frac{2}{b-a}\int_{x_1}^{x_2} C(t_l,x)\cos\left(k\pi\frac{y-a}{b-a}\right)\mathsf{d}y, \tag{D.29}$$

where the continuation values can be approximated by

$$\bar{C}(t_l,X_F(t_l)) \approx \mathrm{e}^{-r\Delta_t}{\sum_{k=0}^{N-1}}'\,\mathsf{Re}\left[\phi_X\left(\frac{k\pi}{b-a}\right)\mathrm{e}^{\mathrm{i}k\pi\frac{x-a}{b-a}}\right]V_k(t_{l+1}), \tag{D.30}$$

following Theorem 4.6. Substitution of Equation (D.30) into Equation (D.29) and interchanging sum and integral yields,

$$C_k(t_l,x_1,x_2) \approx \mathrm{e}^{-r\Delta_t}{\sum_{j=0}^{N-1}}'\,\mathsf{Re}\left[\phi_X\left(\frac{j\pi}{b-a}\right)V_j(t_{l+1})\mathcal{A}_{j,k}(t_l,x_1,x_2)\right],$$

where

$$\mathcal{A}_{j,k}(t_l, x_1, x_2) = \frac{2}{(b-a)} \int_{x_1}^{x_2} \mathrm{e}^{\mathrm{i}j\pi \frac{x-a}{b-a}} \cos\left(k\pi \frac{x-a}{b-a}\right) \mathrm{d}x.$$

In particular, the complex identity,

$$\mathrm{e}^{\mathrm{i}x} = \cos(x) + \mathrm{i}\sin(x), \tag{D.31}$$

holds for $x \in \mathbb{R}$. Hence,

$$\mathcal{A}_{k,j}(t_l, x_1, x_2) = \frac{2}{(b-a)} \int_{x_1}^{x_2} \cos\left(j\pi \frac{x-a}{b-a}\right) \cos\left(k\pi \frac{x-a}{b-a}\right) + \mathrm{i}\sin\left(j\pi \frac{x-a}{b-a}\right) \cos\left(k\pi \frac{x-a}{b-a}\right) \mathrm{d}x. \tag{D.32}$$

Recall the cosine and sine multiplication rules;

$$\cos\theta\cos\varphi = \frac{\cos(\theta-\varphi)+\cos(\theta+\varphi)}{2}, \text{ and } \sin\theta\cos\varphi = \frac{\sin(\theta+\varphi)+\sin(\theta-\varphi)}{2}.$$

Combining the rules for the cosine and sine multiplication and the integral given in Equation (D.32), yield,

$$\mathcal{A}_{k,j}(t_l, x_1, x_2) = \frac{1}{(b-a)} \int_{x_1}^{x_2} \cos\left((j-k)\pi \frac{x-a}{b-a}\right) + \mathrm{i}\sin\left((j-k)\pi \frac{x-a}{b-a}\right) + \cos\left((j+k)\pi \frac{x-a}{b-a}\right) + \mathrm{i}\sin\left((j+k)\pi \frac{x-a}{b-a}\right) \mathrm{d}x.$$

Using the identity in Equation (D.31) and computing the integrals yields,

$$\mathcal{A}_{k,j}(t_l, x_1, x_2) = -\frac{\mathrm{i}}{\pi}\left(\tilde{\mathcal{A}}^{+}_{j,k}(t_l, x_1, x_2) + \tilde{\mathcal{A}}^{-}_{j,k}(t_l, x_1, x_2)\right),$$

where

$$\tilde{\mathcal{A}}^{-}_{k,j}(t_l, x_1, x_2) = \begin{cases} \dfrac{x_2 - x_1}{b-a}\mathrm{i}\pi & j = k, \\ \dfrac{\mathrm{e}^{\mathrm{i}(j-k)\pi\frac{x_2-a}{b-a}} - \mathrm{e}^{\mathrm{i}(j-k)\pi\frac{x_1-a}{b-a}}}{(j-k)} & j \neq k, \end{cases}$$

and

$$\tilde{\mathcal{A}}^{+}_{k,j}(t_{M-2}, x_1, x_2) = \begin{cases} \dfrac{x_2 - x_1}{b-a}\mathrm{i}\pi & j = k = 0, \\ \dfrac{\mathrm{e}^{\mathrm{i}(j+k)\pi\frac{x_2-a}{b-a}} - \mathrm{e}^{\mathrm{i}(j+k)\pi\frac{x_1-a}{b-a}}}{(j+k)} & \text{otherwise.} \end{cases}$$

Note that in particular the matrices $\tilde{\mathcal{A}}^{\pm}_{j,k}(t_{M-2}, x_1, x_2)$ can be written in a simple way, namely,

$$\tilde{\mathcal{A}}^{-}_{j}(t_{M-2}, x_1, x_2) = \begin{bmatrix} \mathrm{A}_0 & \mathrm{A}_1 & \cdots & \mathrm{A}_{N-1} \\ \mathrm{A}_{-1} & \mathrm{A}_0 & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathrm{A}_{1-N} & \cdots & & \mathrm{A}_0 \end{bmatrix}, \; \tilde{\mathcal{A}}^{+}_{j}(t_{M-2}, x_1, x_2) = \begin{bmatrix} \mathrm{A}_0 & \mathrm{A}_1 & \cdots & \mathrm{A}_{N-1} \\ \mathrm{A}_1 & \mathrm{A}_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathrm{A}_{N-1} & \cdots & & \mathrm{A}_{2N+2} \end{bmatrix},$$

where,

$$\mathrm{A}_j(t_{M-2}, x_1, x_2) = \begin{cases} \dfrac{x_2 - x_1}{b-a}\mathrm{i}\pi & j = 0, \\ \dfrac{\mathrm{e}^{\mathrm{i}j\pi\frac{x_2-a}{b-a}} - \mathrm{e}^{\mathrm{i}j\pi\frac{x_1-a}{b-a}}}{j} & j \neq 0, \end{cases}$$

which concludes the proof. □

D.13. Proof Theorem 4.8

Proof. Recall the definition of $\bar{H}(T,x)$,

$$\bar{H}(T,x) = \begin{cases} \mathrm{e}^x & \text{for } x < \log(1+\alpha), \\ \max(1+\alpha, \gamma \mathrm{e}^x) & \text{for } x \geq \log(1+\alpha). \end{cases}$$

Furthermore, for the payoff in case of a non-default event, it holds that

$$\max(1+\alpha, \gamma \mathrm{e}^x) = \begin{cases} 1+\alpha & \text{for } x < \log\left(\frac{1+\alpha}{\gamma}\right) \\ \mathrm{e}^x & \text{for } x \geq \log\left(\frac{1+\alpha}{\gamma}\right). \end{cases}$$

Therefore, the total domain $[a,b]$ can be divided into three subdomains on which the payoff functions hold,

$$\underbrace{[a, \log(1+\alpha)]}_{\mathrm{e}^x} \cup \underbrace{\left[\log(1+\alpha), \log\left(\frac{1+\alpha}{\gamma}\right)\right]}_{1+\alpha} \cup \underbrace{\left[\log\left(\frac{1+\alpha}{\gamma}\right), b\right]}_{\gamma \mathrm{e}^x} = [a,b].$$

The Integral

$$V_k(T) = \frac{2}{b-a} \int_a^b \bar{H}(T,x) \cos\left(k\pi \frac{x-a}{b-a}\right) \mathrm{d}x,$$

can thus be divided into the sub-integrals,

$$V_k(T) = \frac{2}{b-a} \left[\int_a^{\log(1+\alpha)} \mathrm{e}^x \cos\left(k\pi \frac{x-a}{b-a}\right) \mathrm{d}x + (1+\alpha) \int_{\log(1+\alpha)}^{\log(\frac{1+\alpha}{\gamma})} \cos\left(k\pi \frac{x-a}{b-a}\right) \mathrm{d}x \right.$$
$$\left. + \gamma \int_{\log(\frac{1+\alpha}{\gamma})}^{b} \mathrm{e}^x \cos\left(k\pi \frac{x-a}{b-a}\right) \mathrm{d}x \right]$$

Applying Lemma C.1 yields directly,

$$V_k(T) = \frac{2}{b-a} \left[\chi_k(a, \log(1+\alpha)) + (1+\alpha)\psi_k\left(\log(1+\alpha), \log\left(\frac{1+\alpha}{\gamma}\right)\right) \right.$$
$$\left. + \gamma \chi_k\left(\log\left(\frac{1+\alpha}{\gamma}\right), b\right) \right],$$

which concludes the proof. □

D.14. Proof Lemma 4.6

Proof. The coefficients $G_k^{\mathsf{conv}}(t_l, x_1, x_2)$ follow directly from the definition of χ_k given in Lemma C.1. The scaled payoff function of the callable convertible option, given that the bond is called, is given by,

$$\bar{H}_{\mathsf{call}}(t_l, x) = \max(\beta(t_l) + \alpha, \gamma \mathrm{e}^x).$$

The bondholder will choose a forced conversion if,

$$\beta(t_l) + \alpha \leq \gamma \mathrm{e}^x \iff x > \log\left(\frac{\beta(t_l)+\alpha}{\gamma}\right) = d.$$

Hence the coefficients corresponding to the call payoff are given by,

$$G_k^{\mathsf{call}}(t_l, x_1, x_2) = \frac{2\alpha}{b-a} \int_{x_1}^{d} \beta(t) + \alpha \cos\left(k\pi \frac{y-a}{b-a}\right) \mathrm{d}y + \int_d^{x_2} \gamma \mathrm{e}^y \cos\left(k\pi \frac{y-a}{b-a}\right) \mathrm{d}y.$$

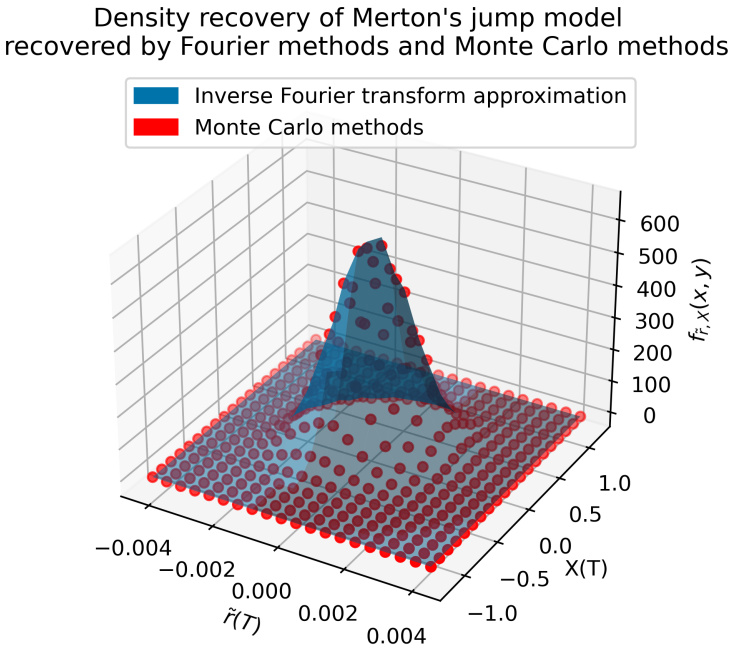
Note that in the case that $x_1 > d$, the first integral is zero and only the second integral is considered overt the total interval $[x_1, x_2]$. Applying Lemma C.1 completes the proof. □

E

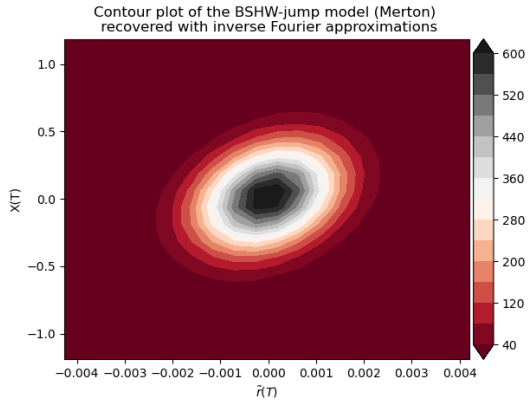
Figures

E.1. Black-Scholes Hull-White density under Merton's jump model

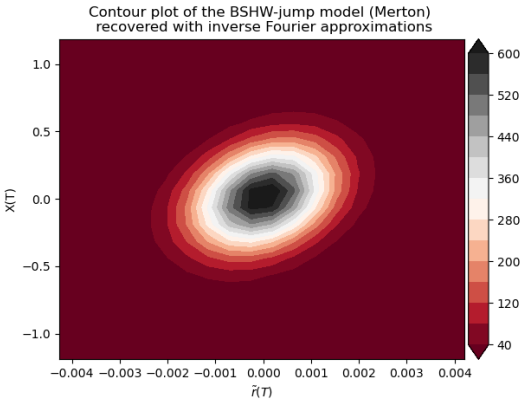
The figures below show the numerical results for the density recovery of the Black-Scholes Hull-White (BSHW) model under Fourier and simulation techniques. The BSHW model is extended with discontinuous jumps, which are modeled using the Merton jump-diffusion model [29].



(a) A three-dimensional plot of the BSHW model, extended with jumps following the jump-diffusion model proposed by Merton. The red dots represent the results, binned, using simulation. The surface corresponds to the results obtained using Fourier techniques.



(b) Contour plot of the numerical density recovery of the BSHW model extended with the jump model proposed by Merton. The results are obtained by using the approximation for the inverse Fourier transform.



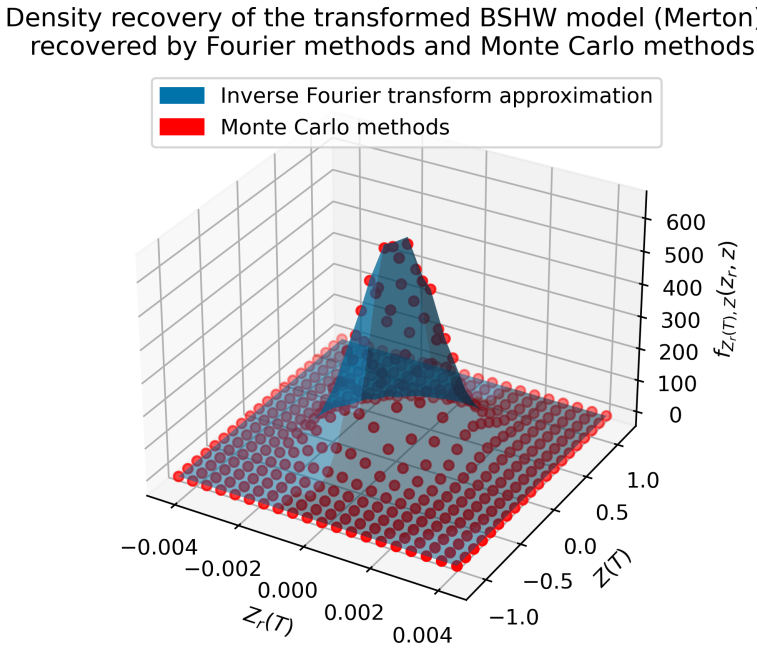
(c) Contour plot of the numerical density recovery of the BSHW model extended with the jump model proposed by Merton. The results are obtained by using simulation as proposed in Chapter 3.

Figure E.1: Different plots that show the result of the numerical approximation of the BSHW model extended with discontinuous jumps as proposed by Merton [29]. The model parameters were set to: $S_0 = 1, \sigma = 0.25, \rho = 0.3, \eta = 0.001, \kappa = 0.05$ and $T = 1$. Furthermore, the jump parameters were $\lambda = 3, \mu_J = -0.03$ and $\sigma_J = 0.03$. The hyper-parameters concerning the Fourier approximation were chosen: $N = 2^9 = 512$ and $L_x = L_y = 10$. The simulations were obtained by simulating $M = 500000$ paths, each containing $N = 50$ nodes.

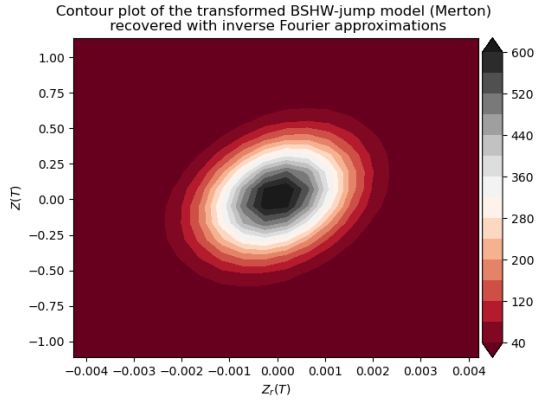
From Figure E.1 it can be seen that the BSHW density is sufficiently captured by the Fourier approximation.

E.2. Transformed Black-Scholes Hull-White density under Merton's jump model

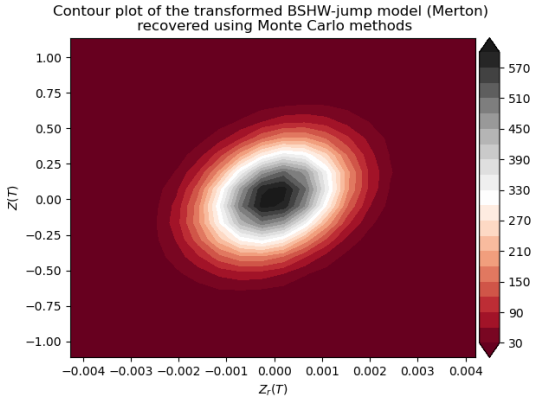
The figures below show the numerical results for the density recovery of the transformed Black-Scholes Hull-White (BSHW) model under Fourier and simulation techniques. The transformation used is described in Section 5.2. The BSHW model is extended with discontinuous jumps, which are modeled using the Merton jump-diffusion model [29]. From Figure E.2 it can be seen that the transformed BSHW density is sufficiently captured by the Fourier approximation.



(a) A three-dimensional plot of the transformed BSHW model, extended with jumps following the jump-diffusion model proposed by Merton. The red dots represent the results, binned, using simulation. The surface corresponds to the results obtained using Fourier techniques.



(b) Contour plot of the numerical density recovery of the transformed BSHW model extended with the jump model proposed by Merton. The results are obtained by using the approximation for the inverse Fourier transform.



(c) Contour plot of the numerical density recovery of the transformed BSHW model extended with the jump model proposed by Merton. The results are obtained by using simulation as proposed in Chapter 3.

Figure E.2: Different plots that show the result of the numerical approximation of the transformed BSHW model extended with discontinuous jumps as proposed by Merton [29]. The model parameters were set to: $S_0 = 1, \sigma = 0.25, \rho = 0.3, \eta = 0.001, \kappa = 0.05$ and $T = 1$. Furthermore, the jump parameters were $\lambda = 3, \mu_J = -0.03$ and $\sigma_J = 0.03$. The hyper-parameters concerning the Fourier approximation were chosen: $N = 2^9 = 512$ and $L_x = L_y = 10$. The simulations were obtained by simulating $M = 500000$ paths, each containing $N = 50$ nodes.

Bibliography

[1] Manuel Ammann, Axel Kind, and Christian Wilde. “Simulation-based pricing of convertible bonds”. In: *Journal of empirical finance* 15.2 (2008), pp. 310–331.

[2] Elie Ayache, Peter A Forsyth, and Kenneth R Vetzal. “Valuation of convertible bonds with credit risk”. In: *The journal of Derivatives* 11.1 (2003), pp. 9–29.

[3] Laura Ballotta and Ioannis Kyriakou. “Convertible bond valuation in a jump diffusion setting with stochastic interest rates”. In: *Quantitative Finance* 15.1 (2015), pp. 115–129.

[4] Jonathan A Batten, Karren Lee-Hwei Khaw, and Martin R Young. “Pricing convertible bonds”. In: *Journal of Banking & Finance* 92 (2018), pp. 216–236.

[5] Bo Becker, Murillo Campello, Viktor Thell, and Dong Yan. “Credit Risk and the Life Cycle of Callable Bonds: Implications for Corporate Financing and Investing”. In: (2021).

[6] Andre Biere and Matthias A Scherer. “Robust calibration of a structural-default model with jumps”. In: *Available at SSRN 1342457* (2009).

[7] Michael J Brennan and Eduardo S Schwartz. “Analyzing convertible bonds”. In: *Journal of Financial and Quantitative analysis* 15.4 (1980), pp. 907–929.

[8] Damiano Brigo and Fabio Mercurio. *Interest rate models-theory and practice: with smile, inflation and credit*. Vol. 2. Springer, 2006.

[9] Carmen M Calvo-Garrido, Sidi Diop, Andrea Pascucci, and Carlos Vázquez. “PDE models for the pricing of a defaultable coupon-bearing bond under an extended JDCEV model”. In: *Communications in Nonlinear Science and Numerical Simulation* 102 (2021), p. 105914.

[10] Peter Carr and Dilip Madan. “Option valuation using the fast Fourier transform”. In: *Journal of computational finance* 2.4 (1999), pp. 61–73.

[11] Darrell Duffie, Jun Pan, and Kenneth Singleton. “Transform analysis and asset pricing for affine jump-diffusions”. In: *Econometrica* 68.6 (2000), pp. 1343–1376.

[12] Darrell Duffie and Kenneth J Singleton. “Modeling term structures of defaultable bonds”. In: *The review of financial studies* 12.4 (1999), pp. 687–720.

[13] Frank J Fabozzi and Francesco A Fabozzi. *Bond markets, analysis, and strategies*. 9th ed. Pearson, 2014.

[14] Fang Fang. “The COS method: An efficient Fourier method for pricing financial derivatives”. In: (2010).

[15] Fang Fang, Henrik Jönsson, Cornelis W Oosterlee, and Wim Schoutens. “Fast valuation and calibration of credit default swaps under Lévy dynamics”. In: *Available at SSRN 1628672* (2009).

[16] Santiago Forte. “Calibrating structural models: a new methodology based on stock and credit default swap data”. In: *Quantitative Finance* 11.12 (2011), pp. 1745–1759.

[17] Lech A Grzelak and Cornelis W Oosterlee. “An equity-interest rate hybrid model with stochastic volatility and the interest rate smile”. In: *The Journal of Computational Finance (1–33) Volume* 15 (2010).

[18] Lech A Grzelak, Cornelis W Oosterlee, and Sacha Van Weeren. “Extension of stochastic volatility equity models with the Hull–White interest rate process”. In: *Quantitative Finance* 12.1 (2012), pp. 89–105.

[19] David Heath, Robert Jarrow, and Andrew Morton. “Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation”. In: *Econometrica: Journal of the Econometric Society* (1992), pp. 77–105.

[20] Jing-Zhi Huang and Ming Huang. “How much of the corporate-treasury yield spread is due to credit risk?” In: *The Review of Asset Pricing Studies* 2.2 (2012), pp. 153–202.

[21] Jonathan E Ingersoll Jr. “A contingent-claims valuation of convertible securities”. In: *Journal of Financial Economics* 4.3 (1977), pp. 289–321.

[22] Steven G Kou. “A jump-diffusion model for option pricing”. In: *Management science* 48.8 (2002), pp. 1086–1101.

[23] Steven G Kou and Hui Wang. “Option pricing under a double exponential jump diffusion model”. In: *Management science* 50.9 (2004), pp. 1178–1192.

[24] Coenraad CW Leentvaar and Cornelis W Oosterlee. “Multi-asset option pricing using a parallel Fourier-based technique”. In: *Reports of the Department of Applied Mathematical Analysis, 07-12* (2007).

[25] Xinting Li, Baochen Yang, Yunpeng Su, and Yunbi An. “Pricing Corporate Bonds with Credit Risk, Liquidity Risk, and Their Correlation”. In: *Discrete Dynamics in Nature and Society* 2021 (2021).

[26] Francis A Longstaff and Eduardo S Schwartz. “Valuing American options by simulation: a simple least-squares approach”. In: *The review of financial studies* 14.1 (2001), pp. 113–147.

[27] Roger Lord, Fang Fang, Frank Bervoets, and Cornelis W Oosterlee. “A fast and accurate FFT-based method for pricing early-exercise options under Lévy processes”. In: *SIAM Journal on Scientific Computing* 30.4 (2008), pp. 1678–1705.

[28] Alexander J McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative risk management: concepts, techniques and tools-revised edition*. Princeton university press, 2015.

[29] Robert C Merton. “On the pricing of corporate debt: The risk structure of interest rates”. In: *The Journal of finance* 29.2 (1974), pp. 449–470.

[30] Robert C Merton. “Option pricing when underlying stock returns are discontinuous”. In: *Journal of financial economics* 3.1-2 (1976), pp. 125–144.

[31] Franco Modigliani and Merton H Miller. “The cost of capital, corporation finance and the theory of investment”. In: *The American economic review* 48.3 (1958), pp. 261–297.

[32] Cornelis W Oosterlee and Lech A Grzelak. *Mathematical Modeling and Computation in Finance: With Exercises and Python and Matlab Computer Codes*. World Scientific, 2019.

[33] Marjon J Ruijter and Cornelis W Oosterlee. “Two-dimensional Fourier cosine series expansion method for pricing financial options”. In: *SIAM Journal on Scientific Computing* 34.5 (2012), B642–B671.

[34] Stephen M Schaefer and Ilya A Strebulaev. “Structural models of credit risk are useful: Evidence from hedge ratios on corporate bonds”. In: *Journal of Financial Economics* 90.1 (2008), pp. 1–19.

[35] Steven E Shreve. *Stochastic calculus for finance II: Continuous-time models*. Vol. 11. Springer Science & Business Media, 2004.

[36] Kenneth J Singleton and Len Umantsev. “Pricing coupon-bond options and swaptions in affine term structure models”. In: *Mathematical Finance* 12.4 (2002), pp. 427–446.

[37] Zhiqiang Zhang, Zhenfang Wang, and Xiaowei Chen. “Pricing convertible bond in uncertain financial market”. In: *Journal of Uncertain Systems* 14.01 (2021), p. 2150007.